

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«Российский государственный гуманитарный университет»
(ФГБОУ ВО «РГГУ»)**

ИНСТИТУТ ЛИНГВИСТИКИ

Кафедра теоретической и прикладной лингвистики

Компьютерная лексикография

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Направление подготовки № 45.05.01 Перевод и переводоведение

Специализация: Лингвистическое обеспечение межгосударственных отношений

Уровень квалификации выпускника *специалист*

Форма обучения *очная*

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2019

Общая и компьютерная лексикография

Рабочая программа дисциплины

Составитель:

к. филол. наук, доцент *С.Ю. Семенова*

УТВЕРЖДЕНО

Протокол заседания кафедры ТиПЛ

№1 _____ от 30/08/2019г. _____

ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины (*модуля*)

1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине (*модулю*)

1.3. Место дисциплины в структуре образовательной программы

2. Структура дисциплины (*модуля*)

3. Содержание дисциплины (*модуля*)

4. Образовательные технологии

5. Оценка планируемых результатов обучения

5.1. Система оценивания

5.2. Критерии выставления оценок

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине (*модулю*)

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

7. Материально-техническое обеспечение дисциплины (*модуля*)

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений

1. Пояснительная записка

1.1. Цель и задачи дисциплины

Цель дисциплины – ознакомить студентов с наиболее значимыми идеями, направлениями и разработками в области компьютерной лексикографии.

Задачи дисциплины:

- сформировать у студентов представление о традиционной лексикографии как об одной из основных (наряду с грамматикой) составляющих описания естественного языка;
- ознакомить студентов с терминологией традиционной лексикографии и дать представление о типологии традиционных словарей;
- создать у студентов представление о компьютерной лексикографии как области на стыке традиционной лексикографии и компьютерной лингвистики;
- охарактеризовать основные направления, сложившиеся в компьютерной лексикографии;
- познакомить студентов с историей возникновения и развития компьютерной лексикографии за рубежом и в нашей стране;
- сформировать у студентов представления о компьютерных технологиях, полезных для автоматизации словарного дела, в том числе о базах данных (об их основных логических моделях), а также о гипертекстовых и мультимедийных технологиях и путях их применения в электронных лексикографических системах;
- привить простейшие навыки использования корпусного материала при определении состава словарной информации для конкретных лексических единиц;
- познакомить студентов с проблематикой и конкретными разработками в области терминографии и тезаурусной лексикографии;
- сформировать представления о видах словарной информации, необходимой для автоматической обработки текста, и о способах формализации такой информации в конкретных разработках;
- раскрыть перед студентами пути построения информационных систем, аналогичных словарным, для систематизации теоретических сведений в разных областях лингвистики (в семантике, типологии и др.).
- научить студентов проектировать словарные базы данных и терминологические банки данных несложной структуры;
- научить студентов пользоваться учебной и научной литературой по компьютерной лексикографии, в том числе публикациями в трудах научных конференций.

1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине (*модулю*):

Коды компетенции	Содержание компетенций	Перечень планируемых результатов обучения по дисциплине
ПК-8	способность применять методику ориентированного поиска информации в справочной, специальной литературе и	Знать: <ul style="list-style-type: none"> • основные понятия и термины традиционной

	компьютерных сетях	<p>лексикографии;</p> <ul style="list-style-type: none"> • основные понятия и термины информатики (в том числе области «Базы данных»), связанные с автоматизацией словарного дела; <p>Уметь:</p> <ul style="list-style-type: none"> • использовать корпусный материал для выявления различных типов словарной информации; • пользоваться учебной литературой по традиционной лексикографии и по компьютерной лингвистике; <p>Владеть:</p> <ul style="list-style-type: none"> • простейшими навыками структурирования лексикографической информации; • простейшими навыками расстановки тезаурусных связей.
ПК-15	способностью к обобщению, критическому осмыслению, систематизации информации, анализу логики рассуждений и высказываний	<p>Знать:</p> <ul style="list-style-type: none"> • основные типы информации, необходимой для автоматической обработки текста; • основные способы применения гипертекстовых и мультимедийных технологий в компьютерной лексикографии, нашедшие реализацию в современных системах; <p>Уметь:</p> <ul style="list-style-type: none"> • проектировать структуру простейших баз и банков данных для хранения лексикографической и терминографической информации; • читать и анализировать

		<p>относительно несложные научные публикации по компьютерной лексикографии;</p> <p>Владеть:</p> <ul style="list-style-type: none"> • методикой ориентированного поиска информации в справочной, специальной литературе и компьютерных сетях
ПК-18	<p>способностью применять методы научных исследований в профессиональной деятельности, анализировать материалы исследований в области лингвистики, межкультурной коммуникации и переводоведения с соблюдением библиографической культуры для решения профессиональных задач</p>	<p>Знать:</p> <ul style="list-style-type: none"> • основные понятия и термины традиционной лексикографии; • основные понятия и термины информатики (в том числе области «Базы данных»), связанные с автоматизацией словарного дела; <p>Уметь:</p> <ul style="list-style-type: none"> • использовать корпусный материал для выявления различных типов словарной информации; • пользоваться учебной литературой по традиционной лексикографии и по компьютерной лингвистике; <p>Владеть:</p> <ul style="list-style-type: none"> • простейшими навыками структурирования лексикографической информации; • простейшими навыками расстановки тезаурусных связей.

