

Министерство образования и науки РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«Новосибирский национальный исследовательский государственный
университет» (Новосибирский государственный университет, НГУ)
Гуманитарный факультет

Программа рассмотрена
на заседании кафедры
фундаментальной и прикладной
лингвистики
29.08.2014

Зав. кафедрой, проф. М.К. Тимофеева

Утверждаю

декан гуманитарного
факультета, профессор
1.09.2014
Л.Г. Панин

Основная образовательная программа
высшего образования

Направление подготовки
035800 – Фундаментальная и прикладная лингвистика

Квалификация (степень) выпускника –
бакалавр

ПРОГРАММА УЧЕБНОГО КУРСА
«СТИЛОМЕТРИЯ»

(64 часа, 2 з.е.)

1. Наименование дисциплины

ПРОГРАММА УЧЕБНОГО КУРСА «СТИЛОМЕТРИЯ»

Программа дисциплины «Стилометрия» составлена в соответствии с требованиями к обязательному минимуму содержания и уровню подготовки дипломированного бакалавра по направлению 035800 «Фундаментальная и прикладная лингвистика» в целях обеспечения реализации учебного процесса в НГУ.

Автор: Алешина Ольга Николаевна, доктор филол. наук, профессор

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.

Цели освоения специального курса

Специальный курс «Стилометрия» имеет своими целями:

- 1) познакомить обучающихся с основными понятиями стилометрических исследований;
- 2) научить их использовать современные стилометрические методы в решении собственных исследовательских задач.

В результате освоения дисциплины обучающийся должен:

- Знать базовую терминологию современных стилометрических исследований.
- Уметь выбрать и правильно применять стилометрические методы в решении собственных исследовательских задач.
- Владеть навыками стилометрического анализа текстового материала из семантически и грамматически аннотированных лингвистических корпусов.

Перечисленные результаты образования являются основой для формирования следующих общекультурных и общепрофессиональных компетенций:

а) общекультурными (ОК)

- стремлением к саморазвитию, повышению своей квалификации и мастерства ОК-6
- использованием основных положений и методов социальных, гуманитарных и экономических наук при решении социальных и ОК-9

профессиональных задач, способностью анализировать социально-значимые проблемы и процессы

б) профессиональными (ПК):

общепрофессиональными:

- знанием основных понятий и категорий современной лингвистики ПК-1
- в области научно-исследовательской деятельности:**
- владением основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов ПК-7

3. Место специального курса в структуре образовательной программы

Изучение спецкурса «Стилометрия» возможно после освоения курса «Современный русский язык» и курсов «Введение в языкознание» и «Корпусная лингвистика».

4. Объем дисциплины в зачетных единицах с указанием количества академических, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся.

Общая трудоемкость дисциплины составляет 2 зачетных единицы, 64 часа. Из них на контактную работу с преподавателем 32 часа (лекции), на самостоятельную работу студентов – 32 часа.

5. Содержание дисциплины, структурированное по темам с указанием отведенного на них количества астрономических часов и видов учебных занятий

№ п/п	Раздел Дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				Лекции	Практикум	Самостоят. Работа	Всего	
1	Тема 1. Библиография и стилометрия: общая характеристика.	5	1-2	4			4	
2	Тема 2. «Стилометрия» и «стилеметрия» в отечественных и	5	3	2			2	

	зарубежных исследованиях.							
3	Тема 3. Стилometрия и функциональная стилистика.	5	4-5	4	2		6	
4	Тема 4. Стилometрия языковых ресурсов: фонетический и графемный уровень.	5	6-7	4			4	
5	Тема 5. Стилometрия языковых ресурсов: лексический уровень.	5	8-9	4			4	
6	Тема 6. Стилometрия языковых ресурсов: грамматический уровень.	5	10-11	4			4	
7	Тема 7. Стилometрия и лингвистическая текстология.	5	12	2			2	
8	Тема 8. Стилometрия и атрибуция авторских текстов.	5	13-14	3			3	
9	Тема 9. Роль стилometрии в гендерологии и переводоведении.	5	15-16	2			2	
10	Тема 10. Перспективы в развитии стилometрических исследований.	5	16	1			1	
	<i>Всего</i>			30			32	
								Итогов ый зачет

