

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВО «СГУ имени Н.Г. Чернышевского»

Институт филологии и журналистики



Программа производственной практики
Прикладная практика 1

Направление подготовки бакалавриата
45.03.01 – Филология

Профиль подготовки
Отечественная филология

Квалификация (степень) выпускника
Бакалавр

Форма обучения
заочная

Саратов
2016

1. Цели производственной практики

Целями производственной практики являются:

- глубокое теоретическое и практическое освоение студентами системы знаний в области компьютерной обработки текста и звучащей речи;
- подготовка обучающихся к самостоятельному использованию технологий обработки текста и звучащей речи;
- развитие навыков лингвистического анализа текстов различных стилей;
- освоение научных, теоретических и методических основ корпусного подхода в лингвистике;
- расширение профессиональных и общекультурных навыков и компетенций.

2. Тип учебной/производственной практики и способ ее проведения

Тип учебной практики: практика по получению первичных профессиональных умений и навыков.

Способ проведения учебной практики: стационарная.

3. Место производственной практики в структуре ООП бакалавриата

Прикладная практика 1 относится к вариативной части блока «Б2. Практики» по направлению подготовки 45.03.01 – «Филология» и является обязательной. Она опирается на знания и навыки, получаемые студентами в ходе освоения курсов «Введение в языкознание», «Современный русский язык», «Информационные технологии в филологии», а также на практическое и теоретическое владение родным и иностранными языками.

Прикладная практика подготавливает обучающихся к использованию продуктов и технологических средств обработки текстов и звучащей речи в их будущей профессиональной деятельности:

- научно-исследовательской;
- производственно-практической;
- проектной;
- экспертиро-аналитической.

Теоретические и прикладные знания, умения, навыки используются обучающимися при освоении прикладных дисциплин, в будущей профессиональной деятельности.

4. Компетенции обучающегося, формируемые в результате прохождения производственной практики

В результате прохождения данной производственной практики обучающийся должен приобрести следующие практические навыки, умения, универсальные и профессиональные компетенции:

- способность работать в коллективе, толерантно воспринимая социальные, этнические, конфессиональные и культурные различия (ОК-6),
- владение базовыми навыками сбора и анализа языковых и литературных фактов, филологического анализа и интерпретации текста (ОПК-4),
- способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности (ОПК-6),
- способность применять полученные знания в области теории и истории основного изучаемого языка (языков) и литературы (литератур), теории коммуникации, филологического анализа и интерпретации текста в собственной научно-исследовательской деятельности (ПК-1),
- владение навыками подготовки научных обзоров, аннотаций, составления рефератов и библиографий по тематике проводимых исследований, приемами библиографического описания; знание основных библиографических источников и поисковых систем (ПК-3),
- владение навыками участия в научных дискуссиях, выступления с сообщениями и докладами, устного, письменного и виртуального (размещение в информационных сетях) представления материалов собственных исследований (ПК-4),
- владение базовыми навыками создания на основе стандартных методик и действующих нормативов различных типов текстов (ПК-8),
- владение базовыми навыками доработки и обработки (например, корректура, редактирование, комментирование, реферирование, информационно-словарное описание) различных типов текстов (ПК-9),
- владение навыками перевода различных типов текстов (в основном научных и публицистических, а также документов) с иностранных языков и на иностранные языки; аннотирование и реферирование документов, научных трудов и художественных произведений на иностранных языках (ПК-10).

- Знать: принципы работы программ, направленных на автоматическую обработку естественного языка, особенности подготовки текстов и автоматизированного рабочего места лингвиста к машинному анализу, основные цифровые форматы данных, в которых выступает текст как объект автоматической обработки; акустические особенности звучащей речи и способы их отображения и анализа с помощью специализированных программных средств; функциональные возможности текстовых корпусов и ограничения, накладываемые на корпусное лингвистическое исследование;
- Уметь: пользоваться специализированным программным обеспечением (фонетическими, морфологическими, синтаксическими анализаторами, конкордансерами, программами синтеза и обработки звучащей речи);

- Владеть: навыками машинного анализа звучащей речи: записи и обработки звуковых файлов, анализа спектрограмм, интонограмм, формантного состава звуковых фрагментов, транскрипционного аннотирования фрагментов звучащей речи (с метками фонем и слов); импортирования звуковых, графических и текстовых файлов из внутренних форматов программ в современные наиболее распространенные форматы данных.

