

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИИ



Федеральное государственное бюджетное образовательное учреждение высшего образования

«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГУМАНИТАРНЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «РГГУ»)

Институт лингвистики

УНЦ компьютерной лингвистики

Рабочая программа дисциплины

«Автоматическая оценка сложности текстов»

Направление подготовки 45.04.03 Фундаментальная и прикладная лингвистика

Магистерская программа: Фундаментальная и компьютерная лингвистика

Квалификация выпускника: магистр

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2021

АВТОМАТИЧЕСКАЯ ОЦЕНКА СЛОЖНОСТИ ТЕКСТА

Рабочая программа дисциплины

Рабочая программа дисциплины

Составитель:

к.ф.н., доцент А.Ч. Пиперски

Ответственный редактор:

д. филол. н., профессор В.И.Подлеская

УТВЕРЖДЕНО

Протокол заседания УНЦ компьютерной
лингвистики

№ 6 от «9» апреля 2021г.

ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины

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

1.3. Место дисциплины в структуре образовательной программы

2. Структура дисциплины

3. Содержание дисциплины

4. Образовательные технологии

5. Оценка планируемых результатов обучения

5.1. Система оценивания

5.2. Критерии выставления оценок

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

7. Материально-техническое обеспечение дисциплины

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений

1. ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

1.1. Цели и задачи дисциплины

Предмет курса – современные представления компьютерной лингвистики об оценке сложности текстов.

Цель курса – освоение студентами базовых понятий и методов оценки сложности текста в компьютерной лингвистике.

Задачи курса:

Курс нацелен на **формирование** у студентов следующих профессиональных **компетенций**:

- владением основными понятиями и категориями современной компьютерной лингвистики;
- владением основными методами автоматического анализа текста.

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

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-1 Владеет основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов	ПК-1.1	Знает основные понятия и категории современной лингвистики; основные методы научно-исследовательской деятельности в области фонологического, морфологического, синтаксического, дискурсивного и семантического анализа и правила их применения. Имеет представление об уровневой структуре естественного языка; основных параметрах разнообразия естественных языков; генетической, ареальной и типологической классификации языков; структуре лингвистической науки и ее основных направлениях; основных классических трудах по лингвистике
ПК-3 Способен использовать лингвистические технологии для проектирования систем автоматической обработки звучащей речи и письменного текста на естественном языке, лингвистических компонентов интеллектуальных и информационных электронных систем	ПК-3.3	Умеет пользоваться существующими системами автоматической обработки текста и звучащей речи, интеллектуальными и информационными электронными системами; проводить их сравнительный анализ; проектировать модули данных систем, составлять технические задания

1.2. Место дисциплины в структуре образовательной программы

Курс «Автоматическая оценка сложности текстов» является дисциплиной по выбору цикла дисциплин ООП ВПО (магистратуры) и относится к части, формируемой участниками образовательных отношений.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: Введение в компьютерную лингвистику.

2. СТРУКТУРА ДИСЦИПЛИНЫ

Общая трудоёмкость дисциплины составляет 2 з.е., 76ч., в том числе контактная работа обучающихся с преподавателем 20ч., промежуточная аттестация 0ч., самостоятельная работа обучающихся 56ч.

№№ раз-дела	Раздел курса	Семестр: недели	Виды учебной работы и трудоёмкость (в часах)			Формы контроля успеваемости
			лек-ции	практические занятия	СРС	
1.	Введение в проблематику сложности текста. Основные подходы к оценке сложности.	3: 1-2		7	16	Контроль посещаемости студентов.
2.	Автоматическая оценка сложности на основе выделения формальных признаков текста.	3: 3-4		7	20	Контроль посещаемости студентов. Проверка качества чтения и конспектирования литературы.
3.	Использование имеющихся ресурсов для оценки сложности текстов.	3: 5-6		6	20	Контроль посещаемости студентов. Проверка качества чтения и конспектирования литературы.
	Зачет					Контрольные вопросы
	Итого:			20	56	

3. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Раздел I. Введение в проблематику сложности текста. Основные подходы к оценке сложности (лекции 4 часа, практические занятия 3 часа, СРС 16 часов).

Сложность текста как практическая и теоретическая проблема. Сложность языка и сложность текста. Сложность в теории информации. Колмогоровская сложность. Анализ различных подходов к проблематике сложности текстов и языков. Анализ различных количественных подходов к сложности текста.

