



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ  
ФЕДЕРАЦИИ

ФГБОУ ВО «ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ПЕДАГОГИЧЕСКИЙ ИНСТИТУТ

Кафедра иностранных языков и лингводидактики

УТВЕРЖДАЮ  
Директор \_\_\_\_\_ А.В. Семиров  
«9» апреля 2026 г.



**Рабочая программа дисциплины**

Наименование дисциплины **Б1.В.ДВ.01 Элективные дисциплины 1. Б1.В.ДВ.01.01 -  
Корпусная лингвистика**

Направление подготовки 44.03.05 Педагогическое образование (с двумя профилями  
подготовки)

Направленность подготовки Иностранный язык (английский) – Иностранный язык  
(китайский)

Квалификация выпускника – Бакалавр

Форма обучения очная

Согласована с УМС ПИ ИГУ

Протокол № 3 от «26» марта 2026 г.

Председатель \_\_\_\_\_ М.С. Павлова



Рекомендовано кафедрой:

Протокол № 5

От «18» марта 2026 г.

Зав. кафедрой \_\_\_\_\_ М.В. Носкова



Иркутск 2026 г.

### **I. Цели и задачи дисциплины:**

Цель дисциплины **Б1.В.ДВ.01.01 Корпусная лингвистика**: овладеть современными технологиями поиска и обработки лингвистических данных.

Задачи:

- введение в проблематику корпусной лингвистики, ознакомление с различными направлениями в исследовании языка, основанными на корпусных данных;
- овладение основными методами и приемами построения и использования электронных корпусов текстов;
- формирование у обучающихся практических навыков работы с существующими корпусами.

### **II. Место дисциплины в структуре ОПОП:**

2.1. Учебная дисциплина Б1.В.ДВ.01.01 Корпусная лингвистика относится к части программы, *формируемой участниками образовательных отношений* и является *элективной* дисциплиной.

2.2. Для изучения данной учебной дисциплины необходимы знания, умения и навыки, формируемые предшествующими дисциплинами:

*Б1.В.01 Практика устной и письменной речи (английский язык)*

*Б1.О.21 Технологии цифрового образования*

*Б1.О.22.01 Методика обучения и воспитания (английский язык)*

2.3. Перечень последующих учебных дисциплин, для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной:

*Б1.О.24 Решение практических задач*

*Б2.О.05(II) Производственная практика. Педагогическая практика*

### **III. Требования к результатам освоения дисциплины:**

Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

<b>Компетенция</b>	<b>Индикаторы компетенций</b>	<b>Результаты обучения</b>
ПК-1. Способен осваивать и использовать базовые научно-теоретические знания и практические умения по предмету в профессиональной деятельности	ИДК <i>ПК1.1</i> : анализирует содержание, сущность, закономерности, принципы и особенности изучаемых явлений и процессов, базовые теории в иноязычной предметной области; основы теоретических дисциплин в объеме, необходимом для решения педагогических и методических задач	<b>Знать:</b> – основы классификации лингвистических корпусов по разным основаниям; – современные представления о человеческой и человеко-машинной коммуникации; – принципы интерпретации количественных данных. <b>Уметь:</b> – анализировать полученные с помощью корпусных технологий языковые данные; – осуществлять подбор программного обеспечения для решения практических задач в области преподавания иностранного языка.

		<p><b>Владеть:</b></p> <ul style="list-style-type: none"> <li>– навыками обращения с системами управления базами данных и знаний различного назначения;</li> <li>– навыками анализа возможностей использования систем управления базами данных и знаний различного назначения.</li> </ul>
	<p>ИДК <i>ПК1.2</i>: использует базовые предметные научно-теоретические представления о сущности, закономерностях, принципах и особенностях изучаемых языковых явлений и процессов</p>	<p><b>Знать:</b></p> <ul style="list-style-type: none"> <li>– основы поиска языковых данных в корпусах современного английского языка,</li> <li>– распределение языковых единиц по регистрам;</li> <li>– принципы использования информации о коллокатах ключевого слова.</li> </ul> <p><b>Уметь:</b></p> <ul style="list-style-type: none"> <li>– использовать корпусную методологию для определения сочетаемости слов при коммуникации в письменной форме;</li> <li>– осуществлять поиск в устном подкорпусе текстов на английском языке для формирования навыков устной разговорной речи.</li> </ul> <p><b>Владеть:</b></p> <ul style="list-style-type: none"> <li>– навыками поиска и получения информации о языковых явлениях из корпусов текстов.</li> </ul>



### **Раздел 3. Использование корпусов**

#### **Тема 4. Сферы применения корпусной лингвистики**

Пользователи корпусов. Национальный корпус русского языка, British National Corpus, Longman-Lancaster Corpus, London-Lund Corpus. Корпусы текстов и лингводидактика (Skell). Международные организации, связанные с корпусной лингвистикой. Лексикографические исследования, основанные на корпусах. Грамматические исследования, основанные на корпусах. Исследования дискурса, основанные на корпусах.