5. Структура и содержание производственной практики

Общая трудоемкость производственной практики составляет **3** зачетных единицы, **108** часов.

№ п/ п	Разделы (этапы) практики	Виды учебной работы на практике, включая самостоятельную работу студентов и трудоемкость (в часах)				Формы текущего контроля
		Всего	Лекц.	Практич	СРС	
1.	Обзор технологий обработки звучащей речи и сфер их применения. Знакомство с интерфейсом программы фонетического анализа	6		6		Коллоквиум
2.	Осциллографический анализ звучащей речи	16		16		Лабораторная работа
3.	Спектральный анализ речевых отрезков	16		16		Лабораторная работа
4.	Акустический анализ высказываний	16		16		Отчетная работа
5.	Составление частотного словаря по своей исследовательской теме	16		16		Лабораторная работа, промежуточная отчетная работа
6.	Работа с НКРЯ по своей исследовательской теме	16		16		Лабораторная работа, промежуточная отчетная работа
7.	Разработка структуры базы данных	6		6		Коллоквиум
8.	Создание базы данных	16		16		Отчетная работа
Всего:		108		108		Зачет с оценкой

Раздел 1. Обзор технологий обработки звучащей речи и сфер их применения. Программные средства обработки звучащей речи. Программы анализа и синтеза речи. Онлайн-инструменты. Знакомство с фонетическим

анализатором Praat: создание (запись) звукового файла, сегментирование аудиофайла, аннотирование фрагментов записи (создание транскрипции), применение акустических фильтров, построение графиков звуковой волны (осцилограмма, спектrogramma, интонограмма, график интенсивности, маркеры периодичности колебаний). Создание и сохранение файлов графических изображений в формате EPS с последующей конвертацией в форматы .jpeg, .png.

Раздел 2. Осциллографический анализ звучащей речи. Анализ осцилограмм отдельных звуков (гласных, сонорных, взрывных, фрикативных согласных, аффрикат), их сопоставление; анализ осцилограмм слогов и слов. Анализ акустических характеристик отдельных звуков и тех же звуков в комбинациях.

Раздел 3. Анализ спектральных характеристик звучащей речи. Анализ спектrogramm гласных и согласных звуков (звонких и глухих, губных, заднеязычных, свистящих, шипящих, аффрикат). Формантный состав гласных звуков: связь значений формантных частот с артикуляционными признаками гласных; сопоставление характеристик отдельно произнесенных гласных и гласных в окружении. Анализ акустической структуры гласного в динамике. Стационарные и переходные участки звукового спектра.

Раздел 4. Акустический анализ высказываний. Осциллографический и спектральный анализ речевых отрезков, анализ интонационного рисунка, изменений силы звука на протяжении фрагмента (контур интенсивности). Аннотирование высказываний (добавление фонетической транскрипции в файл с графиками звукового сигнала) по двум уровням: звуки, слова.

Раздел 5. Составление частотного словаря по своей исследовательской теме. Подготовка пользовательского текстового корпуса (определение принципов создания, отбор текстового материала). Проведение лемматизации с помощью программы Mystem; задание параметров работы программы, оптимальных для поставленных задач, ручная корректировка результатов. Составление частотного словаря лемм с помощью программы Simple Concordance Program, анализ полученного результата (описание структурных зон полученного частотного словаря).

Раздел 6. Работа с НКРЯ по своей исследовательской теме. Определение типа пользовательских подкорпусов и запросов в соответствии с индивидуальными исследовательскими задачами и с учетом возможностей Национального корпуса русского языка (морфологический и семантический поиск, задание шаблонов и пр.). Получение конкорданса из НКРЯ по разработанным запросам. Анализ полученных результатов.