Раздел II. Автоматическая оценка сложности на основе выделения формальных признаков текста (лекции 3 часа, практические занятия 4 часа, СРС 18 часов)

Автоматическое извлечение лингвистических признаков из текста. Основные меры удобочитаемости для английского языка: Automated readability index (1967), ATOS Coleman–Liau index (1975), Dale–Chall readability formula (1948), Flesch–Kincaid readability tests Flesch reading ease (1975), Flesch–Kincaid grade level (1975), FORCAST (1973), Fry readability formula (1968), Gunning fog index (1952), Lexile (1989), Linsear Write Raygor readability estimate (1977),

SMOG (1969), Spache readability formula (1952). Применение вышеперечисленных мер к разноструктурным языкам.

Раздел III. Использование имеющихся ресурсов для оценки сложности текстов (лекции 3 часа, практические занятия 3 часа, СРС 18 часов)

Сопоставление автоматически полученных оценок с Общевропейскими компетенциями владения иностранным языком (CEFR). Анализ автоматически собранного материала из Википедии на английском и простом английском языках.

4. Образовательные технологии

Интерактивные формы обучения в данном курсе предполагают:

1. систематическое использование компьютерных презентаций (как преподавателем в установочной части занятия, так и студентом, выступающим с критическим разбором реферируемого научного сочинения);
2. он-лайн демонстрации работы с лингвистическими базами данных и энциклопедическими интернет-ресурсами (информационно-справочный ресурс по языкам мира Etnolog <http://www.ethnologue.com> (частично-платный ресурс, социолингвистическая, информация, ареал распространения, карты); типологическая база данных WALS, The World Atlas of Language Structures <http://wals.info>; информационно-справочный ресурс по языкам мира <http://glottolog.org> (генеалогическая аффилиация и обширная библиография); многоязычная платформа для составления конкордансов и исследования грамматической и лексической дистрибуции WordSketchEngine <https://www.sketchengine.co.uk/>; лексикологическая база данных WordNet (<http://wordnet.princeton.edu/>; электронная энциклопедия по лингвистике серии Oxford Research Encyclopedia <http://linguistics.oxfordre.com/>; и др.;
3. использование открытых он-лайн аудио- и видео обучающих ресурсов (для академического английского: www.cambridge.org/elt/english-for-academics);
4. практическую работу с электронными анализаторами речи и платформами лингвистического документирования (платформы E-Language Archiving Technology, ELAN <http://www.lat-mpi.eu/tools/elan>; коллекция электронных ресурсов на сайте Summer Institute of Linguistics (www.sil.org), в том числе, программа для акустического анализа речи Speech analyzer, программа IPA Help для обучения и пользования Международной фонетической транскрипцией, комплекс программ для полевой работы; программа для акустического анализа речи PRAAT (<http://www.fon.hum.uva.nl/praat>); открытое он-лайн издание по документированию и консервации языков Language Documentation & Conservation, LD&C <http://www.nflrc.hawaii.edu/ldc>).
5. В период временного приостановления посещения обучающимися помещений и территории РГГУ. для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:
 - видео-лекции;
 - онлайн-лекции в режиме реального времени;
 - электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
 - системы для электронного тестирования;
 - консультации с использованием телекоммуникационных средств.

5. Оценка планируемых результатов обучения

5.1. Система оценивания

При выставлении оценки в ведомость и в зачетную книжку преподаватель должен указать результат в соответствии с традиционной шкалой оценок и со шкалой оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

Распределение баллов по видам учебной деятельности таково:

- посещение семинарских занятий – до 8 баллов,
- уровень активности студента при подготовке к занятиям (конспектирование специальной литературы, готовность отвечать на вопросы по анализу кейсов, активное участие в дискуссиях, коллоквиумах и мозговом штурме и проч.) и во время проведения занятий (участие в обсуждениях и выполнении коллективных заданий) – всего до 32 баллов,
- качество выполнения контрольной работы (текущая аттестация) – до 20 баллов,
- успешность выполнения итогового творческого задания – до 40 баллов.

Оценка «зачтено» выставляется, если студент набрал в сумме не менее 50 баллов. Магистрант, не набравший в сумме 50 баллов, сдаёт зачёт по всему курсу и предъявляет преподавателю собственноручно написанные конспекты специальной литературы и выполненные домашние задания ко всем семинарам.