#### 4.3. Перечень разделов/тем дисциплины

№ п/п	Наименование раздела/темы	Виды учебной работы, включая самостоятельную работу обучающихся, практическую подготовку (при наличии) и трудоемкость (в часах)				Оценочные средства	Формируемые компетенции (индикаторы)	Всего (в часах)
		Контактная работа преподавателя с обучающимися			СРС			
		Лекции	Практ. занятия	Лаб. занятия	(в том числе, внеауд. СР, КСР)			
1.	<b>Тема 1.</b> Основные понятия корпусной лингвистики	6	6		6	Лабораторная работа №1 Лабораторная работа №2	ИДК <sub>ПК1.1</sub> Устанавливает закономерности, принципы и особенности изучаемых явлений и процессов, анализирует базовые теории в предметной области; устанавливает закономерности, определяющие место предмета в общей картине мира.	18
2.	<b>Тема 2.</b> Типы корпусов	4	4		8	Лабораторная работа №3	ИДК <sub>ПК1.2</sub> Применяет системный анализ базовых научно-теоретических представлений для решения профессиональных	16

							задач.	
3.	<b>Тема 3.</b> Технологии корпусной лингвистики	10	10		10	Лабораторная работа №4 Лабораторная работа №5	ИДК <sub>ПК1.2</sub> Применяет системный анализ базовых научно-теоретических представлений для решения профессиональных задач.	30
4.	<b>Тема 4.</b> Сферы применения корпусной лингвистики	12	10		13	Лабораторная работа №6 Контрольная работа №1	ИДК <sub>ПК1.2</sub> Применяет системный анализ базовых научно-теоретических представлений для решения профессиональных задач.	35
	<b>ИТОГО (в часах)</b>	32	30		37			99

#### **4.4. Методические указания по организации самостоятельной работы студентов**

Методические рекомендации и учебно-методические материалы по организации СРС по дисциплине *Б1.В.ДВ.01.01 Корпусная лингвистика* представляют собой комплекс материалов и подходов, позволяющих студенту оптимальным образом организовать процесс изучения данной дисциплины в рамках СРС, а также рекомендации преподавателям по различным аспектам организации СРС.

Выполнение самостоятельной работы предусмотрено во внеаудиторной форме, без участия преподавателя, то есть вне наличия непосредственной обратной связи, корректирующих, контролирующих и оценивающих действий преподавателя. Преподаватель определяет цель, содержание деятельности, сроки выполнения заданий. Самостоятельность студента заключается в индивидуальном подходе при осуществлении заданного преподавателем объема и формата работы, в разработке методов решения поставленных проблем.

По каждой изучаемой теме предусмотрена самостоятельная работа студента, результаты которой напрямую или опосредованно подвергаются проверке и инкорпорированы в требования к уровню освоения дисциплины.

Внеаудиторная самостоятельная работа предполагает не только работу с текущим лекционным материалом и первоисточниками, но и индивидуальную учебно-исследовательскую работу, нацеленную на формирование навыков самостоятельного изучения исследуемой лингвистической проблематики. В этой связи выполнение заданий в ходе самостоятельной работы предусматривает разноаспектный формат работ, который включает в себя работу в научной библиотеке – в том числе чтение научных лингвистических журналов, а также работу с корпусами текстов в Интернете.

В процессе самостоятельной работы студенты выполняют все типы заданий: познавательные, практические, обобщающие, исследовательские.

Специфика дисциплины требует тщательного конспектирования лекций, что существенно облегчит самостоятельную работу студента с рекомендованной литературой при подготовке к практическим занятиям, тестам и зачету.

При освоении дисциплины «Корпусная лингвистика» в процессе самостоятельной подготовки к практическим занятиям студенту необходимо:

- тщательно проработать материал лекций и разделы рекомендованных учебников и научной литературы по соответствующим темам, составить краткий план выступления, уточнить необходимые термины в справочной литературе;
- обратить особое внимание на определения основных научных понятий, формулировки проблем и примеры решения практических задач, приводимых на лекциях и семинарах;
- при подготовке к практическим и семинарским занятиям использовать мультимедийные средства;
- следить за выполнением графика самостоятельной работы, не нарушать сроки отчетности.

#### **V. Учебно-методическое и информационное обеспечение дисциплины:**

Электронная информационно-образовательная среда университета обеспечивает доступ к электронным учебным изданиям и электронным образовательным ресурсам, указанным в рабочих программах дисциплин.

