Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Оренбургский государственный университет имени В.А. Бондаренко»

Кафедра прикладной лингвистики и перевода

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б1.Д.В.Э.2.2 Компьютерная обработка текста»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки
45.03.02 Лингвистика
(код и наименование направления подготовки)

Перевод и переводоведение (английский язык, второй иностранный язык)
(наименование направленности (профиля) образовательной программы)

Квалификация
Бакалавр

Форма обучения
Очная

Год набора 2026

Рабочая программа дисциплины «*Б1.Д.В.Э.2.2 Компьютерная обработка текста*» рассмотрена и утверждена на заседании кафедры

Кафедра прикладной лингвистики и перевода
наименование кафедры

протокол № 7 от "20" 03 2026 г.

Заведующий кафедрой
Кафедра прикладной лингвистики и перевода Е.Д. Андреева
наименование кафедры *подпись* *расшифровка подписи*

Исполнители:
доцент Е.Д. Андреева
должность *подпись* *расшифровка подписи*

должность *подпись* *расшифровка подписи*

СОГЛАСОВАНО:
Председатель методической комиссии по направлению подготовки
45.03.02 Лингвистика А.В. Павлова
код наименование *личная подпись* *расшифровка подписи*

Заведующий отделом формирования фонда и научной обработки документов
С.Н. Вихтинцева
личная подпись *расшифровка подписи*

Уполномоченный по качеству факультета
Т.В. Сапух
личная подпись *расшифровка подписи*

№ регистрации _____

© Андреева Е.Д., 2026
© ОГУ, 2026

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины:

развитие способности применять автоматический лингвистический анализ и синтез текста, а также ознакомление студентов с лингвистическими аспектами общения с компьютером на естественном языке.

Задачи:

- овладеть основами автоматического индексирования и реферирования текстов, предредактирования и постредактирования, статистической обработки специальных текстов, машинного перевода и многоязыковой генерации текстов;
- овладеть навыками составления глоссариев и терминологических баз и работы с ними;
- овладеть основами поиска и извлечения текстовой информации, извлечения знаний из больших массивов информации, работы с различными носителями информации, распределенными базами данных и знаний, глобальными компьютерными сетями;
- приобрести опыт работы с программным обеспечением для автоматического анализа текстов (морфологическими и синтаксическими анализаторами, конкордансами), классификации и тематического моделирования коллекций документов;
- овладеть методикой подготовки к переводу и переводческому проекту в части, касающейся компьютерной обработки информации;
- научиться оформлять, форматировать, конвертировать, верстать текст перевода в разных типах компьютерных редакторов.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока Д «Дисциплины (модули)»

Пререквизиты дисциплины: *Б1.Д.Б.11 Информатика, Б1.Д.Б.12 Информационные технологии и программирование, Б1.Д.В.1 Теория перевода*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
ПК*-2 Способен осуществлять подготовку к выполнению перевода	ПК*-2-В-2 Составляет необходимые глоссарии и терминологические базы, работает с готовыми глоссариями и терминологическими базами ПК*-2-В-6 Осуществляет техническую подготовку текста к переводу путем распознавания, набора, форматирования, извлечения из графики и конвертации в требуемые форматы с использованием доступных программных средств	**Знать:** авторитетные справочно-информационные базы, лексические корпусы, поисковые системы; типы и принципы поисковых запросов; принципы работы с глоссариями и терминологическими базами; программы и ресурсы для распознавания, набора, форматирования текста, извлечения графики, конвертирования форматов. **Уметь:** использовать возможности, предоставляемые глобальными компьютерными сетями, для оптимального решения переводческих задач; пользоваться широким набором поисковых систем, электронных словарей и

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
		переводчиков, справочно-информационных баз, электронных ресурсов, параллельных корпусов; формулировать поисковые запросы и использовать поисковый инструментарий и справочный аппарат, составлять тезаурусы и глоссарии; подготавливать текст к переводу. **Владеть:** основными стратегиями поиска информации в разных типах источников; навыками работы с параллельными корпусами, автоматическими и автоматизированными системами перевода, составления глоссариев, терминологических баз; навыками компьютерной обработки текста.
ПК*-4 Способен редактировать и корректировать текст перевода, проверять его качество и соответствие переводческому заданию, оформлять текст перевода в соответствии с требованиями заказчика	ПК*-4-В-3 Редактирует текст перевода (вносит в текст перевода необходимые смысловые, лексические, грамматические и стилистические изменения), осуществляет сопоставление с оригиналом, вычитку и корректуру текста перевода ПК*-4-В-4 Адекватно применяет методы, процедуры и программные средства контроля качества перевода ПК*-4-В-5 Форматирует, конвертирует, верстает письменный текст перевода, восстанавливает исходное форматирование в соответствии с переводческим заданием; продуцирует устный текст перевода в соответствии с требованиями заказчика	**Знать:** основы редактуры текста; методы, процедуры и программные средства контроля качества перевода; различия текстовых форматов; программы и ресурсы для форматирования и верстки текста. **Уметь:** вносить в текст перевода необходимые смысловые, лексические, грамматические и стилистические изменения, сопоставлять текст с оригиналом, осуществлять вычитку и корректуру текста перевода; приводить текст в соответствие с требованиями заказчика. **Владеть:** навыками контроля качества перевода; навыками обработки внешней формы текста; нормативными требованиями к оформлению перевода в компьютерном текстовом редакторе: редактирование и форматирование текстов различной сложности; внедрение в текст сторонних объектов (шрифтов, таблиц, формул, графиков, фотографий, изображений, диаграмм).
ПК*-5 Способен использовать средства автоматизации переводческой деятельности и цифровые технологии для решения переводческих задач	ПК*-5-В-1 Корректно подготавливает исходный текст для работы в системах машинного перевода и CAT-программах в соответствии с требованиями ПК*-5-В-2 Выбирает соответствующие переводческой задаче цифровые технологии и средства автоматизации ПК*-5-В-3 Адекватно выполняет перевод в системах машинного перевода и CAT-программах ПК*-5-В-4 Адекватно осуществляет постредактирование текста перевода	**Знать:** средства автоматизации переводческого процесса; методику подготовки текста для работы в системах машинного перевода и CAT-программах. **Уметь:** осуществлять предперевод и перевод в системах машинного перевода и CAT-программах, постредактирование текста перевода. **Владеть:** навыками подготовки текста к переводу, автоматизированного перевода и постредактирования машинного перевода; организации и осуществления переводческих проектов.