Тема 1. Библиометрия и стилometрия: общая характеристика. 4 часа

Понятия библиометрии и стилometрии. Общая характеристика количественных методов в лингвистике. Стилometрия и лингвистическая статистика (SAS). Краткая история стилometрии (от трудов Л. Валла и «Принципов стилometрии» В. Любославского до современных работ по теории лингвостатистики и стилистической атрибуции текстов). Объект и предмет современных стилometрических исследований. Понятия «текстовый инвариант», «количественная закономерность текста», «лингвостатистический маркер», «наиболее употребительная единица». Частота и частотность единиц в языке и тексте. Высокочастотные единицы языка и текста. Ранговая статистика встречаемости слов в большой текстовой коллекции. Вероятностно-статистический анализ текста. Опыты выявления и исследования устойчивых и переменных статистических общеязыковых и

частнотекстовых характеристик. Дискуссия вокруг закона Ципфа. «Цепи Маркова» и их применение в стилометрии. Концептуальный анализ текста и текстовые лингвостатистические данные.

Стилометрия и вычислительная лингвистика. Разработка структурно-статистических методов и алгоритмов исследования текстов. Краткая характеристика частных проблем автоматического анализа текстов. Многопараметрический статистический анализ результатов ассоциативных экспериментов.

Тема 2. «Стилометрия» и «стилеметрия» в отечественных и зарубежных исследованиях.

2 часа

Стилометрия и лингвистическая экспертиза текста: узкий и широкий подходы. Центры стилометрических исследований в СССР и России (Киев, Тарту, Рига, Петербург, Москва, Новосибирск). Основные понятия «стилеметрии» в концепции Г. Я. Мартыненко. Стилометрия и ее связь с направлениями «внутренней» и «внешней» лингвистики.

Тема 3. Стилометрия и функциональная стилистика. 4 часа

Возможности применения методов количественной лингвистики в функционально-стилистических исследованиях. Вопрос о количестве стилей в современном русском литературном языке (с учетом стилометрических данных). Общая характеристика результатов информационно-статистического анализа текстов различной функциональной принадлежности (научных, официально-деловых, публицистических, фольклорных поэтических и фольклорных прозаических, художественных поэтических и художественных прозаических).

Тема 4. Стилометрия языковых ресурсов: фонетический и графемный уровень. 4 часа

Фонетическая стилометрия. Диагностика собственно лингвистических и экстралингвистических факторов звучащей речи. Статистико-дистрибутивный анализ дискурса и текста (на уровне фонем и графем). Опыт типологии упорядоченности фонем в «звуковых цепочках» разных языков (по Ю. А. Тамбовцеву). Эвфония, какофония и другие понятия фоники и просодики. Просодические средства организации стилистически различных текстов. Применение комбинаторно-статистического анализа в изучении поэтических текстов. «Модели» лингвистической поэтики.

Тема 5. Стилометрия языковых ресурсов: лексический уровень. 4 часа

Понятия лексической стилометрии и количественной лексикологии. Лексико-номинативные свойства языка и текста в статике и динамике: общая характеристика. Лексические маркеры. Лексическая структура текста и авторские ассоциации. Статистико-семантическая атрибуция текста. Исследования устойчивости лексико-статистических характеристик дискурса и текста. Вопрос об «окаzionale лексическом» в тексте. Выражение в тексте коммуникативных тактик и стратегий лексическими средствами (на примерах «языка любви» и «языка вражды»).