При использовании в образовательном процессе печатных изданий библиотечный фонд укомплектован печатными изданиями из расчета не менее 0,25 экземпляра каждого из изданий, указанных в рабочих программах дисциплин, на одного обучающегося из числа лиц, одновременно осваивающих соответствующую дисциплину.

Обучающимся обеспечен доступ к современным профессиональным базам данных

и информационным справочным системам, состав которых обновляется.

а) перечень литературы

**Захаров, В. П.**

Корпусная лингвистика [Текст] : учебник / В. П. Захаров, С. Ю. Богданова ; С.-Петербург. гос. ун-т. - 3-е изд., перераб. - СПб. : Изд-во С.-Петербург. ун-та, 2020. - 233 с. : ил., табл. ; 21 см. - Библиогр.: с. 211-225. - Предм. указ.: с. 231-233. - ISBN 978-5-288-05997-1 : (20 экз.)

**Богданова, С. Ю.**

Квантитативная лингвистика и НИТ [Текст] : учеб.-метод. пособие / С. Ю. Богданова ; Иркут. гос. ун-т, Пед. ин-т, Каф. иностр. яз. и лингводидактики. - Иркутск : Аспринт, 2017. - 87 с. : ил. ; 20 см. - Библиогр.: с. 82-86. - ISBN 978-5-4340-0173-1 : всего 46.

**Богданова, С. Ю.**

Квантитативная лингвистика и НИТ [Электронный ресурс] : учеб.-метод. пособие / С. Ю. Богданова. - ЭВК. - Иркутск : Аспринт, 2017. - 87 с. - Режим доступа: ЭЧЗ "Библиотех". - Неогранич. доступ. - ISBN 978-5-4340-0173-1.

**Гребенщикова, А. В.**

Основы квантитативной лингвистики и новых информационных технологий [Текст] : учеб. пособие / А. В. Гребенщикова. - М. : Флинта ; М. : Наука, 2014. - 151 с. ; 21 см. - Библиогр.: с. 147-151. - ISBN 978-5-9765-2140-7. - ISBN 978-5-02-038860-4 : 238.04 р. (10).

**Волков, В. В.**

Филология в системе современного гуманитарного знания [Электронный ресурс] : учебное пособие / В. В. Волков. - Электрон. текстовые дан. - Москва : Флинта, 2014. - 222 с. - ЭБС "Айбукс". - неогранич. доступ. - ISBN 978-5-9765-2015-8.

**Потапова, Р. К.**

Новые информационные технологии и лингвистика [Текст] : учеб. пособие для студ. вузов / Р. К. Потапова ; Московский гос. лингв. ун-т. - 5-е изд. - М. : Либроком, 2012. - 364 с. ; 21 см. - Библиогр.: с. 363-364. - ISBN 978-5-397-03206-3 : (10 экз) .

г) базы данных, информационно-справочные и поисковые системы

Национальный корпус русского языка – URL: <https://ruscorpora.ru>

Corpus of Contemporary American. – URL: <https://www.english-corpora.org/coca/>

Coronavirus Corpus. – URL: <https://www.english-corpora.org/corona/>

Sketch Engine for Language Learning. – URL: <https://www.sketchengine.eu/skell/>

Voyant Tools. – URL: <https://voyant-tools.org>

## VI. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 6.1. Помещения и оборудование

Помещения – учебные аудитории для проведения учебных занятий, предусмотренных учебным планом ОПОП ВО бакалавриата, оснащены оборудованием и техническими средствами обучения.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду ФГБОУ ВО «ИГУ».

#### Оборудование

В целом в Педагогическом институте ИГУ:

Количество персональных компьютеров – 920

Количество персональных компьютеров, имеющих доступ к сети Интернет – 831

Количество терминалов, с которых имеется доступ к сети Internet - 831

Количество компьютерных классов – 12

Имеется локальная сеть и Wi-fi.

Далее приведена подробная характеристика материально-технического оснащения аудитории 9п *Учебного корпуса №13* по адресу *664011, Иркутская область, г. Иркутск, ул. Сухэ-Батора, д. 9* Педагогического института ИГУ.

Компьютер Z-Comp CORE 2 Duo E740( Системный блок в комплекте, Монитор Samsung 7) - 30шт; коммутатор DGS-1005D/5-портовый, ноутбук HP-Compag 682s T2390, проектор Panasonic PT – LB30 NTE.

## 6.2. Лицензионное и программное обеспечение

Программное обеспечение: ОС: Windows XP (Номер Лицензии Microsoft 19683056), Антивирус Kaspersky (Форус Контракт № 04-114-16 от 14 ноября 2016 г. KES Счет № РСЦЗ-000147 и АКТ от 23 ноября 2016 г. Лиц. № 1В08161103014721370444).

