***На правах рукописи***

Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«Оренбургский государственный университет»**

Кафедра романской филологии и методики преподавания французского языка

Методические указания для обучающихся по освоению дисциплины

*«Б1.Д.В.Э.3.2 Компьютерная обработка текста»*

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

*45.03.02 Лингвистика*

(код и наименование направления подготовки)

*Перевод и переводоведение (французский язык, второй иностранный язык)*

(наименование направленности (профиля) образовательной программы)

Тип образовательной программы

*Программа академического бакалавриата*

Квалификация

*Бакалавр*

Форма обучения

*Очная*

Год набора 2025

Составитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ И.Ю. Моисеева

Методические указания рассмотрены и одобрены на заседании кафедры романской филологии и методики преподавания французского языка

Заведующий кафедрой \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Методические указания являются приложением к рабочей программе по дисциплине «Б1.Д.В.Э.3.2 Компьютерная обработка текста», зарегистрированной в ЦИТ под учетным номером\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Содержание**

[1 Общие сведения о дисциплине «Компьютерная обработка текста» 4](#_Toc19187546)

[2 Методические указания по практическим занятиям 6](#_Toc19187547)

[3 Методические указания по выполнению индивидуального задания 9](#_Toc19187548)

[4 Методические указания по самостоятельной работе 10](#_Toc19187549)

[5 Методические указания по промежуточной аттестации 12](#_Toc19187550)

**1 Общие сведения о дисциплине «Компьютерная обработка текста»**

Компьютерная обработка текста (КОТ) является частью информационных технологий и обеспечивают сбор, обработку, хранение и передачу информации с помощью ЭВМ.

Основу современной КОТ составляют 3 технологических достижения: возможность хранения информации на машинных носителях, развитие средств связи и автоматизация обработки информации с помощью компьютера.

Практически КОТ реализуется применением программно-технических комплексов (ПТК), состоящих из персональных компьютеров (ПК) или рабочих станций (РС) с необходимым набором периферийных устройств, включенных в локальные и глобальные вычислительные сети и обеспеченных необходимыми программными средствами (ПС). Использование названных элементов увеличивает степень автоматизации как научных исследований, так и учебных процессов, что служит основой их совершенствования.

КОТ повышает уровень эффективности работ в переводческой деятельности за счет следующих факторов:

1. Упрощение и ускорение процессов обработки, передачи, представления и хранения информации.
2. Увеличение объема полезной информации с накопителем типовых решений и обобщением опыта научных разработок.
3. Обеспечение глубины, точности и качества решаемых задач. Возможность реализации задач ранее не решаемых. Постановка исследований и получение результатов, недостижимых другими средствами.
4. Возможность анализа большого числа вариантов синтеза объектов (текстов перевода) и принятия решений.

КОТ в настоящее время используется практически во всех сферах деятельности человека. Задача дисциплины – обобщить знания по КОТ применительно к переводческой деятельности и конечному продукту объекта перевода – тексту.

Дисциплина «Компьютерная обработка текста» осваивается практически на примерах и заданиях, ориентированных на решение творческих задач по компьютерному переводу. Моделируя текст перевода с помощью компьютерного редактора, студент должен знать специфику и методологию формообразования и структурные факторы инструментального арсенала компьютерной программы. Выдаваемые на занятиях задания сопровождаются вводными инструкциями, в которых излагаются сведения об эффективных приемах компьютерной обработки текста, наиболее эффективных практических алгоритмах в этой сфере. В начале работы над заданием студенты знакомятся с основными инструментами и возможностями интерфейса программы, ограничениями и возможными затруднениями в процессе выполнения практического задания.

По ходу выполнения и по результатам практических заданий проводится просмотр с обязательным обсуждением качеств работы, с указанием достоинств и недостатков выполненных заданий. Обсуждение результатов работы над учебными заданиями формирует у студентов навыки аргументации своих решений и воспитывает ответственность за результаты работы. Кроме того участие студентов в устном обсуждении развивает навыки убедительной устной речи и оперирование профессиональным тезаурусом.