5.2. Критерии выставления оценок

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	«отлично»/ «зачтено (отлично)»/ «зачтено»	Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения. Свободно ориентируется в учебной и профессиональной литературе. Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной атте-

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>станции. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей. Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами. Достаточно хорошо ориентируется в учебной и профессиональной литературе. Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	«удовлетворительно»/ «зачтено (удовлетворительно)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F,FX	«неудовлетворительно»/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Баллы за участие в практических занятиях. Поскольку преподавание дисциплины строится на коллективном обсуждении значимых вопросов теории и практики, уже само присутствие на практических занятиях позволяет студенту составить общее представление о проблематике курса; поэтому собственно присутствие на практических занятиях оценивается (в сумме не более 10 баллов). Основные баллы (до 50) даёт активное участие в практических занятиях, которое может проявляться в следующих формах:

- 1) участие в обсуждении прочитанной литературы на практических занятиях (30 баллов);
- 2) выполнение практических заданий по программированию (20 баллов);

Баллы за итоговую аттестацию. Зачёт состоит из двух частей, в сумме оцениваемых в 40 баллов:

- 1) пересказ научной статьи по теме курса;
- 2) выполнение практического задания.

Типовые задания, используемые для оценки степени освоенности учебного материала

Даны три текста на английском языке, по экспертной оценке, имеющие разный уровень сложности. Оценить их сложность автоматически и сравнить с экспертной оценкой.

Текст 1

Amazing adventurers

Do you ever dream about climbing Mount Everest or visiting Antarctica? If so, you're not alone. Every year, thousands of people try to climb the world's highest mountains or walk across continents. Let's take a look at some of the 21st century's greatest adventurers.

Amazon adventurer

Ed Stafford from the UK is the first person to walk along the Amazon River from the mountains of Peru to the mouth of the river in Brazil. His amazing journey took two years and four months. There are many dangerous animals in the rainforest, like snakes and crocodiles, but Ed was lucky; he was only bitten by ants and mosquitoes. On his trip, Ed had to find fruit and nuts or catch fish each morning. Sometimes food was hard to find and Ed was often tired and hungry.

Technology was very important for Ed. He used a radio to ask the people of the rainforest for food and help. Many people came to meet him and guide him through the rainforest. While he walked, Ed wrote a blog to tell the world about climate change and destruction of the rainforest.

A mountain climber

Did you know that more than 4,000 people have climbed Everest? Gerlinde Kaltenbrunner from Austria is one of them. She is one of the world's greatest climbers and has climbed all the world's mountains over 8,000 metres. It's very difficult to climb in cold weather and storms, but Gerlinde loves it. She started climbing as a teenager in the mountains near her home. When she left school she became a nurse but always went climbing in her free time. Now she spends her time climbing and helping a charity for poor children in Nepal.

More than one adventure

Some adventurers are always looking for a new challenge. Meagan McGrath from Canada has climbed mountains, ridden a bike across Canada and run races in the desert. But her most incredible journey was a skiing trip to the South Pole. As she skied, Meagan pulled a sledge with a tent and all her food. She skied through terrible storms and freezing temperatures for forty days till she arrived at the South Pole.

Erik Weihenmayer from the United States has climbed mountains and ridden a bike through deserts. Amazingly, Erik is blind and he wants other blind people to have active lives too. He has taken groups of young blind people climbing in Nepal.

Where next?

Technology is a big help for adventurers but the world is still a dangerous place and it's very important to prepare well. If you dream of being an adventurer, there will always be continents to walk across and mountains to climb!

Текст 2

Amazing adventurers

Have you ever dreamt of climbing Mount Everest or visiting Antarctica? If so, you're not alone. Every year, thousands of people try to climb the world's highest mountains or walk across continents. In the past, explorers had compasses and maps, but today's adventurers have satellite phones and GPS. They also use their travels to let the world know about climate change and help people in the countries they visit. Let's take a look at some of the 21st century's greatest adventurers.

Amazon adventurer

Ed Stafford from the UK is the first person to walk the length of the Amazon River. He started by a small stream in the Andes mountains of Peru and arrived at the river's mouth in Brazil, two years and four months later. Snakes, crocodiles and jaguars live in the Amazon rainforest, so it's a dangerous place. Luckily, Ed avoided the big animals, but he was bitten by ants and mosquitoes every day. On his trip, Ed had to find food each morning. Sometimes the fruit, nuts and fish he ate were hard to find and Ed often felt weak and hungry.

Technology was essential for Ed. He used a radio to ask local people for food and permission to enter their land. Many of them came to meet him and guide him through the dense rainforest. As he walked, Ed wrote a blog about his daily experiences.

Ed used his walk to let the world know about climate change and raise money for environmental charities in Brazil and Peru.