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
	ПК*-5-В-5 Выступает как участник и организатор переводческих проектов	

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 3 зачетные единицы (108 академических часов).

Вид работы	Трудоемкость, академических часов	
	4 семестр	всего
Общая трудоёмкость	**108**	**108**
Контактная работа:	**34,25**	**34,25**
Практические занятия (ПЗ)	34	34
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: *- выполнение индивидуального творческого задания (ИТЗ);* *- самоподготовка (проработка и повторение материала учебников и учебных пособий;* *- изучение разделов курса в системе электронного обучения;* *- подготовка к практическим занятиям)*	**73,75**	**73,75**
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	**зачет**	

Разделы дисциплины, изучаемые в 4 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Системы компьютерной (автоматической) обработки текста	18		6		12
2	Поиск в источниках лингвистического материала	18		6		12
3	Системы аннотирования и реферирования	14		4		10
4	Машинный и автоматизированный перевод как система обработки текста	32		10		22
5	Оформление текста перевода в компьютерном текстовом редакторе	26		8		18
	Итого:	108		34		74
	Всего:	108		34		74

4.2 Содержание разделов дисциплины

1. Системы компьютерной (автоматической) обработки текста

Лингвистический автомат. Классификация систем КОТ по выполняемым функциям. Системы первого типа (программы Pragma, Lingvo, Белазар, Promt, Socrat, Google Translate и др.). Системы генерации (синтеза) текстов по особому формальному описанию (программы SciGen, Textgen, VIRGO 4,

Кобра 9.1). Трудности создания полноценно работающих с языком систем КОТ. Основные направления применения систем КОТ. Опознание языка документа и нормализация входного текста. Подблок опознания формата текста и его частей, а также определение их жанровой и тематической принадлежности. Моделирование языка с помощью конечного автомата. Формальные языки. Распознавание как поиск. Построение лексикона на основе конечного автомата.

2. Поиск в источниках лингвистического материала

Типы поиска. Поиск в словарях. Макроструктура словаря. Микроструктура словаря. Типология лингвистических словарей. Статистика речи и автоматический анализ текста. Лингвистический мониторинг функционирования языка. Компьютерное моделирования языка и речи. Понятие корпусной лингвистики. Национальный корпус русского языка. Британский национальный корпус (British National Corpus, BNC). Лингвистические исследования на основе корпуса. Конкорданс как инструмент переводчика.

3. Системы аннотирования и реферирования

Подходы и методы автоматического реферирования. Экстрактивные методы. Статистические методы. Логико-математические методы. Дистрибутивный метод. Метод содержательных аспектов. Метод текстовых связей. Методы с опорой на знания.

4. Машинный и автоматизированный перевод как система обработки текста

Устройство программ машинного и автоматизированного перевода. Подъязыки технической документации. Предредактирование. Постредактирование. Основные системы Translation Memory.

5. Оформление текста перевода в компьютерном текстовом редакторе

Типы текстовых редакторов. Microsoft Office Word. Ассистент слияния. Автотекст. Стили. Кадры. Шаблоны. Формы. Макрокоманды. Главный документ. Таблицы. Связь с другими документами. Извлечение графики и внедрение сторонних объектов. Верстка. Конвертация файлов. Форматирование по заданным параметрам.

4.3 Практические занятия (семинары)

№ занятия	№ раздела	Тема	Кол-во часов
1-2	1	Системы КОТ	4
3	1	Конечные автоматы	2
4	2	Типы поиска	2
5	2	Лингвистические словари	2
6	2	Корпуса. Параллельные корпуса. Конкорданс	2
7-8	3	Системы аннотирования и реферирования	4
9-10	4	Машинный перевод. Постредактирование.	4
11-13	4	Автоматизированный перевод	6
14-15	5	Основные параметры форматирования текста в компьютерном текстовом редакторе	4
16-17	5	Внедрение сторонних объектов	4
		Итого:	34

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1. Баймуратова, У.С. Электронный инструментарий переводчика: учебное пособие для студентов, обучающихся по программам высшего профессионального образования по направлению подготовки 035700.62 Лингвистика / У. С. Баймуратова. – Оренбург: Университет, 2013. – 120 с.