Тема 6. Стилометрия языковых ресурсов: грамматический уровень. 4 часа

Исследование грамматики числовыми методами (понятие «грамматики зависимостей»). Словообразовательная стилометрия. Вопрос об индивидуальности морфемного облика окаzionale единиц. Морфологическая интерпретация текста и ее влияние на морфологическую стилометрию. Корреляционные характеристики частей речи в связных текстах. Морфологические маркеры и вероятностная атрибуция текста по морфологическим признакам. Роль единиц различных частей речи в передаче текстовой информации. Лингвоперсонологическое функционирование частей речи (на материале художественных текстов). Синтаксическая стилометрия. Функционально-прагматическая концепция синтаксической организации текста. Графическое представление синтаксических структур и стилистическая диагностика текстов.

Тема 7. Стилометрия и лингвистическая текстология. 2 часа

Стилометрия и лингвотекстология. Вероятностная атрибуция в лингвотекстологии. Когнитивные и прикладные аспекты современной лингвотекстологии (на примерах решения проблем подлинности «Слова о полку Игореве» и атрибуции древнерусских риторических произведений). Опыты проведения отечественными лингвистами статистико-типологического анализа современных текстов.

Тема 8. Стилометрия и атрибуция авторских текстов. 3 часа

Идентификационные лингвистические (лингвоперсонологические) признаки текстов. Методы автороведческих экспертиз: общих («распознавания образов» (М. А. Марусенко), психолингвистической диагностики (В. П. Белянин)) и т.д. и «частноуровневых» (фоноскопических, лексических и т.д.). Использование стилометрических данных в криминалистике и судебном речеведении (К. И. Бринев). Стилометрия и авторский художественно-публицистический дискурс. «Грамматика» идиостиля. Темпоральная структура текста как компонент идиостиля. Лингвистические

методы атрибуции и датировки литературных произведений (на примерах неоднозначного решения проблем «Корнель - Мольер», атрибуции публицистических произведений, приписываемых А. С. Пушкину).

Тема 9. Роль стилометрии в гендерологии и переводоведении. 2 часа.

Стилометрия и гендерная лингвистика. «Женская» и «мужская» литература с учетом статистико-вероятностной модели гендернообусловленного дискурса. Стилометрия и вопросы теории перевода. Статистические критерии оценки качества перевода (по Е. Н. Бугулову и др.). Статистические параметры трансляции ритмико-фонетических особенностей звучащей речи при переводе.

Тема 10. Перспективы в развитии стилометрических исследований. 1 час

Расширение задач стилометрических исследований в связи с развитием корпусной лингвистики. Связь достоверности результатов текстологических трудов со стилометрическими данными об исследуемых текстах. Возможности разноаспектного исследования корпусов разговорных и письменных текстов с привлечением стилометрических данных.

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Учебные материалы для самостоятельной работы и презентации лекций рассылаются студентам в электронном виде

7. Фонд оценочных средств для промежуточной аттестации обучающихся по дисциплине.

В ходе изучения программных тем обучаемые слушают лекции. Зачет сдается по итогам всего специального курса.

Список вопросов к зачету:

1. Объект и предмет стилометрии.
2. Вероятностно-статистический анализ текста: общая характеристика.
3. Связь стилометрии с направлениями «внутренней» и «внешней» лингвистики.
4. Функционально-стилевая система современного русского языка с учетом стилометрических данных.
5. Фонетическая стилометрия: общая характеристика.

6. Возможности применения комбинаторно-стилистического анализа в изучении поэтических текстов.
7. Вопросы статистико-семантической атрибуции авторского текста.
8. Корреляция грамматических форм в тексте. Характеристика частотного словаря (по выбору обучаемого).
9. Решение проблемы подлинности «Слова о полку Игореве».
10. Использование стилометрических данных в анализе художественного текста (по выбору обучаемого).
11. Женский и мужской социолекты с позиций стилометрических данных.
12. Применение стилометрических данных в сопоставительных и типологических исследованиях фольклорных текстов.

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.