A mountain climber

Four thousand climbers, aged between 13 and 80, have been to the top of Everest. Climbing high mountains requires a lot of preparation and is very dangerous, but some of the world's best climbers are now looking for new challenges.

Gerlinde Kaltenbrunner from Austria fell in love with climbing as a teenager. When she left school, she worked as a nurse and climbed in her free time. Starting with Everest, she has been climbing all the world's fourteen peaks over 8,000 metres. To increase the challenge, Gerlinde climbs without using oxygen tanks. Low oxygen levels can make climbers ill, so Gerlinde has to climb slowly. Gerlinde is passionate about Nepal and raises money for a charity for poor children and orphans there.

More than one adventure

Some of today's adventurers go from challenge to challenge. Meagan McGrath from Canada has climbed the highest mountain on each continent, ridden a bike across Canada and run a long-distance race in the Sahara Desert. But perhaps her most incredible journey was a skiing trip to the South Pole. She pulled a tent and all her food on a sledge behind her as she skied. On the first day, she fell into a glacier and had to be rescued. Many people would have given up, but Meagan decided to carry on. Skiing through ice storms, she arrived at the South Pole forty days later.

Erik Weihenmayer from the United States is another unstoppable adventurer. He's climbed mountains, ridden a bike through the deserts of Morocco and kayaked through the Grand Canyon. Amazingly, Erik has been blind since the age of 13. Apart from his travels, he helps people with disabilities to live active lives and takes groups of young blind people on climbing expeditions.

Where Next?

Despite new technologies, adventurers still have to live with terrible weather, lack of food and wild animals. Preparation and training are essential, but if you have a sense of adventure, there are continents to cross and hundreds of mountains to climb.

Текст 3

Amazing adventurers!

Have you ever dreamt of climbing Mount Everest or walking to the South Pole? If so, you're not alone. Every year, thousands of people try to climb the world's highest mountains or walk across continents. Unlike the explorers of the past who used maps and compasses, today's adventurers travel with modern technology like GPS and satellite phones. Many adventurers are nature lovers who use their travels to help raise awareness about a range of environmental issues, while others are keen to help people in need and raise money for charities. Let's take a look at some of the 21st century's greatest adventurers.

Amazon adventurer

Ed Stafford from the UK is the first person to walk the length of the Amazon River. He started by a small stream in the Andes mountains of Peru and arrived at the river's mouth in Brazil, two years and four months later, having walked 6,000 kilometres.

The Amazon rainforest is home to poisonous snakes, crocodiles and jaguars, so Ed was in constant danger. Luckily, he survived with nothing worse than a few thousand mosquito and ant bites. On his trip, Ed had to find food to eat every day. A lot of the time, the fruit, nuts and fish he ate were hard to find and he often felt weak and exhausted.

Ed's walk would have been impossible without technology. He used a radio to ask the people of the rainforest for food and permission to cross their land. Many of them came to meet him and helped guide him through the most difficult terrain. As he walked Ed wrote a blog, recording his day-to-day experiences. He used the media interest in his trip to protest about the destruction of the rainforest and raise money for environmental and children's charities in Brazil and Peru.

A mountain climber

Over 4,000 climbers, aged from thirteen to eighty have been to the top of Everest. Though climbing high mountains in freezing conditions and violent storms is still extremely dangerous, the world's best climbers now look for new challenges.

Gerlinde Kaltenbrunner from Austria fell in love with mountain climbing as a teenager. When she left school, she worked as a nurse, but kept climbing in her free time. Having climbed Everest, she decided to climb all fourteen of the world's 8,000 metre peaks. To increase the challenge, Gerlinde climbs without using oxygen tanks. This is risky as low oxygen levels at the top of high mountains can affect brain and body functioning. Gerlinde uses her fame as a climber to support a charity for poor children and orphans in Nepal.

Hungry for adventure

Not content with one amazing trip, some of today's adventurers go from challenge to challenge. Meagan McGrath from Canada has climbed the highest mountain on each continent, ridden a bike across Canada and run a long-distance race in the Sahara Desert in 45°C heat. But perhaps her most remarkable journey was a skiing trip to the South Pole. On the first day, she fell into a glacier and had to be rescued. Many of us would have given up at that point, but Meagan decided to carry on. She reached the South Pole forty days later, having pulled a sledge with a tent and all her food behind her through freezing conditions and ice storms.