2. Переходько, И.В. Компьютерные технологии в переводе [Электронный ресурс]: учебное пособие для обучающихся по образовательной программе высшего образования по направлению подготовки 45.03.02 Лингвистика / И.В. Переходько. – Электрон. текстовые дан. (1 файл: 15813 Кб). – Оренбург : ОГУ, 2018. – 110 с. – Загл. с тит. экрана. – Adobe Acrobat Reader 6.0. – Режим доступа: http://artlib.osu.ru/web/books/metod_all/85926_20181205.pdf – ISBN 978-5-7410-2208-5.

5.2 Дополнительная литература

1. Зубов, А.В. Информационные технологии в лингвистике / А.В. Зубов, И.И. Зубова. – М. : Академия, 2004. – 207 с. – ISBN 5-7695-1531-7.

2. Савина, О.Ю. Компьютерные технологии в переводе: учебно-методическое пособие для студентов направления 45.03.02 «Лингвистика» (уровень бакалавриата). Профиль подготовки «Перевод и переводоведение», форма обучения – очная / О.Ю. Савина ; отв. ред. И.В. Соловьёва. – Тюмень : Тюменский государственный университет, 2017. – 28 с. – Режим доступа: по подписке. – URL: http://biblioclub.ru/index.php?page=book&id=573752. – Текст : электронный.

3. Захарова, Т.В. Практические основы компьютерных технологий в переводе[Электронный ресурс]: учебное пособие для обучающихся по образовательным программам высшего образования по направлениям подготовки 45.03.02 Лингвистика, 45.03.01 Филология / Т.В. Захарова, Е.В. Турлова. – Электрон. текстовые дан. (1 файл: 2.24 Мб). – Оренбург : ОГУ, 2017. – 108 с. –Загл. с тит. экрана. – Adobe Acrobat Reader 6.0 – ISBN 978-5-7410-1736-4. – Режим доступа: http://artlib.osu.ru/web/books/metod_all/36381_20170503.pdf

4. Семенов, А.Л. Современные информационные технологии и перевод : учеб. пособие для вузов / А. Л. Семенов. – М. : Академия, 2008. – 224 с. – ISBN 978-5-7695-4459-0.

5. Формализация исследовательских процедур анализа семантики языковых единиц : коллективная монография / М.В. Каменский, Т.Н. Ломтева, Н.С. Кабылкина и др. ; под общ. ред. М.В. Каменского. – Ставрополь : СКФУ, 2016. – 170 с. – ISBN 978-5-9296-0838-4 [Электронный ресурс]. – URL: http://biblioclub.ru/index.php?page=book&id=466913 – Текст : электронный.

5.3 Периодические издания

Вестник компьютерных и информационных технологий : журнал. – Москва : Агентство «Роспечать», 2010-2023.

5.4 Интернет-ресурсы

https://ru.smartcat.com/ – система TM Smartcat

https://www.multitran.ru/– *словарь для переводчиков «Мультитран»*

http://www.translate.ru– *онлайн-переводчик и словарь PROMT*

http://translate.google.com – *онлайн-переводчик Google Translate*

http://www.ruscorpora.ru – Национальный корпус русского языка

http://www.natcorp.ox.ac.uk/ – Британский национальный корпус

https://nsu.ru/xmlui/bitstream/handle/nsu/8982/17645.pdf?sequence=1&isAllowed=y – Учебное пособие «Автоматическая обработка текстов на естественном языке и компьютерная лингвистика», авторы Е.И. Большакова [и др.]

https://www.mitpressjournals.org/loi/coli – сайт журнала Computational Linguistics.

https://ai.google/research/pubs?area=Natural Language Processing – база материалов по обработке естественных языков на портале Google AI.

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы

1. Операционная система РЕД ОС.
2. Пакет офисных приложений «МойОфис Образование»
3. Для работы с ресурсами Интернет - веб-браузер Яндекс https://yandex.ru/.

4. ГАРАНТ Платформа F1 [Электронный ресурс]: справочно-правовая система. / Разработчик ООО НПП «ГАРАНТ-Сервис», 119992, Москва, Воробьевы горы, МГУ, [1990–2026]. – Режим доступа в сети ОГУ http://garant.net.osu.ru

5. Автоматизированная интерактивная система сетевого тестирования – АИССТ (зарегистрирована в РОСПАТЕНТ, Свидетельство о государственной регистрации программы для ЭВМ №2011610456, правообладатель – Оренбургский государственный университет), режим доступа – http://aist.osu.ru.

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий семинарского типа, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Аудитории оснащены комплектами ученической мебели, техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой, подключенной к сети «Интернет» и обеспеченной доступом в электронную информационно-образовательную среду ОГУ.