а) основная литература:

Баранов А.Н. Лингвистическая экспертиза текста: теоретические основания и практика: учеб. пособие. М.: Флинта; Наука, 2007.

Мартыненко Г.Я. Основы стилометрии. Л.: ЛГУ, 1988.

б) дополнительная литература:

российская (и советская):

Баранов А.Г. Функционально-прагматическая концепция текста. Ростов н/Д: Изд-во Рост. ун-та, 1993.

Белянин В. П. Основы психолингвистической диагностики: Модели мира в литературе. М.: Тривола, 2000.

Березкин А. М. Когнитивные и прикладные аспекты современной текстологии // Русская литература. — 2005. — № 1. — С. 67-80.

Болотнова Н.С. Лексическая структура художественного текста в ассоциативном аспекте. Томск: Б.и., 1994.

Зализняк А. А. «Слово о полку Игореве»: взгляд лингвиста. М.: Языки слав. культуры, 2004.

Лихачев Д.С. Текстология: На материале русской литературы X-XVII веков. — 3-е изд., перераб. и доп. — СПб.: Алетейя, 2001.

Айымбетов М. Опыт лингвостатистического анализа лексики и морфологии каракалпакского публицистического текста. Нукус: Билим, 1991.

Актуальные проблемы квантитативной лингвистики и автоматического анализа текстов. Тарту: ТГУ, 1981.

Баранов Ю. Н. Теоретические основы применения лингвистических знаний в криминалистике при производстве фоноскопических и автороведческих экспертиз: Дисс. ... канд. филол. наук. Челябинск, 2004.

Бринев К.И. Теоретическая лингвистика и судебная лингвистическая экспертиза. Барнаул: Алтайская гос. пед. акад., 2009.

Бугулов Е. Н. Статистические критерии оценки качества перевода. Киев: ИТФ, 1990.

Головина Т.А. Лингвоперсонологическое функционирование частей речи: статистический аспект : на материале художественных текстов : Дисс. ... канд. филол. наук. Кемерово, 2008.

Демьянков В.З. Морфологическая интерпретация текста и ее моделирование. М.: Изд-во МГУ, 1994.

Денисова А.В. Дискурсивный аспект исследования лингвистической экспертизы: на материале «языка вражды»: Автореф. дисс... канд. филол. наук. Краснодар, 2009.

Долотин К.И. Диагностика экстралингвистических факторов звучащей речи: экспериментально-статистическое исследование: Автореф. дис. ... доктора филол. наук. М., 2007.

Ермоленко Г.В. Статистико-семантическая атрибуция русских текстов второй половины XIX - начала XX в.: Автореф. дисс. ... канд. филол. наук. Ленинград, 1989.

Квантитативная лингвистика и автоматический анализ текстов. Тарту: ТГУ, 1985 (и др. выпуски).

Квантитативная лингвистика и семантика (Квалисем-99): II Межвуз. конф. : Тез. докл. / [Редкол.: С. М. Полянский (отв. ред.) и др.]. Новосибирск: Новосиб. гос. пед. ун-т, 1999.

Лагута О.Н., Лаврентьев А.М. Лингвистические приложения // Русские простонародные легенды и рассказы: сб. 1861 г. - Новосибирск: Наука, 2005. С. 185-317.

Левицкий В. В. Квантитативные методы в лингвистике. Винница: Нова книга, 2007.

Лингвостатистика и вычислительная лингвистика: Труды по лингвостатистике. Тарту: ТГУ, 1982.

Мартыненко Г. Я. Многопараметрический статистический анализ результатов ассоциативного эксперимента. СПб.: Изд-во С.-Петербур. ун-та, 2003.

Марусенко М.А. Атрибуция анонимных и псевдонимных литературных произведений методами теории распознавания образов: Автореф. дисс. ... канд. филол. наук. Л., 1991.

Назарова Т.В. Блок коммуникативных характеристик речи говорящих на русском языке в системе идентификационных лингвистических признаков. М.: ЭКЦ МВД России, 2009.