Раздел 7. Разработка структуры базы данных. Разработка структуры базы данных для предметной области «Средства создания экспрессии в рекламных слоганах»: описание предметной области в терминах инфологической модели «сущность-связь», выделение сущностей и их атрибутов, определение типа связей между сущностями.

Раздел 8. Создание базы данных. Корректировка разработанной общей структуры БД с учетом специфики материала обучающегося.

Реализация структуры в СУБД Access: создание таблиц, определение типа данных, создание связей между таблицами, обеспечение целостности, создание запросов (2 на выборку с комбинацией условий и данных из связанных таблиц, 1 на обновление).

Формы проведения учебной/производственной практики лабораторная

Место и время проведения учебной/производственной практики

Местом проведения производственной практики является компьютерный класс Института филологии и журналистики СГУ им. Н.Г. Чернышевского, практика проводится в 7 семестре (2 недели).

Формы промежуточной аттестации (по итогам практики)

Промежуточная аттестация по итогам практики осуществляется в форме собеседования по результатам подготовленных отчетных работ. Обучающемуся выставляется зачет с оценкой. Время проведения аттестации – декабрь 7 семестра.

6. Образовательные технологии, используемые на производственной практике

Занятия по прикладной практике ориентированы на активное приобретение обучающимися знаний в области автоматической обработки естественных языков и навыков использования специализированного программного обеспечения, что обуславливает использование в образовательном процессе как компьютерной техники и мультимедийного оборудования (проекторы, звуковоспроизводящие устройства), так и различных информационных технологий, включающих а) ресурсы сети Интернет (текстовые корпуса со встроенным механизмом анализа текстов, специализированные веб-сайты, интернет-курсы, разработанные в системе Moodle (<http://course.sgu.ru>); б) аудиоредакторы (Audacity, Adobe Audition, Praat); в) программы синтеза речи на русском и английском языках.

Сообщение нового материала по проблемам технологий обработки звучащей речи ориентировано на максимальное привлечение знаний, полученных обучающимися в предшествующих курсах, на активное восприятие обучающимися новых сведений и обязательную обратную связь в ходе изложения материала (диалоги с преподавателем, ответы на вопросы преподавателя, решение предлагаемых им задач, сопоставление, оценка различных ответов). В ходе практических занятий используются элементы кейс-технологии. Практические занятия проводятся в компьютерных классах с использованием изучаемых технологий и других электронных средств (презентации в PowerPoint, электронные БД, энциклопедии, корпуса и другие электронные ресурсы) для наиболее разнообразного представления материала и стимуляции активности обучающихся. Для самостоятельного изучения технологий обработки текстов и звучащей речи предлагаются

использование интернет-ресурсов, свободно распространяемых программ, подготовка к участию в коллоквиумах по предлагаемым темам курса.

Для внеаудиторной работы предлагаются задания, ориентированные на использование Интернет-ресурсов и специализированных программ, на самостоятельный сбор и обработку материала с использованием современных информационных технологий.

Для студентов с ОВЗ и инвалидов применяются технологии дистанционного обучения (контроль самостоятельной и письменных работ по электронной почте), задания с использованием Интернет-ресурсов (см. разд. 9). По всем разделам курса студентам с ОВЗ и инвалидам предоставляются электронные презентации.

7. Учебно-методическое обеспечение самостоятельной работы студентов на производственной практике

Для активации учебной деятельности студентов и успешного усвоения курса используются следующие виды самостоятельной работы:

- чтение учебной и научной литературы, рекомендованной преподавателем, с последующим проведением индивидуальных отчетов и коллоквиумов;
- выполнение самостоятельных работ.

