

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Филологический факультет

УТВЕРЖДЕНО:

Декан

И. В.Тубалова

Оценочные материалы по дисциплине

Корпусная лингвистика

по направлению подготовки

45.04.01 Филология

Направленность (профиль) подготовки:

Юридическая лингвистика

Форма обучения

Очная

Квалификация

Лингвист-эксперт

Год приема

2025

СОГЛАСОВАНО:

Руководитель ОП

А.В. Шевчик

Председатель УМК

Ю.А. Тихомирова

Томск – 2025

1. Компетенции и индикаторы их достижения, проверяемые данными оценочными материалами

Целью освоения дисциплины является формирование следующих компетенций:

БК-2 Способен использовать научные методы для решения профессиональных задач.

ОПК-3 Способен применять широкий спектр методов и приемов филологической работы с текстами различного типа.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

РОБК-2.1 Знает основные методы научных исследований

РОБК-2.2 Умеет выстраивать систематическую и логическую цепочку анализа и принимаемых решений в контексте задачи профессиональной деятельности

РООПК-3.1 Знает методы и приемы профессиональной обработки, филологического анализа и оценки текстов различного типа

РООПК-3.2 Умеет осуществлять выбор методов и приемов профессиональной обработки, анализа и оценки текстов различного типа и применять их для решения профессиональных задач конкретной области филологии

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

- опросы в тестовой форме;
- домашние задания;
- самостоятельные работы.

Опрос в тестовой форме (РООПК-3.1)

1. Обязательными свойствами хорошего лингвистического корпуса являются:

- а) представительность;
- б) присутствие разметки;
- в) снятая грамматическая омонимия;
- г) баланс между разными типами текстов.

2. Разметка Обучающего корпуса отличается от аннотации других частей НКРЯ тем что:

- а) она содержит разметку ошибок;
- б) она проведена в соответствии со школьной грамматикой;
- б) в ней присутствует синтаксическая разметка;
- г) она произведена вручную.

3. Список всех употреблений определенного языкового выражения в контексте и с ссылкой на источник при необходимости - _____.

Ключи: 1) а, б, г; 2) а; 3) конкорданс.

Критерии оценивания: тест считается пройденным, если обучающий ответил правильно как минимум на половину вопросов.

Домашние задания (РОБК-2.1, РОБК-2.2, РООПК-3.1, РООПК-3.2):

Домашнее задание 1

1. Опишите, какие перспективы открывает перед вашим собственным исследованием «корпусная идеология».

2. Охарактеризуйте один из известных вам корпусов по признакам сбалансированности и репрезентативности.

Домашнее задание 2

3. Изучите системы экстралингвистической разметки двух корпусов (по выбору). Соотнесите целевую направленность корпуса и систему его метаразметки.

Критерии оценивания:

Результаты выполнения домашнего задания определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» выставляется, если в ответе приводится не менее 3 аргументов, они верны и подкреплены релевантными примерами из лингвистических корпусов.

Оценка «хорошо» выставляется, если в ответе приводится 1-2 аргумента, они верны и подкреплены релевантными примерами из лингвистических корпусов.

Оценка «удовлетворительно» выставляется, если в ответе приводится 1-2 аргумента, но они содержат фактические ошибки до 50%, подкреплены примерами из лингвистических корпусов, релевантность которых достигает 50%.

Оценка «неудовлетворительно» выставляется, если в ответе не приводится аргументов, они не подкреплены примерами из лингвистических корпусов или приводится 1-2 аргумента, но они содержат фактические ошибки более 50%, подкреплены примерами из лингвистических корпусов, релевантность которых не достигает 50%.

Самостоятельная работа (РОБК-2.1, РОБК-2.2, РООПК-3.1, РООПК-3.2)

1. Зайдите на сайт Национального корпуса русского языка, и выясните, каков объем представленных в нём корпусов в настоящий момент.

2. На странице выбора метаразметки НКРЯ (ruscorpora.ru/mycorpora-main.html) соберите подкорпус текстов Вашего любимого автора или писателя, поэта, чьё творчество Вы изучаете.

3. Изучите систему семантических помет в НКРЯ. Проведите поиск по одной из выбранных семантических групп. Зафиксируйте найденные ошибки. Чем они обусловлены?

Критерии оценивания:

Результаты самостоятельной работы определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценка «отлично» выставляется, если в ответе на каждый вопрос приводится не менее 3 аргументов, они верны и подкреплены релевантными примерами из лингвистических корпусов.

Оценка «хорошо» выставляется, если в ответе на каждый вопрос приводится 1-2 аргумента, они верны и подкреплены релевантными примерами из лингвистических корпусов.

Оценка «удовлетворительно» выставляется, если в ответе на каждый вопрос приводится 1-2 аргумента, но они содержат фактические ошибки до 50%, подкреплены примерами из лингвистических корпусов, релевантность которых достигает 50%.

Оценка «неудовлетворительно» выставляется, если в ответах не приводится аргументов, они не подкреплены примерами из лингвистических корпусов или приводится 1-2 аргумента, но они содержат фактические ошибки более 50%, подкреплены примерами из лингвистических корпусов, релевантность которых не достигает 50%.