## VII. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В образовательном процессе используются активные и интерактивные формы, в том числе дистанционные образовательные технологии, используемые при реализации различных видов учебной работы, развивающие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств и формирующие компетенции.

### Наименование тем занятий с использованием образовательных технологий

№	Тема занятия	Вид занятия	Форма / Методы интерактивного обучения	Кол-во часов
1	Корпусные менеджеры (Sketch Engine). Токенизация, лемматизация, парсинг.	Лекция	Видео-конференция	2
2	Принципы разметки текстов. Морфологическая разметка (POS-tagging). Семантическая разметка (UCREL). Синтаксическая разметка.	Практическое занятие	Онлайн-семинар	4
3	Корпусы текстов и лингводидактика (работа с корпусом Skell)	Практическое занятие	Творческое задание	2
Итого часов				8

Учебный процесс опирается на использование интерактивных методов обучения, организуется с учетом включенности в процесс познания всех студентов группы, каждый вносит свой особый индивидуальный вклад, в ходе работы идет обмен знаниями, идеями, способами деятельности с обязательной обратной связью и возможностью взаимной оценки и контроля. Интерактивные методы пробуждают у обучающихся интерес, поощряют активное участие каждого в учебном процессе, обращаются к чувствам каждого обучающегося, способствуют эффективному усвоению учебного материала, оказывают многоплановое воздействие на обучающихся, осуществляют обратную связь

(ответная реакция аудитории), формируют у обучающихся мнения и отношения, жизненные навыки, способствуют изменению поведения.

Учебный процесс строится с учетом возможности доступа к современным корпусным базам данных и направлен на формирование у обучающихся практических навыков работы с существующими системами представления знаний, в частности, создание исследовательского корпуса текстов для решения лингводидактических задач.

## **VIII. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ**

### **8.1. Оценочные средства для проведения текущего контроля успеваемости Устный опрос**

Вопросы:

- 1) Дайте определения терминов: корпус, корпусная лингвистика, разметка, репрезентативность
- 2) Перечислите типы корпусов.
- 3) Какой корпус текстов был первым? Укажите его основные характеристики.
- 4) Какие два типа корпусов можно выделить по критерию репрезентативности? В чем их отличие?
- 5) Каково основное назначение параллельных корпусов текстов?
- 6) В чем заключается сложность создания корпусов устной речи?

#### **Демонстрационный вариант Письменного опроса**

1. Перечислите технологии в разных отраслях лингвистики, повлиявшие на развитие современной корпусной лингвистики
2. Как осуществляется предварительная обработка текстов при создании корпуса?
3. В чем заключаются сложности составления корпусов устной речи?
4. Как можно различить синонимичные слова с помощью корпусов текстов?
5. Что подразумевается под «репрезентативностью корпуса»?
6. Дайте определения токенизации, лемматизации, парсинга.

#### **Демонстрационный вариант Лабораторной работы № 3**

#### **БРИТАНСКИЙ НАЦИОНАЛЬНЫЙ КОРПУС (BNC)**

<http://corpus.byu.edu>

Задание 1.

(1) Создайте файл и назовите его «Фамилия ЛР 3». Зайдите на сайт <http://corpus.byu.edu>. Выберите в списке корпус British National Corpus. Зарегистрируйтесь и сохраните логин и пароль. Прочитайте раздел Введение (Introduction) и внизу нажмите на значок “five minute guided tour” (демонстрационный тур), чтобы узнать больше о корпусе.

(2) Выберите слово *mysterious* в демонстрационном туре. Слово появится во вкладке FREQUENCY вверху экрана. Нажмите на вкладку SEARCH слева. Над словом *mysterious* Вы увидите в кнопку “List”. Нажмите расположенную рядом кнопку “Chart” (таблица). Нажмите кнопку “Section” (раздел) под окном поиска, там должно быть написано “Ignore” (не важно/любой). Нажмите на кнопку “See frequency by section” (посмотреть частотность по разделам/регистрам). Таблица показывает, что слово *mysterious* наиболее частотно в регистре «Художественная литература» (Fiction) (27,22 tokens (instances) per 1 million words), следом идет «Разное» (Miscellaneous), затем регистры Newspapers, Academic and Spoken. Скопируйте таблицу в свой файл.

(3) Нажмите на синюю полосу рядом с Fiction или на голубой столбик внизу таблицы. Это зависит от того, в горизонтальном или вертикальном меню вы находитесь. Там будут показаны первые 100 строк конкорданса со словом *mysterious*, и Вы сможете изучить контексты, в которых встречается данное прилагательное. Скопируйте первые 10 результатов в свой файл. Ответьте в своем файле письменно на вопрос: Какие типы существительных определяет прилагательное *mysterious*?