1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Компьютерная лексикография» относится к вариативной части блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин: «Информатика и информационные технологии в лингвистике», «Лингвистические основы информационного поиска».

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик: «Организация переводческого дела и переводческих проектов в современном мире», «Основы реферирования текстов на первом иностранном языке». Знания, полученные в процессе освоения курса, будут затем развиты при изучении курсов лексикологии и стилистики первого иностранного языка, основ теории второго иностранного языка.

2. Структура дисциплины

Структура дисциплины (модуля) для очной формы обучения

Общая трудоёмкость дисциплины составляет 2 з.е., 72 ч., в том числе контактная работа обучающихся с преподавателем 28 ч., самостоятельная работа обучающихся 44 ч.

№ п/п	Раздел дисциплины/темы	Семестр	Виды учебной работы (в часах)						Формы текущего контроля успеваемости, форма промежуточной аттестации (<i>по семестрам</i>)
			Контактная				Промежуточная аттестация	Самостоятель- ная работа	
			Лекции	Семинар	Практические занятия	Лабораторные занятия			
	Тема 1. Вводная часть курса. Обзор проблематики традиционной лексикографии	6	1					2	Устный ответ
	Тема 2. Понятие и основные направления компьютерной лексикографии	6	1					2	Устный ответ
	Тема 3. История компьютерной лексикографии	6	2					4	Устный ответ
	Тема 4. Общая и компьютерная лексикография и информационные технологии	6	2					4	Устный ответ
	Тема 5. Мультимедийные словари	6	1		1			2	Устный ответ
	Тема 6. Терминологическое банки данных	6	1					2	Устный ответ
	Тема 7. Лексическое наполнение терминологического банка данных	6	1		2			4	Устный ответ
	Тема 8. Тезаурусные терминологические системы	6	1					2	Устный ответ
	Тема 9. Толкование термина	6			1			2	Устный ответ

	Тема 10. Системы лингвистической терминологии	6	1				2	Устный ответ
	Тема 11. Общая и компьютерная лексикография корпусная лингвистика	6			1		2	Устный ответ
	Тема 12. Словарные ресурсы для автоматической обработки текста	6			3		4	Устный ответ
	Тема 13. Семантические словари для автоматической обработки текста	6			4		4	Устный ответ
	Тема 14. Информационные системы для теоретических исследований в лингвистике	6	2		4		4	Устный ответ
	<i>Зачёт</i>						2	<i>Устный ответ на контрольные вопросы</i>
	итоги:		12		16		44	

3. Содержание дисциплины

Тема 1. Вводная часть курса. Обзор проблематики традиционной лексикографии

Традиционная лексикография как раздел языкознания. Лексикография теоретическая и практическая. Проблематика теоретической лексикографии. Социальные функции, выполняемые практической лексикографией. Принципы классификации и типы традиционных словарей. Основные понятия и термины традиционной лексикографии: словарная статья, словник, зона словарной статьи, толкование, помета, метаязык, текстовая иллюстрация.

Тема 2. Понятие и основные направления компьютерной лексикографии

Определение компьютерной лексикографии. Основные типы компьютерных лексикографических систем, их потенциальные пользователи. Словарные продукты, традиционно получаемые в автоматизированном режиме при анализе текста: частотные словари, указатели, конкордансы.

Тема 3. История компьютерной лексикографии

Обзор истории развития вычислительной техники и информационных технологий.

Программное сопровождение машинных словарей. Автоматизация пополнения информационных систем как одно из направлений развития информатики.

История возникновения компьютерной лексикографии. Первые опыты Р.Бузы. Крупные общелексические машинные словари, созданные до начала 90-х годов (до эпохи персональных компьютеров и Интернета). Основные идеи, реализованные в них.

Краткая история машинного перевода. Роль словарей в системах МП. Создание коллекций полнотекстовых документов. Машинные фонды. Машинный фонд русского языка.

Тема 4. Общая и компьютерная лексикография и информационные технологии

Автоматизированное рабочее место (АРМ). АРМ переводчика, лингвиста-теоретика, прикладного лингвиста.