Целью дисциплины является развить способности применять компьютерные технологии при осуществлении компьютерной обработки текста на основе сформированных представлений о современных методах получения, обработки и хранения информации*.*

Задачи дисциплины:

- овладеть основами автоматического индексирования и реферирования текстов, предредактирования и постредактирования, статистической обработки специальных текстов, машинного перевода и многоязыковой генерации текстов,

- овладеть лингвистической терминологией в области современных методов обработки текста на естественном языке,

- овладеть основами системы 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 в части, касающейся компьютерной обработки информации,

- научиться оформлять текст перевода в разных типах компьютерных редакторов.

Объектом исследования данной дисциплины является текст, работа с которым проводится в следующих направлениях.

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.

Источники и носители информации

Типы носителей информации. Базы данных и знаний. Устройство глобальных компьютерных сетей. Поиск и обработка информации с помощью компьютера. Лингвистическая и техническая подготовка к выполнению перевода: составление глоссариев, тезаурусов, поиск справочной информации, работа с электронными словарями. Принципы работы терминологических баз.

Типы лингвистических словарей. Oнлайн словари

Макроструктура словаря. Микроструктура словаря. Типология лингвистических словарей. Статистика речи и автоматический анализ текста. Лингвистический мониторинг функционирования языка. Компьютерное моделирования языка и речи.

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.

Машинный и автоматизированный перевод как система обработки текста

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

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.

Параллельный корпус как инструмент переводчика

Понятие корпусной лингвистики. Национальный корпус русского языка. Британский национальный корпус (British National Corpus, BNC). Лингвистические исследования на основе корпуса. Конкорданс как инструмент переводчика.

**2 Методические указания по практическим занятиям**

Практические занятия представляют собой определенную форму работы, использующую применение, углубление и развитие теоретических знаний совместно с формированием необходимых умений и навыков, самостоятельное применение учебника, наглядных пособий, компьютера и т.д. Проведение практических занятий при изучении курса компьютерной обработки текста позволяет сформировать у обучающихся общепрофессиональные и профессиональные умения, повышает уровень специальной профессиональной компетентности на основе знаний о способах использования возможностей компьютерных технологий, в том числе сети Интернет, в процессе работы с текстами и перевода с учетом требований информационной безопасности. Вместе с развитием умений и навыков в ходе выполнения заданий на практических занятиях систематизируются, усиливаются и уточняются теоретические знания, развивается способность и готовность применять теоретические знания на практике, формируются интеллектуальные умения, общепрофессиональные и профессиональные компетенции.

Целью проведения практических занятий по дисциплине «Компьютерная обработка текста» является установление связей теории с практикой в форме экспериментального подтверждения положений теории, обучение студентов основным приемам решения лингвистических задач различного класса с использованием компьютерной техники, современных информационных технологий и интернет ресурсов, контроль самостоятельной работы студентов по освоению курса. Цель практических занятий достигается лучше, если проведению контрольных работ предшествует некая подготовительная работа.

При домашней подготовке к выполнению заданий практических занятий студент должен повторить изученную тему.

Практическое занятие проводится в специально оборудованных компьютерами учебных аудиториях. Продолжительность практического занятия – не менее 2-х академических часов.

Проведению практического занятия должна предшествовать проверка теоретической подготовленности обучающихся.

В процессе практического занятия как вида учебного занятия студенты отвечают на заранее подготовленные вопросы по теме и выполняют несколько практических заданий под руководством преподавателя в соответствии с изучаемым содержанием учебного материала.

После выполнения практического задания студент демонстрирует результат выполнения преподавателю, отвечает на вопросы. Преподаватель оценивает работу.

Практическое задание выполняется каждым студентом самостоятельно.

Студенты, пропустившие занятия, выполняют практические задания во внеурочное время.

Для успешной реализации знаний, полученных в курсе дисциплины «Компьютерная обработка текста» и прохождения тестирования студенту необходимо иметь представление о терминах глоссария компьютерной обработки текста: компьютерная сеть, сетевые технологии, информационные системы, протокол, домен, архив FTP, лингвистический корпус, электронный словарь, машинный перевод, контрастивная текстология, аналитико-синтетическая обработка текста, модель перевода, текстовый редактор, автоматический перевод, автоматизированный перевод.