(4) Вернитесь к панели поиска слева путем выбора Search и нажмите на “List” (список) вместо “Chart” (таблица). Нажмите на надпись “Find matching strings” (найти соответствующие цепочки). Теперь Вы видите количество употреблений справа на панели. Если Вы нажмете на слово *mysterious*, то получите первые 100 строк конкорданса во всем корпусе, а не только в подкорпусе Fiction. Скопируйте первые 10 результатов в свой файл.

### Демонстрационный вариант Контрольной работы № 1

Выберите правильный ответ из правой колонки

1.	Какое определение корпуса является наиболее точным?	<p>а) информационно-справочная система, основанная на собрании текстов на некотором языке в электронной форме.</p> <p>б) репрезентативное собрание текстов, обычно в читаемом машиной формате и включающем информацию о ситуации, в которой текст был произведен, такую как информация о говорящем, авторе, адресате или аудитории.</p> <p>в) своеобразного рода словесное единство, расположенное на машинном носителе, стандартизовано представляющее словесный материал, позволяющее применять стандартные программы его обработки, характеризующееся конечным размером и репрезентативностью как результатом особой процедуры отбора.</p>
2.	Что представляет собой корпусная лингвистика?	<p>а) отрасль компьютерной лингвистики, занимающаяся составлением корпусов текстов для лексикографических целей</p> <p>б) деятельность, требующаяся для составления и использования корпуса, направленная на исследование естественного использования языка</p> <p>в) вид деятельности, связанной с использованием текстов для того, чтобы ответить на вопросы о использовании языка</p>
3	Какой корпус содержит только устную речь?	<p>а) London-Lund Corpus</p> <p>б) Longman-Lancaster Corpus</p> <p>в) British National Corpus</p>
4.	Какой минимальный объем считается репрезентативным для национального корпуса языка?	<p>а) 160 млн. словоупотреблений</p> <p>б) 1 млн. словоупотреблений</p> <p>в) 100 млн. словоупотреблений</p>
5.	Что обозначает термин «коллокация» в корпусной	<p>а) любое сочетание слов</p> <p>б) устойчивое сочетание слов</p>

	лингвистике?	в) сочетание слов, расположенных рядом в тексте
6.	Области лингвистических исследований, повлиявшие на развитие современной корпусной лингвистики	а) историческая лингвистика, грамматика, социолингвистика б) когнитивная лингвистика, лексикография, социолингвистика в) фонология, лексикография, лингводидактика
7.	Какие данные можно получить из корпусов текстов?	а) суждения о грамматике, экстралингвистическая информация б) экспериментальные данные, данные полевых исследований в) эмпирическая поддержка, информация по частотности, метаинформация (экстралингвистическая информация)
8.	Что такое конкорданс?	а) указатель, связывающий каждое употребление с контекстом б) список встречающихся в тексте словоформ, расположенных в алфавитном порядке в) тип словаря
9.	Что такое token?	а) словоформа б) начальная форма слова в) основа слова
10	Какой процесс устанавливает лексему, парадигма которой включает данную словоформу?	а) стемминг б) токенизация в) лемматизация
11	Для чего нужны нормированные подсчеты?	а) они указывают частотность встречаемости определенной словоформы б) они преобразуют количество случаев встречаемости слова по стандартной шкале в) они применяются в лексикографии для выявления новых значений слов
12	Чтобы слова можно было ассоциировать друг с другом, они должны:	а) примыкать друг к другу или располагаться на некотором расстоянии б) располагаться на некотором расстоянии в) примыкать друг к другу
13	В каком из регистров чаще встречаются номинализации с суффиксом –ness?	а) в академической прозе б) в художественной литературе в) в устной речи
14	Для каких исследований можно использовать неаннотированные корпусы?	а) в области морфологии б) в области грамматических категорий в) в области синтаксиса
15	Какой из регистров демонстрирует предпочтение к статической, а не к динамической упаковке информации, отраженной в большей частотности номинализаций и более высоком соотношении	а) устная речь б) художественная литература в) академическая проза

## Ключи

1.б 2.б 3.а 4.в 5.в 6.а 7.в 8.а 9.а 10.в 11.б 12.а 13.б 14.а 15.в

## 8.2. Оценочные средства для промежуточной аттестации

1. Вид промежуточной аттестации - зачет.
2. Форма проведения - устный ответ.