Базы данных. Системы управления базами данных (СУБД). Реляционные, иерархические, сетевые логические модели данных и СУБД. Преимущества словарной базы данных над словарем в текстовом формате. Представление о базах знаний.

Компьютерные сети, история их развития. Интернет и его возможности для развития компьютерной лексикографии. Гипертекстовые технологии и системы.

Понятие семантической сети. Семантическая сеть как инструмент описания полисемии. Лексикографическая информация разных типов в семантической сети. Система WordNet и ее отечественный аналог RussNet.

Тема 5. Мультимедийные словари

Мультимедиа в узком и широком смысле. Возможности применения мультимедиа в словарном деле. Примеры мультимедийных словарей и других мультимедийных лингвистических ресурсов. Роль мультимедиа для документации малых и исчезающих языков. Визуализация как одно из направлений современной информатики.

Тема 6. Терминологические банки данных

Виды терминологических банков данных (ТБД). Особенности ТБД по сравнению с общелексическими машинными словарями. Типовая структура записи в ТБД. Примеры ТБД, созданных до начала 1990-х годов.

Тема 7. Лексическое наполнение терминологического банка данных

Терминоведение и терминография. Свойства термина как лексической единицы особого рода. Семантические процессы в терминологии (терминологизация, детерминологизация, ретерминологизация). Термины и номенклатура. Словообразовательные модели в терминологии.

Описательные и нормативные терминологические системы. Понятия стандартизации, унификации, нормализации и гармонизации терминологии. Стандартизация на международном, государственном, ведомственном уровнях. Терминологические ГОСТы как особый вид словарей. Требования, предъявляемые к стандартизованному термину.

Тема 8. Тезаурусные терминологические системы

Значения полисемичного термина *тезаурус*. Основные типы тезаурусных связей. Общелексические и терминологические тезаурусы. Тезаурус как инструмент информационного поиска. Тезаурусы в библиотечном деле. Предметная рубрика и ее структура. Авторитетный файл предметных рубрик.

Тема 9. Толкование термина

Требования, предъявляемые к толкованию термина (дефиниции). Основные логические типы определений и основные типы и компоненты дефиниций в терминографии.

Тема 10. Системы лингвистической терминологии

Особенности лингвистической терминологии. Задачи исследований в области лингвистической терминологии. Примеры словарей и баз данных лингвистических терминов. Тезаурусы по лингвистике.

Тема 11. Общая и компьютерная лексикография корпусная лингвистика

Понятие и основные параметры корпуса текстов. Примеры корпусов. Морфологическая, синтаксическая, семантическая разметка. Корпус параллельных текстов. Взаимодействие компьютерной лексикографии и корпусной лингвистики, сходство технологий.

Тема 12. Словарные ресурсы для автоматической обработки текста

Классические этапы автоматического анализа текста – графематический, морфологический, синтаксический, семантический. Типы словарной информации для обеспечения этих этапов. Машинные словари основ и окончаний, словоформ, моделей управления, словосочетаний, имен собственных. Этап трансфера в системах машинного перевода и переводные эквиваленты в машинном словаре.

Тема 13. Семантические словари для автоматической обработки текста

Виды семантических машинных словарей (по целям использования). Требования к семантическим словарям, нацеленным на автоматическую обработку текста. Указание семантического класса – основной путь представления смысла слова. Проблемы таксономии в лексикографии.

Словарь РУСЛАН как пример формализованного машинного семантического словаря. Представление об информационно-лингвистической модели (ИЛМ) Н.Н.Леонтьевой как теоретической концепции данного словаря.

Структура и метаязык словаря РУСЛАН. Представление семантической структуры слова. Семантические характеристики и смысловые отношения. Описание ситуаций, обозначаемых предикатными словами. Описание валентностей и лексической сочетаемости. Представление энциклопедической информации. Подбор иллюстративного материала.

Тема 14. Информационные системы для теоретических исследований в лингвистике

Проблема структуризации толкований. Система «Лексикограф» как пример семантической базы знаний. База данных «Русский глагол» системы «Лексикограф». Отражение процессов семантической деривации. Теория лексического значения глагола Е.В. Падучевой как теоретическое основание базы данных. Базы данных предметных имен и наречий в системе «Лексикограф».

Словарь английских глаголов речи А.Вежбицкой как пример «бумажной» семантической базы данных.

Возможности и примеры применения информационных систем в других областях языкознания – в фонетике, диалектологии, фразеологии, типологии, в мониторинге лингвистических исследований.