Практические занятия по системам компьютерной обработки текста предполагают ознакомление с видами лингвистических автоматов, применение знаний для создания конечных автоматов для обработки разного рода текстовой информации и моделирования языка с помощью лингвистических автоматов (Pragma, Lingvo, Белазар, Promt, Socrat, Google Translate, SciGen, Textgen, VIRGO 4, Кобра 9.1). Студенты должны научиться работать с текстовой информацией (получать, обрабатывать и хранить) с применением компьютерных технологий и использовать лингвистические автоматы для переводческой и научно-исследовательской деятельности.

Задания по теме «Источники и носители информации» предполагает ознакомление с типами источников информации, поиск информации в глобальных компьютерных сетях. Студенты должны научиться формировать точные поисковые запросы, сортировать выдаваемую поисковой системой информацию, определять релевантную и нерелевантную информацию, осуществлять проверку информации на достоверность. Практические задания также предполагает освоение методики подготовки к выполнению перевода. Преподаватель задает тему планируемого перевода, студентам необходимо подготовиться к выполнению перевода по заданной тематике: составить глоссарии и тезаурусы, найти справочную информацию. Также в рамках данной темы студенты должны освоить навыки работы с существующими терминологическими базами и создания собственных терминологических баз, в первую очередь с помощью программ ТМ.

При выполнении практических заданий по теме «Типы лингвистических словарей. Онлайн словари» студенту следует ознакомиться со спецификой работы с некоторыми словарями и переводчиками из теоретических источников по заданной тематике, представленных в рабочей программе по данной дисциплине, а также создав конспект теоретического материала, представленного для ознакомления по заданной теме преподавателем. Принципы работы электронных словарей студент может найти на следующих сайтах:

- [http://www.lingvo.ru/](http://infourok.ru/go.html?href=http%3A%2F%2Fwww.lingvo.ru%2F);

- [http://www.wordreference.com/](http://infourok.ru/go.html?href=http%3A%2F%2Fwww.wordreference.com%2F);

- [http://dictionary.cambridge.org/](http://infourok.ru/go.html?href=http%3A%2F%2Fdictionary.cambridge.org%2F);

- [http://www.merriam-webster.com/](http://infourok.ru/go.html?href=http%3A%2F%2Fwww.merriam-webster.com%2F) и других.

Контроль знаний по данной теме осуществляется с помощью задания, предполагающего работу на компьютерном устройстве и включающего вопросы об онлайн словарях, а также выполнения перевода текста с помощью одного из онлайн словарей или переводчиков, представленного для работы преподавателем.

В рамках темы «Системы аннотирования и реферирования» студенты получают задание составить реферат / аннотацию текста «ручным» способом, затем с помощью автоматических программ реферирования и аннотирования. На материале данного задания определяются функции автоматических систем аннотирования и реферирования, их практическое применение в работе переводчика.

Практические задания по машинному переводу и системам памяти перевода основаны на теоретических знаниях студентом алгоритма работы Omega T, WordFast Anywhere, SDL Trados Studio. Описание принципа работы данных систем представлены в учебниках и учебных пособиях рабочей программы по компьютерным технологиям в переводе. Преподаватель дает обзорное занятие по системе памяти перевода Omega T и дает практическое задание по данной теме для выполнения в аудитории. Задание включает в себя работу с текстом и отдельными разделами текста при помощи Omega T.

Для актуализации письменной переводческой деятельности и качественной компьютерной обработки текста исходного и переводящего языка студенту необходимо осуществлять работу с текстовыми редакторами. Данная работа основывается на теоретических знаниях различных онлайн и офлайн текстовых редакторов (прежде всего Microsoft Office Word) и успешной ориентации в следующих теоретических положениях по данной тематике:

Что такое текстовый редактор?

Перечислите основные элементы интерфейса соответствующей программы.

