МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный экономический университет»

УТВЕРЖДАЮ
Проректор по научной работе
_____/Гербашко Е.А./
«28» августа 2020 г.

УТВЕРЖДАЮ
Проректор по учебной и методической работе
_____/Шубаева В.Г./
«28» августа 2020 г.

КВАНТИТАТИВНЫЕ МЕТОДЫ В ЛИНГВИСТИКЕ

Рабочая программа дисциплины

Направление подготовки	45.06.01 Языкознание и литературоведение
Направленность (профиль) программы	Языкознание
Уровень высшего образования	подготовка кадров высшей квалификации
Форма обучения	очная

Составитель:

_____ д.ф.н., доцент Клепикова Т.А.

Санкт-Петербург
2020

СОДЕРЖАНИЕ

1. ЦЕЛЬ И ЗАДАЧИ ДИСЦИПЛИНЫ....3
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ 3
3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ....3
4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ....4
5. СОДЕРЖАНИЕ РАЗДЕЛОВ И ТЕМ ДИСЦИПЛИНЫ....4
6. ЗАНЯТИЯ СЕМИНАРСКОГО ТИПА....5
7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ....6
7.1. Методические указания для обучающегося по освоению дисциплины....6
7.2. Организация самостоятельной работы....6
8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ....7
9. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ....7
9.1. Учебно-методическое и информационное обеспечение дисциплины....7
9.2. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА....8
10. ОСОБЕННОСТИ ОСВОЕНИЯ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ....9
11. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ....9

1. ЦЕЛЬ И ЗАДАЧИ ДИСЦИПЛИНЫ

Цель дисциплины: углубление знаний и совершенствование исследовательских навыков аспирантов направления подготовки 45.06.01 – Языкознание и литературоведение в области квантитативных методов анализа языковых и речевых явлений.

Задачи:

- формирование представления о лингвистике как о доказательной науке, обладающей соответствующей методологией анализа;
- углубление знаний о предмете, задачах, методах современной квантитативной лингвистики;
- формирование системы знаний, умений и навыков, позволяющих проводить научно-обоснованный и доказательный анализ языка с применением квантитативных методов;
- освоение методологии доказательной лингвистики, опирающейся на квантитативные методы (статистико-вероятностные методы, корпусные технологии, компьютерные методы синтаксического, морфологического, лексикографического анализа).
- совершенствование владения информационными технологиями для решения научно-исследовательских задач лингвистического плана.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина Б1.В «Квантитативные методы в лингвистике» относится к вариативной части Блока 1, является обязательной для освоения обучающимся после выбора обучающимся направленности (профиля) программы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы представлены в таблице 3.1.

Таблица 3.1 – Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

Код и наименование компетенции выпускника	Уровень освоения компетенции	Планируемые результаты обучения (показатели освоения компетенции)
ОПК-1 способностью самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий	*Второй уровень (углубленный)* ***(ОПК-1)-2***	***Декомпозиция III*** ***Знать:*** *теоретические и методологические основания корпусной лингвистики; основы применения методов лингвистической статистики и корпусной лингвистики при проведении междисциплинарных исследований З2(III) (ОПК-1)-2* ***Уметь****: проводить междисциплинарные исследования с применением методов лингвистической статистики и корпусной лингвистики У2(III) (ОПК-1)-2* ***Владеть****: навыками проведения междисциплинарных исследований с применением методов лингвистической статистики и корпусной лингвистики В2 (III) (ОПК-1)-2*

ПК-3 Способность использовать методологию лингвистического исследования в научно-исследовательской деятельности	ПК-3	***Декомпозиция III*** ***Знать:*** *принципы доказательной лингвистики, основы применения статистико-вероятностных методов в лингвистике З(III) (ПК-3)* ***Уметь:*** *определять пути решения лингвистических задач с использованием релевантных методов лингвистической статистики и корпусной лингвистики У(III) (ПК-3)* ***Владеть:*** *навыками анализа фактов языка с использованием релевантных методов лингвистической статистики и корпусной лингвистики В (III) (ПК-3)*

4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 1 зачетную единицу, 36 часов.

Форма промежуточной аттестации: *зачет – 2 год обучения.*

Распределение фонда времени по темам дисциплины по очной форме обучения представлено в таблице 4.1.:

Таблица 4.1 – Распределение фонда времени по темам дисциплины

Номер и наименование тем *и/или разделов/тем*	Объем дисциплины (ак. часы)			
	Контактная работа			СРО
	ЗЛТ	ПЗ	ЛР	
1	*2*	*3*	*4*	*5*
Тема 1. Лингвистика как номотетическая наука	2	2	-	5
Тема 2. Основные цели, задачи, категории лингвистической статистики.	2	2	-	5
Тема 3. Корпусы и методика корпусного анализа	2	2	-	5
Тема 4. Создание корпуса. Конкордансеры	2	2	-	5
Всего по дисциплине:	**8**	**8**		**20**

*ЗЛТ – занятия лекционного типа, ПЗ – все виды занятий семинарского типа, кроме лабораторных работ, ЛР – лабораторные работы, СРО – самостоятельная работа обучающегося.