4. Образовательные технологии

Образовательные технологии

№ п/п	Наименование раздела	Виды учебных занятий	Образовательные технологии
1	2	3	4
2.	Вводная часть курса. Обзор проблематики традиционной лексикографии	Лекция 1. Семинар 1. Самостоятельная работа.	Вводная лекция с использованием видеоматериалов Консультирование и проверка домашних заданий посредством электронной почты
3.	Понятие и основные направления компьютерной лексикографии	Лекция 1. Семинары 1.	Развернутая беседа с показом и обсуждением материала
4.	История компьютерной лексикографии	Лекция 1. Семинар 1.	Развернутая беседа с обсуждением материала
5.	Общая и компьютерная лексикография информационные технологии	Лекция 1. Семинар 1.	Развернутая беседа с обсуждением материала
6.	Мультимедийные словари	Семинары 2. Самостоятельная работа.	Развернутая беседа с обсуждением материала
7.	Терминологические банки данных	Лекция 1. Семинары 2.	Развернутая беседа с обсуждением материала
8.	Лексическое наполнение терминологического банка данных Тезаурусные терминологические системы	Лекция 1. Семинары 2. Самостоятельная работа.	Развернутая беседа с обсуждением материала
9.	Толкование термина	Лекция 1.	Развернутая беседа с

		Семинары 2. Самостоятельная работа.	обсуждением материала
10.	Системы лингвистической терминологии	Лекция 1.	Развернутая беседа с обсуждением материала
11.	Общая и компьютерная лексикография корпусная лингвистика	Лекция 1. Семинар 1.	Развернутая беседа с обсуждением материала
12.	Словарные ресурсы для автоматической обработки текста	Лекция 1. Семинары 2. Самостоятельная работа.	Развернутая беседа с обсуждением материала
13.	Семантические словари для автоматической обработки текста	Лекция 1. Семинары 1. Самостоятельная работа.	Развернутая беседа с обсуждением материала
14.	Информационные системы для теоретических исследований в лингвистике	Лекция 1.	Развернутая беседа с обсуждением материала

5. Оценка планируемых результатов обучения

5.1. Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль:		
- опрос	5 баллов	30 баллов
- участие в дискуссии на семинаре	5 баллов	10 баллов
- тест (темы 1-7)	10 баллов	10 баллов
- тест (темы 8-14)	10 баллов	10 баллов
Промежуточная аттестация		40 баллов
Устный опрос по контрольным вопросам		
Итого за семестр зачёт		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E

20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

5.2. Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ А,В	«отлично»/ «зачтено (отлично)»/ «зачтено»	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	«удовлетворительно»/ «зачтено (удовлетворительно)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	«неудовлетворительно»/ не зачтено	Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

Контрольные вопросы по курсу (ПК-8; ПК-15; ПК-18)

1. Краткая характеристика традиционной лексикографии как раздела общего языкознания. Лексикография теоретическая и практическая. Проблематика теоретической лексикографии. Социальные функции, выполняемые практической лексикографией.
2. Основные понятия традиционной лексикографии: словарная статья, словник, зона словарной статьи, помета, толкование, метаязык, иллюстрация.
3. Основные противоположения, на которых основывается классификация традиционных словарей (словарь академического типа vs. справочник и т.д.). Конкретные виды словарей, примеры.
4. Понятие, направления компьютерной лексикографии.
Основные типы компьютерных лексикографических систем.
5. Электронные коллекции документов. Корпусная лингвистика, ее связь с компьютерной лексикографией. Понятия корпуса, разметки, репрезентативности. Примеры корпусов.
6. Программы поддержки лексикографических работ. Основные функции таких программ.
7. Информационная система. Информационные ресурсы. База данных. База знаний. Понятие системы управления базами данных (СУБД). Преимущества машинного словаря, созданного в виде базы данных, над словарем в текстовом формате.
8. Логические модели баз данных: реляционная, иерархическая, сетевая. Примеры соответствующих СУБД.
9. Представление об истории развития вычислительной техники, поколениях ЭВМ.