Примеры заданий

По разделам 1-4

Задание к отчетной работе:

I. Анализ акустических характеристик отдельных звуков

1. Запишите отдельные звуки, произнесенные тремя разными голосами (двумя женскими и одним мужским (можно детским)).
2. Проанализируйте осцилограммы звуковых колебаний: сравните графики гласных, согласных и сонорных звуков.
3. Обратите внимание на огибающую «гребенки», на участки периодических и непериодических колебаний – для каких звуков они характерны?
4. Сравните осцилограммы одного звука, произнесенного разными голосами? Какие визуальные характеристики графиков варьируются? Проведите аналогичное сравнение на материале спектрограмм.
5. Проанализируйте спектрограммы записанных звуков, добавьте в график информацию о формантном составе звучащего отрезка. Найдите стационарные и переходные участки звучания.
6. Для гласных звуков: выделите на спектrogramме стационарный участок звучания и вычислите для него: F_0 , F_1 и F_2 . Сравните эти показатели а) для разных гласных и б) для одних и тех же гласных, произнесенных разными голосами.

II. Анализ акустических характеристик звуков в окружении

1. Запишите в режиме «Моно» предложенные в Вашем варианте слоги и слова.
2. Проанализируйте осцилограмму и спектрограмму слогов; можно ли по графикам определить границы звуков, составляющих слог? Сделайте разметку звукового отрезка: добавьте слой аннотации «Звуки». Сравните характеристики отдельных звуков и тех же звуков в составе слога.
3. Проанализируйте осцилограмму и спектрограмму целых слов: а) сравните длительность ударных и безударных слогов; б) сопоставьте ранее рассмотренные

- характеристики отдельных звуков и звуков, соответствующих фонемам, которые затронуты позиционными изменениями, или (для глухих согласных) занимающих конечную позицию; в) постройте для слов графики интенсивности и интонограммы, прокомментируйте полученные изображения.
4. Сделайте анализ осцилограмм и спектрограмм русскоязычных слов и их англоязычных «соответствий», прокомментируйте графики.

III. Анализ интонационных характеристик звукового отрезка

1. Запишите предложение, предложенное в Вашем варианте, в виде а) повествовательного (нейтрального и с акцентами на разных словах), б) вопросительного, в) восклицательного высказывания.
2. Предложите анализ осцилограммы (обратите внимание на длительность отдельных участков (слов, слогов)), интонограммы, графика интенсивности колебаний звуковой волны.

По разделам 5-8.

Задания к отчетной работе:

1. По материалу своей исследовательской работе сформируйте текстовый корпус, объемом не менее 700 словоупотреблений. Используя программы Mystem и Simple Concordance Program, создайте частотный словарь лемм для сформированного корпуса. Прокомментируйте распределение лемм по зонам полученного словаря.
2. Составьте небольшой текст (ок. 1/2 страницы шрифтом Times New Roman кегль 14 пт.) о том, как вы можете или по каким причинам не можете использовать в своей исследовательской работе Национальный корпус русского языка. В НКРЯ составьте запрос (или несколько запросов), соотносящийся с темой вашей исследовательской работы (оценивается в том числе то, насколько полно вы используете поисковые возможности НКРЯ – возможность создания пользовательского подкорпуса, поиск по грамматическим и семантическим параметрам, задание шаблонов, возможности уточнение поиска, количественные данные и т.п.). Кратко (!) опишите процедуру составления запросов, предоставьте скриншоты. Прокомментируйте полученные результаты. Какие тенденции позволяет выявить использование НКРЯ?
3. По разработанной на занятии структуре создайте в Access базу данных для предметной области «Средства создания экспрессии в рекламных слоганах» по одному из типов продуктов (алкогольные напитки, безалкогольные напитки, еда, детское питание, бытовая химия, косметика, автомобили, электронные устройства, бытовая техника, услуги (в т.ч. телекоммуникационные), лекарства, одежда и обувь). Составьте запросы: какие экспрессивные средства языкового уровня X использовались в рекламе товаров страны Y после года Z? какие экспрессивные средства использовались для рекламы товаров типа X (запрос должен быть реализован таким образом, чтобы при активации всплывало диалоговое окно, в которое пользователь вводит тип продукта)? измените у компании X страну.

8. Данные для учета успеваемости студентов в БАРС

Таблица 1. Таблица максимальных баллов по видам учебной деятельности.