Erik Weihenmayer from the United States is another multi-adventurer. He's ridden a bike through the deserts of Morocco, kayaked through the Grand Canyon and climbed Everest. Amazingly, Erik has been blind since the age of 13. Apart from his travels, he tries to encourage people with disabilities to live active lives and takes groups of young blind people on climbing expeditions.

Where next?

Despite new technologies, crossing continents and climbing mountains still has many risks. Preparation and fitness training are absolutely essential, but if you have a sense of adventure, there are endless possibilities and still hundreds of unclimbed peaks in the Andes and Himalayas.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Список источников и литературы

Основная литература

1. Мизернов И.Ю., Граценко Л.А. Анализ методов оценки сложности текста. // Новые информационные технологии в автоматизированных системах, 2015 – 10 стр.
2. Оборнева И.В. Автоматическая оценка сложности учебных текстов на основе статистических параметров. Автореферат, 2006, Москва – 20 стр.
3. Daskalu M. et al. ReaderBench, an Environment for Analyzing Text Complexity and Reading Strategies. AIED, 2013 – 10 p.
4. Hakuta C.K., Santos M. Understanding language. – Stanford University, 2012 – 145 p.
5. Hiebert E.H., Mesmer A.E. Upping the Ante of Text Complexity in the Common Core State Standards: Examining Its Potential Impact on Young Readers 2013 – 8 p.
6. Janusz Kacprzyk. Analyzing Discourse and Text Complexity for Learning and Collaborating // Studies in Computational Intelligence, Vol.534 – 2014 – 293 p.

Рекомендованная литература

1. Оборнева, И.В. Автоматизированная оценка сложности учебных текстов на основе статистических параметров: диссертация ... к.п.н. Москва, 2006.
2. Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association
3. Adams, M.J. (2009). The challenge of advanced texts: The interdependence of reading and learning. In E.H. Hiebert (Ed.), Reading more, reading better (pp. 163–189). New York: Guilford.
4. Alexander, P.A., Schallert, D.L., & Hare, V.C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. Review of Educational Research, 61(3), 315–343.
5. Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). Becoming a nation of readers: The report of the Commission on Reading. Champaign, IL: Center for the Study of Reading; Washington, DC: National Academy of Education, National Institute of Education.
6. Anderson, T.H., & Armbruster, B.B. (1984). Content area textbooks. In R.C. Anderson, J. Osborn, & R.J. Tierney (Eds.), Learning to read in American schools: Basal readers and content texts (pp. 193–224). Hillsdale, NJ: Erlbaum.
7. Armbruster, B.B. (1984). The problem of “inconsiderate text.” In G.G. Duffy, L.R. Roehler, & J. Mason (Eds.), Comprehension instruction: Perspectives and suggestions (pp. 202–217). New York: Longman.
8. Armbruster, B.B. (1996). Considerate texts. In D. Lapp, J. Flood, & N. Farnan (Eds.), Content area reading and learning: Instructional strategies (2nd ed., pp. 47–58). Boston: Allyn & Bacon.
9. Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. Language & Communication, 21(3), 285–301. doi:10.1016/S0271-5309(01)00005-2