10. История возникновения компьютерной лексикографии. Примеры общелексических машинных словарей, созданных до начала 1990-х годов (по выбору – один из словарей). Основные идеи, реализованные в таких словарях.
11. Компьютерные сети, история их развития (кратко). Возможности использования Интернета для развития компьютерной лексикографии.
12. Мультимедиа в узком и широком смысле. Примеры современных мультимедийных словарных систем. Визуализация как направление современной информатики.
13. Терминологические банки данных (ТБД), их виды, особенности по сравнению с общелексическими машинными словарями.
14. Примеры терминологических банков данных.
15. Примеры терминологических тезаурусов.
16. Типовая структура записи терминологического банка данных.
17. Автоматизированное рабочее место (АРМ). АРМ переводчика, лингвиста-теоретика, прикладного лингвиста.
18. Терминоведение и терминография. Определение термина (по А.Н. Баранову). Термины и номенклатура.
19. Свойства термина как особой лексической единицы. Семантические процессы в терминологии (терминологизация, детерминологизация, ретерминологизация).
20. Примеры словообразовательных моделей в терминологии. Крупные семантические категории терминов (процессы, величины, отрасли и др.; примеры терминов, относящихся к этим категориям).
21. Особенности лингвистической терминологии. Направления исследований в области лингвистической терминологии. Примеры словарей лингвистических терминов.
22. Стандартизация, унификация, нормализация и гармонизация терминологии. Уровни стандартизации: международный, государственный, ведомственный. Терминологические ГОСТы как особый вид нормативных словарей.
23. Примеры нормализации терминов для задачи информационного поиска.
24. Требования, предъявляемые к стандартизованному термину.
25. Требования, предъявляемые к толкованию термина (дефиниции).
26. Основные типы дефиниций. Компоненты интенциональной дефиниции.
27. Словарные ресурсы, предназначенные для автоматической обработки текста. Организация морфологической информации в словаре.
28. Представление о стратегиях автоматического синтаксического анализа (кратко). Словарная информация для синтаксического анализа. Модели управления и словосочетания в машинном словаре.
29. Машинный семантический словарь РУСЛАН. Его структура, метаязык (кратко). Представление об информационно-лингвистической модели (ИЛМ) Н.Н. Леонтьевой как теоретической концепции данного словаря.
30. Семантические характеристики слова в словаре РУСЛАН. Моделирование компонентной семантической структуры слова.
31. Описание валентностей слова в словаре РУСЛАН. Смысловые отношения.
32. Энциклопедическая информация в словаре РУСЛАН. Информационный вес слова. Иллюстративный материал в данном словаре.
33. Лексикографические базы знаний, их назначение. Система "Лексикограф" как пример лексикографической базы знаний. Теоретическая концепция данной системы, представление о теории лексического значения глагола Е.В. Падучевой (кратко).
34. Описание глагольной лексики в системе "Лексикограф". Отражение процессов семантической деривации (кратко).
35. Структуризация толкований. Словарь английских глаголов речи А. Вежбицкой как пример "бумажной" семантической базы данных.
36. Описание именной лексики в системе "Лексикограф" (кратко).
37. Описание наречий в системе "Лексикограф" (кратко).

38. Семантические сети. Система WordNet и ее отечественный аналог RusNet.

6. Учебно-методическое и информационное обеспечение дисциплины

6. Список источников и литературы

Основная

- 1.
2. Баранов Анатолий Николаевич.
Введение в прикладную лингвистику / А. Н. Баранов ; МГУ им. М. В. Ломоносова, Филол. фак. - Изд. 3-е. - М. : URSS : ЛКИ, 2007. - 358 с. : схем.,табл. ; 21 см. - (Новый лингвистический учебник). - Библиогр. : с. 326-345. - Указ. : с. 346-354. - ISBN 978-5-382-00231-6 : 217.47.
3. Zubov Александр Васильевич.
Основы искусственного интеллекта для лингвистов : учеб. пособие для студентов вузов, обучающихся по специальности "Теорет. и прикладная лингвистика" / А. В. Zubov, И. И. Зубова. - М. : Логос, 2007. - 319 с. : рис., табл. ; 21 см. - (Новая университетская библиотека). - Библиогр.: с. 291-319 (455 назв.). - ISBN 5-98704-118-X : 198.00.
4. Прикладное языкознание : Учебник / Л.В. Бондарко [и др.] ; под ред. А.С. Герд ; С.-Петербург. гос. ун-т. - : Изд-во С.-Петерб. ун-та, 1996. - 525 с. : рис.,табл. - ISBN 5-288-01369-1 : 23.5.
5. Соссюр, Ф. Курс общей лингвистики / Ф. Соссюр ; переводчик А. М. Сухотин ; под редакцией Р. О. Шор. — Москва : Издательство Юрайт, 2020. — 303 с. — (Антология мысли). — ISBN 978-5-534-05835-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/455043> (дата обращения: 04.04.2020).
6. Шайкевич Анатолий Янович.
Введение в лингвистику : учеб. пособие для студентов вузов, обучающихся по специальности 021700 - Филология / А. Я. Шайкевич. - М. : Академия, 2005. - 393 с. : рис., карты ; 22 см. - (Высшее профессиональное образование. Филология). - Указ.: с. 378-391. - ISBN 5-7695-0930-9 : 179.96.
7. Шайкевич Анатолий Янович.
Введение в лингвистику : учеб. пособие для студентов вузов, обучающихся по специальности "Филология" / А. Я. Шайкевич. - 2-е изд., испр. - М. : Академия, 2010. - 393 с. : рис., табл., карты ; 22 см. - (Высшее профессиональное образование. Филология). - Указ.: с. 378-391. - ISBN 978-5-7695-5829-0 : 582.00.