5. СОДЕРЖАНИЕ РАЗДЕЛОВ И ТЕМ ДИСЦИПЛИНЫ

Раздел 1. Принципы доказательной лингвистики

Тема 1. Лингвистика как номотетическая наука

Предмет, задачи, методология, направления квантитативной лингвистики. Доказательная парадигма в лингвистике. Лингвистика как номотетическая наука. Общенаучный метод моделирования в лингвистике. Особенности номотетических моделей системы языка и речемыслительной деятельности человека. Гипотетические и воспроизводящие модели. Воспроизводящие инженерно-лингвистические модели. Основные этапы решения лингвистических задач методом моделирования. Эпистемиологическая и эмпирическая верификация лингвистических моделей. Экспертные системы. Системы представления знаний. Лингвистическая составляющая проблемы создания экспертных систем.

Аксиоматические теории, исчисления, формальные грамматики. «Барьер отторжения» между естественным языком и «языком» компьютера.

Тема 2. Основные цели, задачи, категории лингвистической статистики

Предмет, цели, задачи, методология лингвистической статистики (вычислительная, квантитативная, количественная, статистическая лингвистика). Цели и организация лингвостатистических исследований. Основные понятия теории вероятностей и математической статистики. Теория множеств и её расширение для нечётких множеств. Основные статистические категории: выборка и совокупности, типы переменных, их классификация применительно к соответствующему уровню лингвистического анализа. Лингвистическая переменная.

Планирование статистического эксперимента и проверка гипотез. Параметрические и непараметрические критерии значимости. Корреляционный, дисперсионный, кластерный, факторный анализ в лингвистических исследованиях. Меры центральной тенденции и рассеяния. Выборочные статистики и параметры распределения: методы оценивания.

Направления лингвостатистического анализа. Квантитативный анализ в лексике, грамматике, морфологии, дискурсе, фоностатистика. Квантитативные методы в социолингвистике, сравнительной типологии и других областях лингвистического анализа.

Частота как характеристика употребительности единицы в тексте. Закон Ципфа-Мандельброта-Пиотровского. Частотные словари, словари-индексы, машинные фонды лексики. Характеристики частотных словарей и индексы покрытия текста. Частота слова и продуктивность. Частота слова и многозначность. Частота слова и его возраст. Лингвистические автоматы. Автоматические словари слов и оборотов в лингвистических автоматах разной степени сложности и системах машинного перевода. Методы отбора лексики и процедуры лексико-морфологического анализа текста при машинном переводе и автоматической обработке информации.

Атрибуция текстов и стилеметрия. Контент-анализ, анализ метафоры. Социолингвистические и психолингвистические переменные, их оценивание. Психометрические методы. Лингвистические идентификация личности при помощи квантитативных методов.

Раздел 2. Корпусные технологии в лингвистике

Тема 3. Корпусы и методика корпусного анализа

Корпус как поисковая система. Базы данных и корпусы: принципы дифференциации. Создание и типология корпусов. Первое и второе поколение корпусов. Типы корпусов: устные и письменные, одноязычные и многоязычные, аннотированные и неаннотированные. Параллельные корпусы.

Принципы создания корпусов. Дизайн корпуса. Корпусоподобные интерфейсы. Репрезентативность корпуса. Методы извлечения информации из корпуса. Типы извлекаемой информации. Метаданные. Разметка. Виды разметок: частеречная разметка, грамматический парсинг, просодическая аннотация, анафорическая аннотация, семантическая разметка и под. Treebank. Лемма, лемматизация. Конкорданс. Корпусный менеджер. Стандартизация в корпусной лингвистике. Международные стандарты.

Меры связанности единиц: MI, log-likelihood, t-test, z-test.

Тема 4. Создание корпуса. Конкордансер.

Конкордансер как поисковая система. Конкордансеры WordSmith, AntConc и AntPConc. Принципы работы с конкордансером. Создание корпуса. Анализ данных. Формирование запроса. Извлечение статистически релевантной информации о морфологических, лексических, синтаксических, стилистических и иных явлениях в языке и речи.