Что такое строка заголовка? Что она содержит? Охарактеризуйте.

Что такое «Строка меню»?

Из каких вкладок она состоит? Охарактеризуйте каждую вкладку

Что такое панель инструментов? Для чего она предназначена? Что в неё входит?

Какие правила следует соблюдать при наборе текста с клавиатуры?

Какими способами можно перемещаться по тексту?

Перечислите правила выделения текста.

Практические занятия по работе с текстовым редактором включают в себя обзорное занятие по данной теме, работу с текстами оригинала и перевода с применением данной программы и проработкой всех теоретических положений на основе данных текстов.

На практических занятиях по лингвистическим корпусам студенты учатся работать с существующими корпусами языков и лингвистических данных, используют корпуса для проведения лингвистических исследований и составления конкордансов больших текстов для перевода. Полученные знания и навыки применяются для выполнения научно-исследовательских работ студентов.

**3 Методические указания по выполнению индивидуального задания**

Индивидуальное задание выполняется для систематизации, закрепления и расширения теоретических знаний и практических навыков, полученных в процессе обучения, а также формирования умений их применять при решении типовых заданий по дисциплине «Компьютерная обработка текста».

При выполнении индивидуальных заданий студентам целесообразно придерживаться следующих рекомендаций.

Готовиться к сдаче индивидуального задания целесообразно во время решения аналогичных заданий в течение всего семестра.

При выполнении индивидуального задания применять решения типовых вариантов, конспекты теоретического материалы, записи, выполненные на практических занятиях, основную и дополнительную литературу (Рабочая программа, пункты 5.1 и 5.2), специализированные сайты (Рабочая программа, пункт 5.4), информационные справочные системы современных информационных технологий (Рабочая программа, пункт 5.5), при необходимости осуществлять самостоятельный подбор учебников, методических рекомендаций.

При выполнении индивидуального задания следует обязательно записывать все пояснения, которые необходимы по ходу выполнения задания.

Если выполнение индивидуальное задание вызывает трудности, то допускаются консультации у преподавателя на практических занятиях.

На защите индивидуального задания студент должен быть готов дать объяснения по методам его выполнения.

Можно применять следующий алгоритм выполнения индивидуального задания:

1. сформировать понимание проблемной ситуации (знакомство с общей формулировкой задания, определение для себя новых, незнакомых понятий и действий);
2. формализовать задачу (уточнить условие задачи, входные данные, сформулировать перед собой цель, провести декомпозицию цели, сформулировать подзадачи);
3. спланировать самостоятельную деятельность по реализации задания (в письменном или печатном виде сформировать общий график работы, возможно, обсудить его с преподавателем).

**4 Методические указания по самостоятельной работе**

Самостоятельная учебная деятельность является необходимым условием успешного образования. Многие профессиональные навыки, способность мыслить и обобщать, делать выводы и строить суждения, выступать и слушать других, – все это развивается в процессе самостоятельной работы студентов. Самостоятельная учебная работа является равноправной формой учебных занятий, наряду с лекциями, семинарами, практическими и лабораторными занятиями, экзаменами и зачетами, но реализуемая во внеаудиторное время.

Самостоятельная работа – вид учебной деятельности, базирующийся на выполнении студентами комплекса усложняющихся профессионально-ориентированных заданий при консультационно-координирующей помощи преподавателя, ориентированный на формирование результатов обучения, выраженных соответствующими компетенциями.

Изучение дисциплины предполагает не только познавательную деятельность, которую студенты осуществляют во время практических занятий, но и самостоятельную работу, осуществляемую вне аудиторных занятий. Успешное усвоение учебного материала возможно только при комплексном подходе, состоящем в получении новой информации в ходе практического занятия; ее понимания и обобщения; записи в собственной интерпретации в виде текста, схем, таблиц; самостоятельного изучения и конспектирования рекомендованной учебной литературы; выполнения различных практических заданий.