Дополнительная

1. Апресян Ю.Д. О проекте активного словаря (АС) русского языка // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог» (2008); Бекасово, 4-8 июня 2008 г. Вып.7 (14). – М.: РГГУ, 2008. С. 23-31. <http://www.dialog-21.ru/media/1725/4.pdf/>
2. Большакова Е.И., Большаков И.А. Компьютерный словарь русских паронимов, основанный на формальном критерии паронимии // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2013). Т.1. М.: Изд-во РГГУ, 2013. С. 130-139. <http://www.dialog-21.ru/media/1231/bolshakovaei.pdf/>
3. Иомдин Б.Л., Лопухина А.А., Носырев Г.В. К созданию частотного словаря значений слов // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2014). Вып. 13. М.: Изд-во РГГУ, 2014. С. 199-210. <http://www.dialog-21.ru/digests/dialog2014/materials/pdf/IomdinBL.pdf/>

4. Клыгина Е.А., Крейдлин Г.Е. База данных «Тело и телесность в языке и культуре (идеология, структура и наполнение)» // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 30 мая-3 июня 2012 г.). Вып. 11 (18): В 2 т. Т.1.: Основная программа конференции. М.: Изд-во РГГУ, 2012. С. 256-267. <http://www.dialog-21.ru/media/1334/63.pdf/>
5. Кюсева М.В., Резникова Т.И., Рыжова Д.А. Типологическая база данных адъективной лексики // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2013). Т.1. М.: Изд-во РГГУ, 2013. С. 37-376. <http://www.dialog-21.ru/media/1259/kyusevamv.pdf/>
6. Поляков А.Е., Савчук С.О., Сичинава Д.В. Грамматический словарь для автоматического анализа текстов XVIII—XIX века: первые результаты // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2013). Т.1. М.: Изд-во РГГУ, 2013. С. 574-586. <http://www.dialog-21.ru/media/1282/polyakovaecut.pdf/>
7. Семенова С.Ю. Информация энциклопедического характера в прикладном семантическом словаре // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Междунар. конференции «Диалог 2009». Вып. 8 (15).С. 444-450. <http://www.dialog-21.ru/media/1619/69.pdf/>
8. Филипенко М.В. Наречия в системе «Лексикограф» // Компьютерная лингвистика и интеллектуальные технологии. Тр. междунар. конференции Диалог'2004 («Верхневолжский, 2 — 7 июня 2004 г.»). С. 650-655. <http://www.dialog-21.ru/media/2508/filipenko.pdf/>

7. Материально-техническое обеспечение дисциплины (модуля)

Для обеспечения дисциплины используется материально-техническая база образовательного учреждения: компьютерные классы и научная библиотека РГГУ.

Программное обеспечение

№п /п	Наименование ПО	Производитель	Способ распространения
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
4	AutoCAD 2010 Student	Autodesk	свободно распространяемое
5	Archicad 21 Rus Student	Graphisoft	свободно распространяемое
6	SPSS Statistics 22	IBM	лицензионное
7	Microsoft Share Point 2010	Microsoft	лицензионное
8	SPSS Statistics 25	IBM	лицензионное
9	Microsoft Office 2013	Microsoft	лицензионное
10	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
11	Microsoft Office 2013	Microsoft	лицензионное
12	Windows 10 Pro	Microsoft	лицензионное
13	Kaspersky Endpoint Security	Kaspersky	лицензионное
14	Microsoft Office 2016	Microsoft	лицензионное
15	Visual Studio 2019	Microsoft	лицензионное
16	Adobe Creative Cloud	Adobe	лицензионное

1. Перечень БД и ИСС

№п /п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2019 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2019 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru
4	Компьютерные справочные правовые системы Консультант Плюс, Гарант

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
 - обеспечивается индивидуальное равномерное освещение не менее 300 люкс;
 - для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
 - письменные задания оформляются увеличенным шрифтом;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.
- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;

- экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;
 - в форме аудиофайла.
- для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
 - устройством для сканирования и чтения с камерой SARA CE;
 - дисплеем Брайля PAC Mate 20;
 - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
 - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемые эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

*Приложение 1***АННОТАЦИЯ ДИСЦИПЛИНЫ**

Дисциплина *Общая и компьютерная лексикография* реализуется в Институте лингвистики на факультете Фундаментальной и прикладной лингвистики кафедрой Теоретической и прикладной лингвистики.

Цель дисциплины ознакомить студентов с наиболее значимыми идеями, направлениями и разработками в области компьютерной лексикографии.