10. Bakken, J.P., & Whedon, C.K. (2002). Teaching text structure to improve reading comprehension. *Intervention in School and Clinic*, 37(4), 229–233.
doi:10.1177/105345120203700406
11. Baumann, J.F. (1986). Effect of rewritten content textbook passages on middle grade students' comprehension of main ideas: Making the inconsiderate considerate. *Journal of Reading Behavior*, 18(1), 1–21.
12. Beck, I.L., McKeown, M.G., Omanson, R.C., & Pople, M.T. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. *Reading Research Quarterly*, 19(3), 263–277.
13. Bormuth, J. R. 1966. "Readability: A new approach." *Reading research quarterly* 1:79–132.
14. Bormuth, J. R. 1969. Development of readability analysis: Final Report, Project no 7-0052, Contract No. OEC-3-7-0070052-0326. Washington, D. C.: U. S. Office of Education, Bureau of Research, U. S. Department of Health, Education, and Welfare.
15. Bormuth, J. R. 1971. Development of standards of readability: Towards a rational criterion of passage performance. Washington, D. C.: U. S. Office of Education, Bureau of Research, U. S. Department of Health, Education, and Welfare.
16. Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts. *The Journal of Experimental Education*, 71(2), 126–148. doi:10.1080/00220970309602060
17. Brock, C.H., Boyd, F.B., & Moore, J.A. (2003). Variation in language and the use of language across contexts: Implications for literacy learning. In J. Flood, D. Lapp, J.R. Squire, & J.M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 446–458). Mahwah, NJ: Erlbaum.
18. Chall, J.S., & Dale, E. (1995). *Manual for the new Dale-Chall readability formula*. Cambridge, MA: Brookline.
19. Chall, J.S., Bissex, G.L., Conard, S.S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty: A practical guide for teachers and writers*. Cambridge, MA: Brookline.
20. Ciardiello, A.V. (2002). Helping adolescents understand cause/effect text structure in social studies. *The Social Studies*, 93(1), 31–36. doi:10.1080/00377990209599877
21. Cohen, S.A., & Steinberg, J.E. (1983). Effects of three types of vocabulary on readability of intermediate grade science textbooks: An application of Finn's transfer feature theory. *Reading Research Quarterly*, 19(1), 86–101.
22. Council of Chief State School Officers. (2012, January 26). *The Common Core State Standards: Supporting districts and teachers with text complexity* [Webinar]. Wasington, DC: Author. Retrieved from http://www.ccsso.org/Resources/Digital_Resources/The_Common_Core_State_Standards_Supporting_Districts_and_Teachers_with_Text_Complexity.html
23. Crossley, S.A., Dufty, D.F., McCarthy, P.M., & McNamara, D.S. (2007). Toward a new readability: A mixed model approach. In D.S. McNamara & J.G. Trafton (Eds.), *Proceedings of the 29th annual conference of the Cognitive Science Society* (pp. 197–202). Austin, TX: Cognitive Science Society.
24. Dale E, Chall J (1948). "A Formula for Predicting Readability". *Educational Research Bulletin* 27: 11–20+28
25. Dale, E., & O'Rourke, J. (1976). *The living word vocabulary, the words we know: A national vocabulary inventory*. Elgin, IL: Dome.

26. Davison, A., & Kantor, R.N. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17(2), 187–209.
27. Dolch, E. W. 1939. "Fact burden and reading difficulty." *Elementary English review* 16:135–138.
28. Dreher, M.J., & Singer, H. (1989). Friendly texts and text-friendly teachers. *Theory Into Practice*, 28(2), 98–104. doi:10.1080/00405848909543387
29. Duke, N.K. (2000). 3.6 minutes per day: The scarcity of informational texts in first
30. Felker, D. B., F. Pickering, V. R. Charrow, V. M. Holland, and J. C. Redish. 1981. Guidelines for document designers. Washington, D. C: American Institutes for Research.
31. Fisher, D., & Frey, N. (2009). Background knowledge: The missing piece of the comprehension puzzle. Portsmouth, NH: Heinemann.
32. Flesch, Rudolf. 1948. A new readability yardstick. *Journal of Applied Psychology* 32: 221–233.
33. Fry, E. (2002). Readability versus leveling. *The Reading Teacher*, 56(3), 286–291.
34. Fry, Edward. 1977. *Elementary Reading Instruction*. New York: McGraw-Hill.
35. grade. *Reading Research Quarterly*, 35(2), 202–224. doi:10.1598/RRQ.35.2.1
36. Graesser, A.C., McNamara, D.S., & Kulikowich, J.M. (2011). Coh-Metrix: Providing multi-level analyses of text characteristics. *Educational Researcher*, 40(5), 223–234. doi:10.3102/0013189X11413260
37. Graesser, A.C., McNamara, D.S., & Louwrese, M.M. (2011). Methods of automated text analysis. In M.L. Kamil, P.D. Pearson, E.B. Moje, & P.P. Afflerbach (Eds.), *Handbook of reading research* (Vol. IV, pp. 34–53). New York: Routledge.
38. Gray, W.S., & Leary, B.E. (1935). What makes a book readable, with special reference to adults of limited reading ability: An initial study. Chicago: University of Chicago Press.
39. Gunning, T. G. (2003). *Building Literacy in the Content Areas*. Boston: Allyn & Bacon.
40. Gunning, T.G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders*, 23(3), 175–189. doi:10.1097/00011363-200307000-00005
41. Halbert, M. G. 1944. "The teaching value of illustrated books." *American school board journal* 108, no. 5:43–44.
42. Hembree, R. (1992). Experiments and relational studies in problem solving: A metaanalysis. *Journal for Research in Mathematics Education*, 23(3), 242–273.
43. Hiebert, E.H. (2011). The Common Core's staircase of text complexity: Getting the size of the first step right. *Reading Today*, 29(3), 26–27.
44. Hiebert, E.H., & Martin, L.A. (2001). The texts of beginning reading instruction. In S.B. Neuman & D.K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 361–376). New York: Guilford.
45. Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122–149. doi:10.1037/0033-295X.99.1.122
46. Kincaid JP, Braby R, Mears J (1988). "Electronic authoring and delivery of technical information". *Journal of Instructional Development* 11: 8–13. doi:10.1007/bf02904998
47. Kincaid, J. Peter, Robert P. Fishburne Jr., Richard L. Rogers & Brad S Chissom. 1975. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel]. *Research Branch Report 8-75*, Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN.

48. Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy Enlisted Personnel. Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.
49. Kintsch, W. (with Crothers, E.J. et al.). (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
50. Klare, G. R., J. E. Mabry, and L. M. Gustafson. 1955. "The relationship of patterning (underlining) to immediate retention and to acceptability of technical material." *Journal of applied psychology* 39, no 1:40–42.
51. Landauer, T.K., McNamara, D.S., Dennis, S., & Kintsch W. (Eds.). (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.
52. McNamara, D. S., Graesser, A. C., McCarthy, P. M., Cai Z. 2014. *Automated Evaluation of Text and Discourse with Coh-Matrix*. Cambridge University Press.
53. MetaMetrics. (n.d.). What does the Lexile® measure mean? Durham, NC: Author. Retrieved January 8, 2012, from www.lexile.com/m/uploads/downloadablepdfs/WhatDoestheLexileMeasureMean.pdf
54. National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects: Appendix A: Research supporting key elements of the standards and glossary of key terms*. Washington, DC: Authors.
55. Paul, D.J., Nibbelink, W.H., & Hoover, H.D. (1986). The effects of adjusting readability on the difficulty of mathematics story problems. *Journal for Research in Mathematics Education*, 17(3), 163–171.
56. School Renaissance Institute. (2000). *The ATOS™ readability formula for books and how it compares to other formulas*. Madison, WI: Author. Retrieved January 3, 2011, from www.windsorct.org/sagelmc/ReadabilityComparisonArticle.pdf
57. Smith, D.R., Stenner, A.J., Horabin, I., & Smith, M. (1989). *The Lexile scale in theory and practice: Final report*. Durham, NC: MetaMetrics. (ERIC Document Reproduction Service No. ED307577)
58. Stenner, A. J., I Horabin, D. R. Smith, and R. Smith. 1988. *The Lexile Framework*. Durham, NC: Metametrics.
59. Vernon, M. D. 1946. "Learning from graphic material." *British journal of psychology* 36:145–158.
60. Wikipedia. (2011). Brown corpus. Retrieved January 3, 2011, from en.wikipedia.org/wiki/Brown_Corpus
61. Williamson, G.L. (2006). *Aligning the journey with the destination: A model for K–16 reading standards*. Durham, NC: MetaMetrics.
62. Wolf, M.K., Herman, J.L., & Dietel, R. (2010). *Improving the validity of English language learner assessment systems: Full policy brief (CRESST Policy Brief No. 10)*. Los Angeles: National Center for Research on Evaluation, Standards, & Student Testing, University of California. Retrieved January 3, 2011, from www.cse.ucla.edu/products/policy/PolicyBrief10_FRep_ELLPolicy.pdf
63. Zeno, S.M., Ivens, S.H., Millard, R.T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Планы практических занятий

Практическое занятие 1. Обсуждение проблематики сложности текста.

Цель практического занятия: провести анализ различных подходов к проблематике сложности текстов и языков. Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Практическое занятие 2. Качественный и количественный подход к сложности текста

Цель практического занятия: провести анализ различных количественных подходов к сложности текста. Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Практическое занятие 3. Автоматическое извлечение признаков из текста.

Цель практического занятия: реализовать извлечения признаков с помощью языка Python. Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Практическое занятие 4. Формулы удобочитаемости.

Цель практического занятия: реализовать стандартные формулы удобочитаемости с помощью языка Python. Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Практическое занятие 5. Сравнение автоматических оценок текстов с системой CEFR.

Цель практического занятия: сопоставить автоматически полученные оценки с Обще-европейскими компетенциями владения иностранным языком (CEFR). Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Практическое занятие 6. Википедия как источник простых и сложных текстов.

Цель практического занятия: проанализировать автоматически собранный материал из Википедии на английском и простом английском языках. Выполнение практических заданий и разбор домашних заданий.

Литература:

Fisher, D., Frey N., Lapp, D. (2012). Text Complexity: Raising Rigor in Reading. Newark, DE: International Reading Association.