3. Оценочные материалы итогового контроля (промежуточной аттестации) и критерии оценивания

Зачет с оценкой в первом семестре проводится в письменной форме по билетам. Билет содержит теоретический вопрос и практическое задание. Первая часть представляет собой вопрос, проверяющий РОБК-2.1, РООПК-3.1. Ответ на вопрос первой части дается в развернутой форме. Вторая часть содержит практическое задание, проверяющее РОБК-2.2, РООПК-3.1, РООПК-3.2. Продолжительность зачета 1,5 часа.

Перечень теоретических вопросов по курсу.

1. Что такое лингвистический корпус? В чём его отличие от других типов лингвистических ресурсов?

2. Какие корпусные ресурсы Вам известны? (Не менее 5). Охарактеризуйте один из них по 5 типологическим признакам.
3. По каким признакам классифицируются корпуса? Какие типологии корпусов существуют?
4. Какие параметры включаются в экстралингвистическую аннотацию? Какова их функция?
5. Охарактеризуйте понятие «репрезентативность» корпуса. Для чего нужна *ipm* и как рассчитывается эта характеристика?
6. В чём основные отличия морфологического стандарта НКРЯ от классической теории частей речи?
7. Что такое грамматическая омонимия? Какие способы снятия омонимии Вам известны?
8. Как организована синтаксическая разметка НКРЯ?
9. Как организована семантическая разметка НКРЯ? В чём её преимущества и ограничения?
10. В чём специфика учебных корпусов? Для каких целей они используются?
11. В чём специфика параллельных корпусов? Для каких целей они используются?
12. В чём специфика мультимодальных корпусов? Для каких целей они используются?
13. В чём специфика диалектных корпусов? Для каких целей они используются?
14. В чём отличие corpus-based и corpus-driven подходов к исследованию языка на базе корпуса?
15. Как корпусная лингвистика повлияла на изменение представлений о языке?

Примеры практических заданий:

1. Пользуясь любым корпусом русского языка, расположите падежи имён существительных в соответствии с частотой их употребления в корпусе.
2. Подсчитайте, сколько раз в НКРЯ используется имя существительное *быль*.
3. Соберите с помощью НКРЯ подкорпус примеров с прилагательными-обозначениями цвета в текстах В.П. Астафьева.
4. Сравните относительную частотность использования слова *тятя* в основном и диалектном подкорпусах НКРЯ.
5. Пользуясь НКРЯ, определите приблизительный период возникновения у слова передовой значения «прогрессивный».

Критерии оценивания

Зачёт с оценкой «отлично» выставляется за полный, логично выстроенный ответ. Обучающийся свободно ориентируется в проблематике курса, знает терминологию и основные положения дисциплины, отвечает на дополнительные вопросы, обобщающие содержание курса.

Зачёт с оценкой «хорошо» выставляется за правильные и достаточно полные ответы. Обучающийся свободно ориентируется в большинстве вопросов курса, но испытывает небольшие затруднения при ответе на отдельные вопросы (1 или 2), имеются незначительные погрешности в использовании терминологии.

Зачёт с оценкой «удовлетворительно» выставляется при недостаточно полном ответе, при наличии заметных пробелов в знаниях обучающегося, затруднениях при ответе на более чем 2 вопроса экзаменатора.

Незачёт выставляется в случае отсутствия знаний по курсу, неспособности ответить на вопросы экзаменатора.

Вопрос и задание оцениваются по шкале 3-5 баллов, текущий контроль также учитывается в промежуточной аттестации, оценка выставляется по шкале 3-5 баллов.

Интегральная оценка высчитывается по формуле «оценка за текущий контроль x 0,6 + оценка за промежуточную аттестацию x 0,4» и округляется до целого числа в пользу обучающегося.

4. Оценочные материалы для проверки остаточных знаний (сформированности компетенций)

Тест

1. (РООПК-3.1) Национальный корпус НЕ

А. включает поэтических текстов

В. отражает диалектные особенности говоров

С. отражает состояние языка страны в конкретные временные периоды

Д. отражает государственную идеологию страны

Ключ:

Д. отражает государственную идеологию страны

2. (РООПК-3.1) Тексты, составляющие лингвистический корпус

А. организованы для удобства чтения

В. сконструированы лингвистами

С. не включают речь необразованных носителей языка

Д. размечены лингвистами и программистами

Ключ: Д. размечены лингвистами и программистами

3. (РООПК-3.1) В лингвистический корпус могут входить

А. тексты электронной коммуникации

В. записи устной речи детей

С. транскрипции теле- и радиопередач

Д. произведения художественной литературы

Ключ: А, В, С, Д

4. (РОБК-2.1) Тексты данного корпуса созданы в определенный заданный период времени. О каком виде корпусов идет речь?

А. динамический

В. статический

С. синхронический

Д. исследовательский

Ключ: С. синхронический

5. (РООПК-3.2) Формат вывода конкорданса называется

А. KWIC

В. KVIC

Ключ: А. KWIC

6. (РООПК-3.2) POS – это

А. корпус-менеджер

В. вид разметки

С. название лингвистического корпуса

Д. формат выдачи данных

Ключ: В. вид разметки

7. (РООПК-3.1) Как называется подкорпус НКРЯ, включающий устные тексты, выровненные с соответствующими аудио- и видеофрагментами

А. МультиПАРК

В. МУРКО
С. Устный
D. Обучающий
Ключ: В. МУРКО

Информация о разработчиках

Шевчик Анна Валерьевна, кандидат филологических наук, доцент кафедры русского языка НИ ТГУ.