Задачи :

- сформировать у студентов представление о традиционной лексикографии как об одной из основных (наряду с грамматикой) составляющих описания естественного языка;
- ознакомить студентов с терминологией традиционной лексикографии и дать представление о типологии традиционных словарей;
- создать у студентов представление о компьютерной лексикографии как области на стыке традиционной лексикографии и компьютерной лингвистики;
- охарактеризовать основные направления, сложившиеся в компьютерной лексикографии;
- познакомить студентов с историей возникновения и развития компьютерной лексикографии за рубежом и в нашей стране;
- сформировать у студентов представления о компьютерных технологиях, полезных для автоматизации словарного дела, в том числе о базах данных (об их основных логических моделях), а также о гипертекстовых и мультимедийных технологиях и путях их применения в электронных лексикографических системах;
- привить простейшие навыки использования корпусного материала при определении состава словарной информации для конкретных лексических единиц;
- познакомить студентов с проблематикой и конкретными разработками в области терминографии и тезаурусной лексикографии;
- сформировать представления о видах словарной информации, необходимой для автоматической обработки текста, и о способах формализации такой информации в конкретных разработках;
- раскрыть перед студентами пути построения информационных систем, аналогичных словарным, для систематизации теоретических сведений в разных областях лингвистики (в семантике, типологии и др.).
- научить студентов проектировать словарные базы данных и терминологические банки данных несложной структуры;
- научить студентов пользоваться учебной и научной литературой по компьютерной лексикографии, в том числе публикациями в трудах научных конференций.

Дисциплина (*модуль*) направлена на формирование следующих компетенций:

ПК-8 – способность применять методiku ориентированного поиска информации в справочной, специальной литературе и компьютерных сетях

ПК-15 – способностью к обобщению, критическому осмыслению, систематизации информации, анализу логики рассуждений и высказываний

ПК-18 – способностью применять методы научных исследований в профессиональной деятельности, анализировать материалы исследований в области лингвистики, межкультурной коммуникации и переводоведения с соблюдением библиографической культуры для решения профессиональных задач

В результате освоения дисциплины (*модуля*) обучающийся должен:

Знать

- основные понятия и термины традиционной лексикографии;
- основные понятия и термины информатики (в том числе области «Базы данных»), связанные с автоматизацией словарного дела;
- основные типы информации, необходимой для автоматической обработки текста;
- основные способы применения гипертекстовых и мультимедийных технологий в компьютерной лексикографии, нашедшие реализацию в современных системах;

Уметь

- использовать корпусный материал для выявления различных типов словарной информации;
- пользоваться учебной литературой по традиционной лексикографии и по компьютерной лингвистике;
- проектировать структуру простейших баз и банков данных для хранения лексикографической и терминографической информации;
- читать и анализировать относительно несложные научные публикации по компьютерной лексикографии;

Владеть

- простейшими навыками структурирования лексикографической информации;
- простейшими навыками расстановки тезаурусных связей.

методикой ориентированного поиска информации в справочной, специальной литературе и компьютерных сетях

По дисциплине предусмотрена промежуточная аттестация в форме *зачета*.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы.

ЛИСТ ИЗМЕНЕНИЙ

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
1	<i>Обновлена основная и дополнительная литература</i>	21.06.2017	6
2	Приложение №1		
3	<i>Обновлена основная и дополнительная литература</i>	20.06.2018	6
4	Приложение №2		
5	<i>Обновлена основная и дополнительная литература</i>	22.06.2019	6
6	Приложение №3		
7	<i>Обновлены структура дисциплины, образовательные технологии, основная и дополнительная литература</i>	26.06.2020	6
8	Приложение №4		

Состав программного обеспечения (ПО), современных профессиональных баз данных (БД) и информационно-справочные систем (ИСС) (2017 г.)

1. Перечень ПО

Таблица 1

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
5	Archicad 19 Rus Student	Graphisoft	свободно распространяемое
9	Microsoft Office 2013	Microsoft	лицензионное
10	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
11	Kaspersky Endpoint Security	Kaspersky	лицензионное

*** Оставить используемое ПО в рамках учебной дисциплины**

2. Перечень БД и ИСС

Таблица 2

№п/п	Наименование
	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2017 г. Web of Science Scopus
	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2017 г. Журналы Oxford University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам
	Компьютерные справочные правовые системы Консультант Плюс, Гарант

Состав программного обеспечения (ПО), современных профессиональных баз данных (БД) и информационно-справочные систем (ИСС) (2018 г.)