6. ЗАНЯТИЯ СЕМИНАРСКОГО ТИПА

Таблица 6.1 – Практические занятия/ Семинарские занятия / Лабораторные работы

№ темы	Тема занятия	Вид занятия/Оценочное средство
1	*2*	*3*
1	Лингвистика как номотетическая наука.	СЗ/ Обзор современных источников по теме
2	Основные цели, задачи, категории лингвистической статистики	СЗ/ Устная презентация по теме
3	Корпусы и методика корпусного анализа.	СЗ/ Решение практических задач
4	Создание корпуса. Конкордансеры.	СЗ/ Решение практических задач

* ПЗ – практические занятия, СЗ – семинарские занятия, ЛР – лабораторные работы

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

7.1. Методические указания для обучающегося по освоению дисциплины

Для формирования четкого представления об объеме и характере знаний и умений, которыми надо будет овладеть по дисциплине в самом начале учебного курса обучающийся должен ознакомиться с учебно-методической документацией:

– рабочей программой дисциплины: с целями и задачами дисциплины, ее связями с другими дисциплинами образовательной программы, перечнем знаний и умений, которыми в процессе освоения дисциплины должен владеть обучающийся,

– порядком проведения текущего контроля успеваемости и промежуточной аттестации;

– графиком консультаций преподавателей кафедры.

Систематическое выполнение учебной работы на занятиях лекционных и семинарских типов, а также выполнение самостоятельной работы позволит успешно освоить дисциплину.

В процессе освоения дисциплины обучающимся следует:

– слушать, конспектировать излагаемый преподавателем материал;

– ставить, обсуждать актуальные проблемы курса, быть активным на занятиях;

– задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений;

– выполнять задания практических занятий полностью и установленные сроки.

При затруднениях в восприятии материала следует обратиться к основным литературным источникам. Если разобраться в материале не удалось, то обратится к лектору (по графику его консультаций) или к преподавателю на занятиях семинарского типа.

Обучающимся, пропустившим занятия (независимо от причин), не имеющим письменного решения задач или не подготовившимся к данному занятию, рекомендуется не позже чем в 2 - недельный срок явиться на консультацию к преподавателю и отчитаться по теме.

7.2. Организация самостоятельной работы

Под самостоятельной работой обучающихся понимается планируемая работа обучающихся, направленная на формирование указанных компетенций, выполняемая во внеаудиторное время по заданию и при методическом руководстве преподавателя, без его непосредственного участия.

Методическое обеспечение самостоятельной работы при наличии обучающихся лиц с ограниченными возможностями представляется в формах, адаптированных к ограничениям их здоровья.

Виды самостоятельной работы по дисциплине представлены в таблице 7.2.1.

Таблица 7.2.1 – Организация самостоятельной работы обучающегося

№ темы	Вид самостоятельной работы
1	*2*
1	Подготовка к тематической дискуссии. Подготовка к зачету.
2	Подготовка к зачету. Написание научного исследования по предложенной теме.
3	Решение практической задачи. Подготовка к зачету.
4	Решение практической задачи. Подготовка к зачету.

Каждый вид СРО, указанный в таблице 7.2.1 обеспечен методическими материалами.

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В преподавании дисциплины «*Квантитативные методы в лингвистике*» используются разнообразные образовательные технологии как традиционные, так и с применением активных и интерактивных методов обучения.

Активные и интерактивные методы обучения:

– лекция-дискуссия (тема №1, №4);
– проблемная лекция (тема №2);
– обзор современных источников по теме (тема №1);
– устная презентация по теме (тема №1)
– тематическая дискуссия (тема № 2);
– проблемный семинар (тема № 3);
– проблемный доклад (тема № 4)
– решение практических задач (тема № 3, тема № 4).

Лекция-дискуссия: преподаватель при изложении лекционного материала не только использует ответы слушателей на его вопросы, но и организует свободный обмен мнениями в интервалах между логическими разделами.

Проблемная лекция: преподаватель ставит вопрос или формулирует проблемную задачу и показывает варианты ответов или способов решения, а студенты наблюдают за поиском и определяют свое отношение к полученному материалу.

Проблемный семинар: предполагает активное вовлечение студентов в процесс обсуждения и решения поставленной проблемы, процесс поиска решения направляется и контролируется преподавателем.

Тематическая дискуссия: обсуждение темы (спорного или проблемного характера) в учебной группе. Вопросы для проблемной дискуссии предлагаются учащимся заранее.

Проблемный доклад: предполагает активное участие студентов в процессе обсуждения изложенного одним из студентов теоретического материала, ведение активной дискуссии. Процесс обсуждения проблемы направляется и контролируется преподавателем.

9. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

9.1. Учебно-методическое и информационное обеспечение дисциплины

Таблица 9.1.1 – Учебно-методическое обеспечение дисциплины

Библиографическое описание издания (автор, заглавие, вид, место и год издания, кол. стр.)	Основная/ дополнительная литература	Книгообеспеченность	
		Кол-во. экз. в библ. СПбГЭУ	Электронные ресурсы
1. Клепикова Т.А. Studying lexis with corpora: a	основная	33	ЭБ

glossary: учебное пособие по корпусной лингвистике / Т.А.Клепикова ; М-во образования и науки Рос. Федерации, С.-Петерб. гос. экон. ун-т, Каф. англ. яз. и пер. - Санкт-Петербург : Изд-во СПбГЭУ, 2013. - 102 с.			OPAC.UNECON.RU
2. Трошина А.В. Информационные технологии в лингвистике : учебное пособие / А.В.Трошина, Н.К.Генидзе. — Санкт-Петербург : Изд-во СПбГЭУ, 2017 .— 88 с. : ил. — Сведения доступны также по Интернету: opac.unecon.ru.	основная	35	ЭБ OPAC.UNECON.RU
3. Волосатова Т.М. Информатика и лингвистика [Электронный ресурс] : Учебное пособие .— 1 .— Москва : ООО "Научно-издательский центр ИНФРА-М", 2018 .— 196 с. — ВО - Бакалавриат .— ISBN 978-5-16-010977-0 .	дополнительная	-	ЭБС ZNANIUM
4. Алпатов В. Языкознание: От Аристотеля до компьютерной лингвистики. — Москва : ООО "Альпина нон-фикшн", 2018 .— 253 с.	дополнительная	-	ЭБС ZNANIUM

Таблица 9.1.2 – Перечень современных профессиональных баз данных (СПБД)

№	Наименование СПБД
1	Электронная библиотека Grebennikon.ru – www.grebennikon.ru
2	Научная электронная библиотека eLIBRARRY – www.elibrary.ru
3	Научная электронная библиотека КиберЛеника – www.cyberleninka.ru
4	База данных ПОЛПРЕД Справочники – www.polpred.com
5	Международная реферативная база данных научных изданий Scopus – https://www.scopus.com
6	Международная реферативная база данных научных изданий Web of Science – http://webofscience.com
7	База данных OECD Books, Papers & Statistics на платформе OECD iLibrary – www.oecd-ilibrary.org

Таблица 9.1.3 – Перечень информационных справочных систем (ИСС)

№	Наименование ИСС
1	Справочная правовая система КонсультантПлюс (инсталлированный ресурс СПбГЭУ или www.consultant.ru)
2	Справочная правовая система «ГАРАНТ» (инсталлированный ресурс СПбГЭУ или www.garant.ru)
3	Информационно-справочная система «Кодекс» (инсталлированный ресурс СПбГЭУ или www.kodeks.ru)
4	Электронная библиотечная система BOOK.ru - www.book.ru
5	Электронная библиотечная система ЭБС ЮРАЙТ – www.urait.ru
6	Электронная библиотечная система ЗНАНИУМ (ZNANIUM) - www.znanium.com
7	Электронная библиотека СПбГЭУ– opac.unecon.ru

9.2. Материально-техническое обеспечение учебного процесса

Для реализации данной дисциплины имеются специальные помещения для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Таблица 9.2.1 – Перечень программного обеспечения (ПО)

№ п/п	Наименование ПО
1	Microsoft Windows Professional (КОНТРАКТ № 244/20 «26» июня 2020 г
2	Microsoft Office Professional (КОНТРАКТ № 244/20 «26» июня 2020 г)
3	7-Zip (freeware)

10. ОСОБЕННОСТИ ОСВОЕНИЯ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

В целях освоения учебной программы дисциплины инвалидами и лицами с ограниченными возможностями здоровья Университет обеспечивает:

– для инвалидов и лиц с ограниченными возможностями здоровья по зрению: размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме справочной информации о расписании учебных занятий; присутствие ассистента, оказывающего обучающемуся необходимую помощь; выпуск альтернативных форматов методических материалов (крупный шрифт или аудиофайлы);

– для инвалидов и лиц с ограниченными возможностями здоровья по слуху: надлежащими звуковыми средствами воспроизведение информации;

– для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата: возможность беспрепятственного доступа обучающихся в учебные помещения, туалетные комнаты и другие помещения кафедры, а также пребывание в указанных помещениях.

Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и в отдельных группах или в отдельных организациях.

11. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Фонд оценочных средств для проведения аттестации уровня сформированности компетенций обучающихся по дисциплине оформляется отдельным документом и является приложением к рабочей программе дисциплины.