Носенко И.А. Начала статистики для лингвистов: учеб. пособие. М.: Высш. школа, 1981.

Панин Л.Г. Лингвотекстологическое исследование минейного Торжественника: Рукописи XIV-XVI вв. / Отв. ред. Л.П. Жуковская, А.И. Федоров. Новосибирск: Наука, 1988.

Пищальникова В.А. Концептуальный анализ художественного текста: Учеб. пособие. Барнаул: Изд-во Алт. гос. ун-та, 1991.

Полянская И.В. Проблемы методики атрибуции риторических произведений Древней Руси (опыт количественного анализа): Автореф. дисс. ... канд. филол. наук. М., 1992.

Рахманов Дж. А. Статистико-дистрибутивный анализ азербайджанского текста: Автореф. дисс. ... канд. филол. наук. Баку, 1988.

Родионова Е.С. Лингвистические методы атрибуции и датировки литературных произведений: к проблеме «Корнель – Мольер»: Автореф. канд. филол. наук. СПб., 2008.

Севбо И.П. Графическое представление синтаксических структур и стилистическая диагностика. Киев: Наук. думка, 1981.

Середа Ю.С. Основы информационно-статистического анализа текстов. Нижний Новгород: Поволжье, 2009.

Синелева А.В. Атрибуция «Романа с кокаином»: лингвостатистическое исследование: Автореф. дисс. ... канд. филол. наук. СПб., 2001.

Статистика казахского текста: Труды группы автоматизации лингвистических исследований и статистики речи. Алма-Ата: Гылым, 1990.

Статистика речи и автоматическая переработка текстов / [Составители Е. А. Шингарева и др.]. Л.: ЛГПИ, 1988.

Суркова А. С. Разработка структурно-статистических методов и алгоритмов идентификации текста: Автореф. дисс. ... канд. филол. наук. Нижний Новгород, 2004.

Тамбовцев Ю.А. Некоторые теоретические положения типологии упорядоченности фонем в звуковой цепочке языка и компендиум статистических характеристик основных групп согласных фонем. Новосибирск: Новосиб. клас. ин-т, 2001.

Федосова Т. В. Темпоральная структура текста как компонент идиостиля автора (на материале произведений К. Воннегута и Дж. Фаулза). Горно-Алтайск: Изд-во Горно-Алт.ун-та, 2006.

Хозяинов С.А. Атрибуция публицистических произведений, приписываемых А.С. Пушкину: тексты 1830-1836 гг.: Автореф. дисс. ... канд. филол. наук. СПб., 2008.

Чардин И.С. Лингвистические корпуса с разметкой на основе грамматики зависимостей и их применение при автоматическом синтаксическом анализе: Дисс... канд. филол. наук. М., 2004.

Якубайтис Т. А. Вероятностная атрибуция типа текста по нескольким морфологическим признакам. Рига: ИЭВТ, 1982.

Якубайтис Т. А. Корреляционные характеристики частей речи в связных текстах. Рига: Зинатне, 1980.

зарубежная:

Baayen R. H. Word frequency distributions. Dordrecht [etc.] : Kluwer acad. publ., Cop. 2001.

Franking H. Stylometry: a statistical method for determining authorship, textual integrity, and chronology. University of Kansas, 1988.

Kenny A. The Computation of Style: An Introduction to Statistics for Students of Literature and Humanities. Oxford, Pergamon Press, 1982.

Lebart L., Salem A., Berry L. Exploring textual data. Springer, 1998.

Juola P. Authorship Attribution // <http://www.mathcs.duq.edu/~juola/papers.d/fnt-aa.pdf>.

Mežs I. The Latvian language in the mirror of statistics. Riga: Jāņa sēta Map publ., Cop. 2005.

9. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины

Ляшевская О.Н., Шаров С. А. Новый частотный словарь русской лексики // <http://dict.ruslang.ru/freq.php>

Signature Stylometric System - <http://www.philocomp.net/?pageref=humanities&page=signature>

JGAAP Authorship Attribution Program - <http://www.jgaap.com/>

Лингвистические корпуса:

[American Rhetoric/ Online Speech Bank](http://www.americanrhetoric.com/) <http://www.americanrhetoric.com/>

ARCHER – A Representative Corpus of Historical English Registers: <http://www.llc.manchester.ac.uk/research/projects/archer/>

[Bank of English](http://www.titania.bham.ac.uk/docs/svenguide.html) (COBUILD - Collins Birmingham University International Language Database): <http://www.titania.bham.ac.uk/docs/svenguide.html>

BASE – British Academic Spoken English: <http://www2.warwick.ac.uk/fac/soc/al/research/collect/base/>

BNC – [British National Corpus](http://www.natcorp.ox.ac.uk/) <http://www.natcorp.ox.ac.uk/>

Brown Corpus <http://khnt.aksis.uib.no/icame/manuals/brown/>

CHILDES - Child Language Data Exchange System: <http://childes.psy.cmu.edu/>
COCA – [Corpus of Contemporary American English](http://www.americancorpus.org/) <http://www.americancorpus.org/>
COLT - Bergen Corpus of London Teenage Language: <http://torvald.aksis.uib.no/colt/>
[Complete Corpus of Old English](http://www.doe.utoronto.ca/) <http://www.doe.utoronto.ca/>
[Emotional Speech Corpus](#) (о проекте:
<http://arrow.dit.ie/cgi/viewcontent.cgi?article=1004&context=dmcart>)
IViE - Intonational Variation in English: <http://www.phon.ox.ac.uk/files/apps/IViE/>
L-CIE - [Limerick Corpus of Irish-English](http://www.ul.ie/~lcie/): <http://www.ul.ie/~lcie/>
[Leap - Learning the Prosody of a foreign language](http://www.phonetik.uni-freiburg.de/leap/): <http://www.phonetik.uni-freiburg.de/leap/>
Roget's Thesaurus of English Words and Phrases
<http://www.gutenberg.org/files/10681/10681-body.txt>
TalkBank: <http://talkbank.org/>
The Oxford Text Archive <http://ota.ahds.ac.uk/>
YCOE - The York-Toronto-Helsinki Parsed Corpus of Old English Prose: <http://www-users.york.ac.uk/~lang22/YcoeHome1.htm>
Весь Достоевский <http://petsru.ru/~Dostoevsky/main.htm>
ГАФИС «Брюсов» www.vbryusov.ru
КОРГАЗ – Компьютерный корпус текстов русских газет конца XX века
<http://www.philol.msu.ru/~lex/corpus/>
Корпус русского языка XVIII века (о проекте:
<http://textualheritage.org/content/view/206/93/>)
Лидские корпуса русских текстов corpus.leeds.ac.uk
НКРЯ – Национальный корпус русского языка ruscorpora.ru
СКАТ <http://project.phil.pu.ru/scat/page.php?page=project>
Словарь-конкорданс публицистики Ф.М. Достоевского
<http://dostoevskii.karelia.ru/index.phtml>
Словарь-корпус языка А. С. Грибоедова <http://feb-web.ru/feb/concord/abc/11/gr11-522.htm> и <http://www.inforeg.ru/electron/concord/concord.htm>
СПИ - Историко-параллельный корпус «Слова о полку Игореве»
<http://nevmenandr.net/slovo/>
Упсальский корпус современных русских текстов
<http://www.slaviska.uu.se/korpus.htm>
ХАНКО – Хельсинкский аннотированный корпус русских текстов
<http://www.helsinki.fi/venaja/russian/e-material/hanco/index.htm>

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.

На лекциях и зачете необходимы компьютеры с доступом в Интернет.

Для аудиторных лекционных занятий необходимы компьютерное обеспечение и доступ к Интернету.