1	2	3	4	5	6	7	8	9
Семестр	Лекции	Лабораторные занятия	Практические занятия	Самостоятельная работа	Автоматизированное тестирование	Другие виды учебной деятельности	Промежуточная аттестация	Итого
1	0	30	20	0	0	20	30	100
Итого								100

Программа оценивания учебной деятельности студента

7 семестр
номер семестра

Лабораторные занятия

25-30 – все работы выполнены самостоятельно и без ошибок, грамотно оформлены;

17-24 – работы выполнены самостоятельно, без содержательных ошибок, есть недочеты в оформлении;

11-16 – работы выполнены самостоятельно, однако имеются недочеты или пропуски в ответах (не более чем на 2 задания);

6-10 – работа выполнена самостоятельно, но в ней присутствуют серьезные содержательные недочеты или же работа выполнена не полностью.

0-5 – выполнены не все работы, предусмотренные программой, в них имеются серьезные содержательные недочеты.

Практические занятия

16-20 – высокая посещаемость (не более 1 пропуска), активная работа на практических занятиях, самостоятельность при выполнении заданий и поиске дополнительной учебной информации, высокий уровень подготовки к занятиям.

10-15 – хорошая посещаемость (не более 4 пропусков), активная работа на практических занятиях, самостоятельность при выполнении заданий и поиске дополнительной учебной информации, хороший уровень подготовки к занятиям.

0-9 – посещение не менее 50% занятий, работа на практических занятиях, правильное выполнение основных заданий.

Другие виды учебной деятельности

16-20 – активная подготовка индивидуальных проектов, правильное выполнение всех контрольных работ.

10-15 – подготовка одного индивидуального проекта, правильное выполнение всех контрольных работ.

5-9 – написаны все контрольные работы, но в них имеются серьезные содержательные недочеты.

0-4 – индивидуальные проекты не подготовлены, контрольные работы выполнены не в полном объеме.

Промежуточная аттестация

При проведении промежуточной аттестации

ответ на «отлично» оценивается от 25 до 30 баллов;

ответ на «хорошо» оценивается от 18 до 24 баллов;

ответ на «удовлетворительно» оценивается от 10 до 17 баллов;

ответ на «неудовлетворительно» оценивается от 0 до 9 баллов.

Таким образом, максимально возможная сумма баллов за все виды учебной деятельности студента за 7 семестр по практике «Прикладная практика 1»
номер семестра наименование дисциплины

составляет 100 баллов.
количество баллов

Таблица 2. Таблица пересчета полученной студентом суммы баллов по практике «Прикладная практика 1» в оценку:

80-100 баллов	зачтено («отлично»)
60-79 баллов	зачтено («хорошо»)
40-59 баллов	зачтено («удовлетворительно»)
0-39 баллов	не зачтено («не удовлетворительно»)

9. Учебно-методическое и информационное обеспечение учебной/производственной практики.

а) основная литература:

Гольдин В.Е., Крючкова О.Ю. Электронные лингвистические ресурсы (электронное пособие). Саратов, 2011.

Щипицна Л.Ю. Информационные технологии в лингвистике : учебное пособие. М.: Флинта, 2015.

б) дополнительная литература:

Лингвистический энциклопедический словарь / Под ред. В.Н. Ярцевой. М., 1990.

Марчук Ю.Н. Компьютерная лингвистика: Учеб. пособие / Ю.Н. Марчук. – М.: АСТ: Восток-Запад, 2007 (1 экз. в ЗНБ и 2 экз. на кафедре).

Информатика: Учебник для вузов. Стандарт третьего поколения. СПб.: Питер, 2011.

в) программное обеспечение и Интернет-ресурсы:

Кофф -

- электронные публикации:

Беляева Л. Н. Лингвистические автоматы в современных гуманитарных технологиях: Учебное пособие. СПб.: ООО «Книжный Дом», 2007. 192 с. (Режим доступа: <http://window.edu.ru/library/pdf2txt/234/64234/34931/page1>, ссылка на скачивание: <http://window.edu.ru/resource/234/64234/files/belyaeva2.pdf>).