### Примерный перечень вопросов к зачету:

1. Определения корпусной лингвистики.
2. Основные понятия корпусной лингвистики: корпус текстов, разметка, конкорданс, токен, лемма, коллокация.
3. Традиционные и современные методы лингвистических исследований корпусов.
4. Зависимость валидности данных от предоставляемой информации.
5. Принцип репрезентативности, его применение на корпусах разных типов.
6. Первые компьютерные корпуса.
7. Брауновский корпус США.
8. LOB корпус в Европе.
9. Основные типы корпусов.
10. Корпусы английского и русского языков
11. Классификация корпусов по хронологии, по жанрам, по аннотированности, по целям и задачам, по объему, по языкам.
12. Технологии в разных отраслях лингвистики, повлиявшие на развитие современной корпусной лингвистики.
13. Конструирование корпуса.
14. Основные процедуры предварительной обработки корпусов. Токенизация, лемматизация, парсинг.
15. Принципы разметки текстов.
16. Конкурирующие технологии (rule-based), статистические, основанные на ограничениях (restriction-based), гибридные).
17. Национальный корпус русского языка.
18. British National Corpus.
19. Longman-Lancaster Corpus, London-Lund Corpus.
20. Международные организации, связанные с корпусной лингвистикой.

### Критерии оценивания

Оценки «отлично» заслуживает студент, обнаруживший всестороннее, систематическое и глубокое знание учебного материала дисциплины, умение свободно выполнять задания, предусмотренные программой, усвоивший основную литературу. Как правило, оценка «отлично» выставляется усвоившим взаимосвязь основных понятий дисциплины и их значение для приобретаемой профессии, проявившим творческие способности в понимании, изложении и использовании учебно-программного материала.

Оценки «хорошо» заслуживает студент, обнаруживший полное знание учебного материала, успешно выполняющий предусмотренные в программе задания, усвоивший основную литературу, рекомендованную в программе. Как правило, оценка «хорошо» выставляется учащимся, показавшим систематический характер знаний по дисциплине и способным к их самостоятельному пополнению и обновлению в ходе дальнейшей учебной работы и профессиональной деятельности.

Оценки «удовлетворительно» заслуживает студент, обнаруживший знание основного учебного материала в объеме, необходимом для дальнейшей учебы и предстоящей работы по профессии, справляющийся с выполнением заданий, предусмотренных программой,

знакомый с основной литературой, рекомендованной программой. Как правило, оценка «удовлетворительно» выставляется, допустившим погрешности в ответе на экзамене и при выполнении экзаменационных заданий, но обладающим необходимыми знаниями для их устранения под руководством преподавателя.

Документ составлен в соответствии с требованиями ФГОС по направлению 44.03.05 Педагогическое образование (с двумя профилями подготовки) утвержденного приказом Минобрнауки РФ об утверждении ФГОС ВО № 125 от 22 февраля 2018 г.

**Разработчик:** Богданова Светлана Юрьевна, профессор

*Настоящая программа не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.*

**Педагогический институт**

(отделение гуманитарно-эстетического образования)

**ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ**

**44.03.05 Педагогическое образование (с двумя профилями подготовки), Иностранный язык (английский) – Иностранный язык (китайский) 4 курс**

(наименование образовательной программы: направление, профиль, курс)

Тест

**Формулировка компетенции:** ПК-1. Способен осваивать и использовать базовые научно-теоретические знания и практические умения по предмету в профессиональной деятельности