1. Перечень ПО

Таблица 1

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
5	Archicad 21 Rus Student	Graphisoft	свободно распространяемое
9	Microsoft Office 2013	Microsoft	лицензионное
10	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
11	Microsoft Office 2013	Microsoft	лицензионное
12	Windows 10 Pro	Microsoft	лицензионное
13	Kaspersky Endpoint Security	Kaspersky	лицензионное

** Оставить используемое ПО в рамках учебной дисциплины*

2. Перечень БД и ИСС

Таблица 2

№п/п	Наименование
	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2018 г. Web of Science Scopus
	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2018 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis Электронные издания издательства Springer
	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам
	Компьютерные справочные правовые системы Консультант Плюс, Гарант

Состав программного обеспечения (ПО), современных профессиональных баз данных (БД) и информационно-справочные систем (ИСС) (2019 г.)

1. Перечень ПО

Таблица 1

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
5	Archicad 21 Rus Student	Graphisoft	свободно распространяемое
9	Microsoft Office 2013	Microsoft	лицензионное
10	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
11	Microsoft Office 2013	Microsoft	лицензионное
12	Windows 10 Pro	Microsoft	лицензионное
13	Kaspersky Endpoint Security	Kaspersky	лицензионное

2. Перечень БД и ИСС

Таблица 2

№п/п	Наименование
	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2018 г. Web of Science Scopus
	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2018 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis Электронные издания издательства Springer
	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам
	Компьютерные справочные правовые системы Консультант Плюс, Гарант

1. Структура дисциплины (к п. 2 РПД на 2020)**Структура дисциплины (модуля) для очной формы обучения**

Общая трудоёмкость дисциплины составляет 2 з.е., 76 ч., в том числе контактная работа обучающихся с преподавателем 28 ч., самостоятельная работа обучающихся 48 ч.

№ п/п	Раздел дисциплины/темы	Семестр	Виды учебной работы (в часах)						Формы текущего контроля успеваемости, форма промежуточной аттестации (<i>по семестрам</i>)
			Контактная				Промежуточная аттестация	Самостоятель- ная работа	
			Лекции	Семинар	Практические занятия	Лабораторные занятия			
	Тема 1. Вводная часть курса. Обзор проблематики традиционной лексикографии	6	1					2	Устный ответ
	Тема 2. Понятие и основные направления компьютерной лексикографии	6	1					2	Устный ответ
	Тема 3. История компьютерной лексикографии	6	2					4	Устный ответ
	Тема 4. Общая и компьютерная лексикографии информационные технологии	6	2					4	Устный ответ
	Тема 5. Мультимедийные словари	6	1		1			2	Устный ответ
	Тема 6. Терминологическое банки данных	6	1					2	Устный ответ
	Тема 7. Лексическое наполнение терминологического банка данных	6	1		2			4	Устный ответ
	Тема 8. Тезаурусные терминологические системы	6	1					2	Устный ответ
	Тема 9.	6			1			2	Устный ответ

	Толкование термина								
	Тема 10. Системы лингвистической терминологии	6	1					2	Устный ответ
	Тема 11. Общая и компьютерная лексикография корпусная лингвистика	6			1			2	Устный ответ
	Тема 12. Словарные ресурсы для автоматической обработки текста	6			3			4	Устный ответ
	Тема 13. Семантические словари для автоматической обработки текста	6			4			4	Устный ответ
	Тема 14. Информационные системы для теоретических исследований в лингвистике	6	2		4			4	Устный ответ
	<i>Зачёт</i>							6	<i>Устный ответ на контрольные вопросы</i>
	итоги:		12		16			48	

2. Образовательные технологии (к п.4 на 2020 г.)

В период временного приостановления посещения обучающимися помещений и территории РГГУ. для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

3. Перечень БД и ИСС (к п. 6.2 на 2020 г.)

№п /п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru
4	Компьютерные справочные правовые системы Консультант Плюс, Гарант

4. Состав программного обеспечения (ПО) (к п. 7 на 2020 г.)

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Adobe Master Collection CS4	Adobe	лицензионное
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
4	AutoCAD 2010 Student	Autodesk	свободно распространяемое
5	Archicad 21 Rus Student	Graphisoft	свободно распространяемое
6	SPSS Statistics 22	IBM	лицензионное
7	Microsoft Share Point 2010	Microsoft	лицензионное

8	SPSS Statistics 25	IBM	лицензионное
9	Microsoft Office 2013	Microsoft	лицензионное
10	ОС «Альт Образование» 8	ООО «Базальт СПО	лицензионное
11	Microsoft Office 2013	Microsoft	лицензионное
12	Windows 10 Pro	Microsoft	лицензионное
13	Kaspersky Endpoint Security	Kaspersky	лицензионное
14	Microsoft Office 2016	Microsoft	лицензионное
15	Visual Studio 2019	Microsoft	лицензионное
16	Adobe Creative Cloud	Adobe	лицензионное
17	Zoom	Zoom	лицензионное