Самостоятельная работа студента состоит, во-первых, из деятельности студентов во всех организационных формах учебных занятий и во внеаудиторное время, когда они самостоятельно изучают теоретический материал, определенный содержанием рабочей программы. Во-вторых, выполнения учебных заданий, которые рекомендованы студенту во время обучения: решить практические задания, выполнить контрольные или индивидуальные задания.

Самостоятельная внеаудиторная работа студента может включать в себя: типовые виды (выполняются всеми студентами) и нестандартные (выполняются не всеми, участие в этих видах зависит от наличия сформированности тех или иных компетенций).

К типовым видам относятся:

1) подготовка к занятиям;

2) выполнение типовых или усложняющихся учебных заданий, предусмотренных рабочей программой;

3) выполнение контрольных работ;

4) подготовка к неделе рубежного контроля;

5) подготовка и сдача зачетов и экзаменов.

К нестандартным видам самостоятельной работы можно отнести участие студента в научных исследованиях, проводимых в рамках студенческого научного общества.

Подготовка студента к занятиям должна включать в себя не только непосредственное выполнение домашнего задания. Она должна предусматривать тот факт, что последующее занятие будет направлено на изучение нового теоретического и/или практического материала. Такая подготовка предполагает изучение рабочей программы, установление связи с ранее полученными знаниями, выделение наиболее значимых и актуальных проблем, на изучении которых следует обратить особое внимание.

При подготовке к занятиям необходимо:

1) выполнять подбор, изучение, анализ, классификацию и конспектирование литературы по учебной дисциплине, рекомендованной в рабочей программе, соответственно изучаемой теме (Рабочая программа, пункты 5.1 и 5.2);

2) систематически читать периодическую печать, осуществлять поиск и анализ дополнительной информации в журналах, рекомендованных рабочей программой по изучаемой дисциплине, с целью выяснения наиболее сложных, непонятных вопросов и их уточнения во время консультаций (Рабочая программа, пункт 5.3);

3) осуществлять активный поиск информации по изучаемой теме с использованием возможностей информационно-поисковых систем, а также сайтов, рекомендованных рабочей программой (Рабочая программа, пункты 5.4);

4) осуществлять просмотр вебинаров, организованных или предложенных преподавателем.

**5 Методические указания по промежуточной аттестации**

Итоговой формой контроля знаний, умений и навыков по дисциплине является зачет.

**Подготовка студента к зачету**. «Зачет является традиционной формой проверки знаний, умений, компетенций, сформированных у студентов в процессе освоения всего содержания изучаемой дисциплины. Обычный зачет отличается от экзамена только тем, что преподаватель не дифференцирует баллы, которые он выставляет по его итогам.

Самостоятельная подготовка к зачету должна осуществляться в течение всего семестра, а не за несколько дней до его проведения. Подготовка включает следующие действия. Прежде всего, нужно перечитать все лекции, а также материалы, которые готовились к семинарским и практическим занятиям в течение семестра. Затем надо соотнести эту информацию с вопросами, которые даны к зачету. Если информации недостаточно, ответы находят в предложенной преподавателем литературе. Рекомендуется делать краткие записи. Накануне зачета необходимо повторить ответы, не заглядывая в записи.

На подготовку ответа на зачете дается 60 минут. Сначала необходимо выполнить практическое задание, т.к. теоретический вопрос считается выученным заранее. При условии завершения практической части билета до окончания отведенного на подготовку времени, можно записать краткий конспект ответа на теоретический вопрос.

Результаты сдачи зачета оцениваются в формате «зачтено» и «незачтено».

Оценкой «зачтено» оцениваются ответы студентов, показавших знание основного учебного материала в объеме, необходимом для дальнейшей учебы и в предстоящей работе по профессии, справляющихся с выполнением заданий, предусмотренных программой, но допустившим погрешности в ответе и при выполнении практических заданий, не носящие принципиального характера, когда установлено, что студент обладает необходимыми знаниями для последующего устранения указанных погрешностей под руководством преподавателя..

«Незачтено» выставляется студентам, обнаружившим пробелы в знаниях основного учебного материала, допускающим принципиальные ошибки в выполнении предусмотренных программой заданий.