Номер задания	Содержание вопроса	Ответы Правильный ответ выделен жирным шрифтом	Компетенция (шифр компете нции из УП)	Код и наименование дисциплины, формирующей данную компетенцию												
<p align="center"><b>1. Задание закрытого типа на установление соответствия (проверяется автоматически)</b>  <b>Инструкция:</b> <i>Установите соответствие. К каждой позиции левого столбца выберите позицию в правом столбце. Один вариант в правом столбце лишней.</i></p>																
1.	<p>Установите соответствие между направлением в лингвистике и ученым, принадлежащим к этому направлению.</p> <table border="1" data-bbox="360 936 869 1435"> <thead> <tr> <th data-bbox="360 936 603 1032">Направление в лингвистике</th> <th data-bbox="603 936 869 1032">Ученый</th> </tr> </thead> <tbody> <tr> <td data-bbox="360 1032 603 1151">1. Сравнительно-историческое языкознание</td> <td data-bbox="603 1032 869 1151">А Лев Владимирович Щерба.</td> </tr> <tr> <td data-bbox="360 1151 603 1211">2. Грамматика</td> <td data-bbox="603 1151 869 1211">В Герман Пауль</td> </tr> <tr> <td data-bbox="360 1211 603 1308">3. Корпусная лингвистика</td> <td data-bbox="603 1211 869 1308">С Якоб Гримм</td> </tr> <tr> <td data-bbox="360 1308 603 1368">4. Лексикография</td> <td data-bbox="603 1308 869 1368">D Дуглас Байбер</td> </tr> <tr> <td data-bbox="360 1368 603 1435"></td> <td data-bbox="603 1368 869 1435">E Джеймс Мюррей</td> </tr> </tbody> </table>	Направление в лингвистике	Ученый	1. Сравнительно-историческое языкознание	А Лев Владимирович Щерба.	2. Грамматика	В Герман Пауль	3. Корпусная лингвистика	С Якоб Гримм	4. Лексикография	D Дуглас Байбер		E Джеймс Мюррей	<p><b>1-С</b> <b>2-В</b> <b>3-D</b> <b>4-E</b></p>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика
Направление в лингвистике	Ученый															
1. Сравнительно-историческое языкознание	А Лев Владимирович Щерба.															
2. Грамматика	В Герман Пауль															
3. Корпусная лингвистика	С Якоб Гримм															
4. Лексикография	D Дуглас Байбер															
	E Джеймс Мюррей															
2.	<p>Установите соответствие между процедурой обработки естественного языка и ее определением.</p> <table border="1" data-bbox="360 1570 869 2078"> <thead> <tr> <th data-bbox="360 1570 580 1666">Процедура обработки</th> <th data-bbox="580 1570 869 1666">Определение</th> </tr> </thead> <tbody> <tr> <td data-bbox="360 1666 580 1827">1. Лемматизация</td> <td data-bbox="580 1666 869 1827">А. разделение потока символов в текстах на естественном языке на отдельные значимые единицы (токены).</td> </tr> <tr> <td data-bbox="360 1827 580 1946">2. Токенизация</td> <td data-bbox="580 1827 869 1946">В. процесс образования начальных форм для словоформ.</td> </tr> <tr> <td data-bbox="360 1946 580 2078">3. Парсинг</td> <td data-bbox="580 1946 869 2078">С. приписывание компонентам текста морфологической информации.</td> </tr> </tbody> </table>	Процедура обработки	Определение	1. Лемматизация	А. разделение потока символов в текстах на естественном языке на отдельные значимые единицы (токены).	2. Токенизация	В. процесс образования начальных форм для словоформ.	3. Парсинг	С. приписывание компонентам текста морфологической информации.	<p><b>1-В</b> <b>2-А</b> <b>3-Е</b> <b>4-С</b></p>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика				
Процедура обработки	Определение															
1. Лемматизация	А. разделение потока символов в текстах на естественном языке на отдельные значимые единицы (токены).															
2. Токенизация	В. процесс образования начальных форм для словоформ.															
3. Парсинг	С. приписывание компонентам текста морфологической информации.															

	4. POS-tagging	D отношение абсолютной частоты слова (словоформы) к объему корпуса (в словоупотреблениях).															
		E анализ синтаксической структуры предложения и представления ее в виде дерева зависимостей или структуры составляющих.															
3.	Установите соответствие между названием корпуса текстов и его объемом.		<b>1-С</b> <b>2-А</b> <b>3-В</b> <b>4-Е</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01</i> <i>Корпусная лингвистика</i>												
	<table border="1"> <thead> <tr> <th>Название корпуса</th> <th>Объем в словоупотреблениях (с/у)</th> </tr> </thead> <tbody> <tr> <td>1. BNC</td> <td>A 500 000 с/у</td> </tr> <tr> <td>2. London-Lund</td> <td>B более 1 000 000 000 с/у</td> </tr> <tr> <td>3. COCA</td> <td>C 100 000 000 с/у</td> </tr> <tr> <td>4. The Brown Corpus</td> <td>D 3 000 000 с/у</td> </tr> <tr> <td></td> <td>E 1 000 000 с/у</td> </tr> </tbody> </table>	Название корпуса	Объем в словоупотреблениях (с/у)	1. BNC	A 500 000 с/у	2. London-Lund	B более 1 000 000 000 с/у	3. COCA	C 100 000 000 с/у	4. The Brown Corpus	D 3 000 000 с/у		E 1 000 000 с/у				
Название корпуса	Объем в словоупотреблениях (с/у)																
1. BNC	A 500 000 с/у																
2. London-Lund	B более 1 000 000 000 с/у																
3. COCA	C 100 000 000 с/у																
4. The Brown Corpus	D 3 000 000 с/у																
	E 1 000 000 с/у																
4.	Установите соответствие между типом корпуса и основной процедурой в нем.		<b>1- С</b> <b>2- А</b>														
	<table border="1"> <thead> <tr> <th>Типы корпусов</th> <th>Основная процедура</th> </tr> </thead> <tbody> <tr> <td>1. Параллельные корпуса</td> <td>A создание транскриптов</td> </tr> <tr> <td>2 Устные</td> <td>B разметка ошибок</td> </tr> <tr> <td></td> <td>C выравнивание</td> </tr> </tbody> </table>	Типы корпусов	Основная процедура	1. Параллельные корпуса	A создание транскриптов	2 Устные	B разметка ошибок		C выравнивание								
Типы корпусов	Основная процедура																
1. Параллельные корпуса	A создание транскриптов																
2 Устные	B разметка ошибок																
	C выравнивание																
<b>2. Задание закрытого типа «верно-неверно» (проверяется автоматически)</b>																	
<b>Инструкция:</b> <i>Прочитайте утверждения и отметьте их как верные или неверные</i>																	
5.	Укажите, является данное утверждение правильным (верно) или неправильным (неверно) <b>1</b> Корпусная лингвистика — раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий. <b>2</b> Достаточно большой (репрезентативный) и сбалансированный объем корпуса не		<b>1 верно</b> <b>2 неверно</b> <b>3 верно</b> <b>4 верно</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01</i> <i>Корпусная лингвистика</i>												