Материально-техническое обеспечение занятия.

Доска.

Методические рекомендации для студентов по освоению дисциплины

Курс «Автоматическая оценка сложности текстов» состоит из теоретической и практической части. Для выполнения практической части, включающей в себя реализацию алгоритмов оценки сложности, рекомендуется уверенное владение языками программирования.

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

№ п / п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR

7. Материально-техническое обеспечение дисциплины

Занятия по дисциплине можно проводить с максимальной эффективностью, если проводить их в компьютерном классе с доступом в Интернет, проектором и экраном для презентаций. Необходимо также наличие доски, чтобы преподаватель мог разбирать примеры по ходу объяснения и записывать задания.

Состав программного обеспечения

№п /п	Наименование ПО	Производитель	Способ распространения (<i>лицензионное или свободно распространяемое</i>)
1	Microsoft Share Point 2010	Microsoft	лицензионное
2	Windows 10 Pro	Microsoft	лицензионное
3	Kaspersky Endpoint Security	Kaspersky	лицензионное
4	Microsoft Office 2016	Microsoft	лицензионное
5	Zoom	Zoom	лицензионное

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

При необходимости рабочая программа дисциплины может быть адаптирована для обеспечения образовательного процесса лицам с ограниченными возможностями здоровья, в том числе для дистанционного обучения. Для этого от студента требуется представить заключение психолого-медико-педагогической комиссии (ПМПК) и личное заявление (заявление законного представителя).

В заключении ПМПК должно быть прописано:

- рекомендуемая учебная нагрузка на обучающегося (количество дней в неделю, часов в день);
- оборудование технических условий (при необходимости);
- сопровождение и (или) присутствие родителей (законных представителей) во время учебного процесса (при необходимости);
- организация психолого-педагогического сопровождение обучающегося с указанием специалистов и допустимой нагрузки (количества часов в неделю).

Для осуществления процедур текущего контроля успеваемости и промежуточной аттестации обучающихся, при необходимости могут быть созданы фонды оценочных средств, адаптированные для лиц с ограниченными возможностями здоровья и позволяющие оценить достижение ими запланированных в основной образовательной программе результатов обучения и уровень сформированности всех компетенций, заявленных в образовательной программе.

Форма проведения текущей и итоговой аттестации для лиц с ограниченными возможностями здоровья устанавливается с учетом индивидуальных психофизических особенностей (устно, письменно (на бумаге, на компьютере), в форме тестирования и т.п.). При необходимости студенту предоставляется дополнительное время для подготовки ответа на зачете или экзамене.

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
 - обеспечивается индивидуальное равномерное освещение не менее 300 люкс;
 - для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
 - письменные задания оформляются увеличенным шрифтом;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.
- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;
 - в форме аудиофайла.
- для глухих и слабослышащих:

- в печатной форме;
- в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
 - устройством для сканирования и чтения с камерой SARA CE;
 - дисплеем Брайля PAC Mate 20;
 - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
 - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемые эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина реализуется в Институте лингвистики УНЦ компьютерной лингвистики.

Цель дисциплины:

Предмет курса – современные представления компьютерной лингвистики об оценке сложности текстов. Цель курса – освоение студентами базовых понятий и методов оценки сложности текста в компьютерной лингвистике.

Задачи дисциплины:

- владение основными понятиями и категориями современной компьютерной лингвистики;
- владение основными методами автоматического анализа текста.

Дисциплина направлена на формирование следующих компетенций:

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-1 Владеет основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов	ПК-1.1	Знает основные понятия и категории современной лингвистики; основные методы научно-исследовательской деятельности в области фонологического, морфологического, синтаксического, дискурсивного и семантического анализа и правила их применения. Имеет представление об уровневой структуре естественного языка; основных параметрах разнообразия естественных языков; генетической, ареальной и типологической классификации языков; структуре лингвистической науки и ее основных направлениях; основных классических трудах по лингвистике
ПК-3 Способен использовать лингвистические технологии для проектирования систем автоматической обработки звучащей речи и письменного текста на естественном языке, лингвистических компонентов интеллектуальных и информационных электронных систем	ПК-3.3	Умеет пользоваться существующими системами автоматической обработки текста и звучащей речи, интеллектуальными и информационными электронными системами; проводить их сравнительный анализ; проектировать модули данных систем, составлять технические задания

По дисциплине предусмотрена промежуточная аттестация в форме зачета.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы.

ЛИСТ ИЗМЕНЕНИЙ

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
1	Приложение №		