Волкова И.А. Адаптация и обучение системы общение с ЭВМ на естественном языке. Дис... канд. Физ.-мат. Наук. М., 1982. 78 с. (Морфологическая модель) (режим доступа: <http://axofiber.no-ip.org/download/volkova-dissertation.pdf>).

Пиотровский Р. Г. Инженерная лингвистика: Теория – Эксперимент – Реализация // Изв-я РАН. Сер. Лит. И Яз. Т. 37. №1. 1978. С. 10-19. (Режим доступа: <http://feb-web.ru/feb/izvest/1978/01/781-010.htm>)

Пиотровский Р.Г. Автоматическая переработка текста: теория и практика к концу ХХ в. // НТИ. Сер. 2. – 1998. – № 5. – С. 26–36. (Режим доступа: http://www.viniti.ru/cgi-bin/nti/nti.pl?action=show&year=2_1998&issue=5&page=26)

- Электронные курсы в системе дистанционного образования Moodle (<http://course.sgu.ru>). <http://www.speechpro.ru/technologies> – сайт Центра речевых технологий в Санкт-Петербурге.

- Аудиоредакторы:

<http://audacity.sourceforge.net/download/> – программа Audacity.

<http://www.fon.hum.uva.nl/praat/> – страница проекта Praat (Амстердамский университет).

<http://www.sil.org/computing/speechtools/softdev2/Cecil2/CECdownloads2.htm> – программа WinCecil компании SIL.

- Синтезаторы:

а) онлайн-синтезаторы:

<http://imtranslator.net/translate-and-speak/#window> – онлайн-синтезатор речи Imtranslator (с поддержкой русского языка).

<http://www.nuance.com/vocalizer5/flash/index.html> – онлайн-синтезатор речи Vocalizer 5 (с поддержкой русского языка).

<http://rssradio.ru/> – программа озвучивания новостных RSS-каналов, созданная Центром речевых технологий (онлайн-синтезатор).

б) Программы синтеза речи:

<http://dimio.site50.net/dspeech.html> – TTS-программа DSpeech, ссылка на скачивание (английский язык), <http://dimio.altervista.org/stats/download.php?id=51> – ее русификация с сайта производителя.
<http://espeak.sourceforge.net/> – TTS-программа ESpeak
<http://www.naturalreaders.com/download.htm> – Free Natural Reader 10.0.
(<http://mrtranslate.ru/download/texttospeech.html> – обзор синтезаторов речи).

10. Материально-техническое обеспечение учебной/производственной практики.

Необходимыми средствами обучения является вычислительная техника и программное обеспечение, а именно компьютерный класс, который соответствует следующим требованиям: вычислительная техника: 8-10 персональных компьютеров под управлением операционной системы Microsoft Windows XP, 7, 8.

В Институте филологии и журналистики имеются 2 компьютерных класса открытого доступа в Интернет общей вместимостью 18 рабочих мест. Компьютерные классы обеспечены микрофонами и наушниками, а также необходимым комплектом лицензионного программного обеспечения.

Кафедра располагает презентационным оборудованием.

Программа составлена в соответствии с требованиями ФГОС ВО по направлению и профилю подготовки 45.03.01 Филология (квалификация (степень) «бакалавр»).

Авторы: к.ф.н. доц. Степанова Н.Б., к.ф.н. доц. Трещева Е.Г.

Программа разработана в 2010 г. (одобрена на заседании кафедры теории, истории языка и прикладной лингвистики от 29.09.2010 года, протокол № 3).

Программа актуализирована в 2015 г. (одобрена на заседании кафедры теории, истории языка и прикладной лингвистики от 16.10.2015 года, протокол № 4).

Программа актуализирована в 2016 г. (одобрена на заседании кафедры теории, истории языка и прикладной лингвистики от 27.09.2016 года, протокол № 2).

Зав. кафедрой теории,
истории языка и прикладной лингвистики

О.И.О. Крючкова

Директор
Института филологии и журналистики СГУ

Л.С. Борисова