	<p>гарантирует типичность данных и полноту представления всего спектра языковых явлений.</p> <p><b>3</b> Корпусы могут использоваться для получения справок о характеристиках текста или лексических единиц, статистических данных о языковых единицах и о лингвистических категориях и метаданных.</p> <p><b>4</b> Корпусная лингвистика имеет своим предметом теоретические основы и практические механизмы создания и использования представительных массивов языковых данных.</p>			
6.	<p>Укажите, является данное утверждение правильным (верно) или неправильным (неверно)</p> <p><b>1</b> The Brown Corpus был создан для отражения письменной речи Великобритании начала 1960 х годов с удовлетворительной для того времени степенью репрезентативности.</p> <p><b>2</b> Национальный корпус представляет данный язык на определенном этапе (или этапах) его существования во всем многообразии жанров, стилей, территориальных и социальных вариантов, временных периодов</p> <p><b>3</b> LOB расшифровывается как The London-Oslo-Bergen Corpus</p> <p><b>4</b> Собственно лингвистическая разметка описывает сведения об авторе и о тексте: автор, название, год и место издания, жанр, тематика</p>	<p><b>1 неверно</b>  <b>2 верно</b>  <b>3 неверно</b>  <b>4 неверно</b></p>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика
<p><b>3. Задание с выбором одного правильного ответа из предложенных</b>  <b>Инструкция: Прочитайте предложение и выберите один правильный ответ</b></p>				
7.	<p>Для чего нужны нормированные подсчеты?</p> <p><b>А</b> они указывают частотность встречаемости определенной словоформы</p> <p><b>В</b> они преобразуют количество случаев встречаемости слова по стандартной шкале</p> <p><b>С</b> они применяются в лексикографии для выявления новых значений слов</p>	<b>В</b>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика
8.	<p>Какой минимальный объем считается репрезентативным для национального корпуса языка?</p> <p><b>А</b> 160 млн. словоупотреблений</p> <p><b>В</b> 1 млн. словоупотреблений</p> <p><b>С</b> 100 млн. словоупотреблений</p>	<b>С</b>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика
9.	<p>В каком из регистров чаще встречаются номинализации с суффиксом –ness?</p> <p><b>А</b> в академической прозе</p> <p><b>В</b> в художественной литературе</p> <p><b>С</b> в устной речи</p>	<b>В</b>	ПК-1	Б1.В.ДВ.01.01 Корпусная лингвистика

4. Задания открытого типа с кратким ответом/ вставить термин, словосочетание, дополнить предложенное				
<b>Инструкция:</b> Прочитайте текст и запишите пропущенное слово/словосочетание (термин, словосочетание, дополните предложенное)				
10.	Вставьте пропущенный термин, слово или словосочетание  KWIC – Key ____ In Context	<b>Word / word</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>
11.	Вставьте пропущенный термин, слово или словосочетание  ipm – items per _____	<b>million</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>
12.	Вставьте пропущенный термин, слово или словосочетание NLP – Natural _____ processing.	<b>Language / language</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>
13.	Вставьте пропущенный термин, слово или словосочетание  Теггер (грамматический разметчик) — программа, выполняющая в автоматическом режиме грамматическую (морфологическую) _____ текстов корпуса.	<b>разметку</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>
14.	Вставьте пропущенный термин, слово или словосочетание  Корпусная лингвистика – раздел _____ лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов с использованием компьютерных технологий.	<b>компьютерной</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>
15.	Вставьте пропущенный термин, слово или словосочетание  Корпус – репрезентативное собрание текстов, обычно в _____ формате, включающем информацию о ситуации, в которой текст был произведен, такую как информация о говорящем, авторе, адресате или аудитории.	<b>машиночитае мом</b>	<i>ПК-1</i>	<i>Б1.В.ДВ.01.01 Корпусная лингвистика</i>

**Время на выполнение тестового задания: 40 минут. (Рекомендованное разработчиком)**