

ПРОГРАММА ИТОГОВОГО ГОСУДАРСТВЕННОГО ЭКЗАМЕНА
по направлению
«Теоретическая и компьютерная лингвистика»
(магистратура)

1. Содержание дисциплины

Ниже представлены вопросы, входящие в состав экзаменационных билетов. Компоновка вопросов обновляются каждый год. Обновление состава вопросов происходит в соответствии с обновлением программ учебных курсов.

Билет состоит из двух вопросов, по одному из каждого блока. Время на подготовку устного ответа равняется одному астрономическому часу.

Блок I: Теоретические дисциплины

Теоретическая семантика

1. Семантика основных составляющих предложения
2. Правила семантической композиции
3. Референтные именные группы
4. Кванторные именные группы
5. Семантика глагола и прилагательного
6. Подъем квантора
7. Интенциональность и возможные миры
8. Два аспекта языкового значения: интенционал и экстенционал
9. Модальность и анализ модальных глаголов
10. Модальная база и источник порядка в анализе модальных выражений

Теоретический синтаксис

1. Архитектура грамматики в минимализме
2. Основные синтаксические операции: merge, move, agree (Chomsky 2000)
3. Признаковая база синтаксиса. Работа признаков в операции agree.
4. Теория падежа в Управлении и Связывании (приписывание падежа при управлении)
5. Теория падежа в минимализме (приписывание падежа при согласовании)
6. Альтернативные подходы к анализу падежа (падеж как признак времени, (Pesetsky&Torrego 2001), падеж как согласование по синтаксической категории (Pesetsky 2013))
7. Проблема quirky-подлежащих. Падеж и лицензирование именных групп
8. Эргативность и проблема морфологического падежа
9. Конфигурационная модель падежа Марантца (Marantz 1991)
10. Согласование в конфигурационных моделях (Bobaljik 2008)
11. Расширение конфигурационной модели в типологии Бейкера (Baker 2015): локальные области и условия, связанные с с-командованием
12. Модели согласования как успешной операции и обязательной операции
13. Ограничения на комбинации признаков лица и падежа (Person Case Constraint). Виды PCC, основные подходы к анализу PCC.

Теория и практика дискурсивного анализа

На письменном или устном дискурсе

На русском устном дискурсе

На русском мультимод. дискурсе

1. Что такое референция? Как это понятие соотносится с референциальным выбором? Приведите список внутриязыковых факторов, используемых для объяснения референциального выбора. Как эти факторы соотносятся с когнитивными основаниями референциального выбора? Приведите 2-3 примера, взятые из естественного дискурса, и подробно проанализируйте их со всех вышеупомянутых точек зрения.
2. Какие наиболее базовые типы редуцированных референциальных выражений встречаются в языках? Как они соотносятся по частотности среди языков мира? Приведите примеры из известных Вам языков и проанализируйте их.
3. Что такое устойчивые и чередующие местоимения? Как этот параметр соотносится с противопоставлением свободных и связанных местоимений? Приведите примеры из известных Вам языков и проанализируйте их.
4. Что такое сильные местоимения? Как они соотносятся со слабыми местоимениями в разных языках мира? В каких типах контекстов возникают сильные местоимения? Приведите примеры из известных Вам языков и проанализируйте их.
5. Что такое «чувствительности» (sensitivities) в стратегиях референции? Какие Вам известны непоследовательности языков в отношении таких признаков редуцированных референциальных выражений, как устойчивые/чередующиеся и связанные/свободные? Приведите примеры из известных Вам языков и проанализируйте их.
6. Что такое дискурсивная просодия? Самостоятельно выберите образец устного дискурса и покажите на этом материале максимальное число просодических явлений.
7. Что такое элементарная дискурсивная единица? По каким критериям выделяются ЭДЕ? В чем состоит когнитивное основание ЭДЕ? Приведите и подробно проанализируйте три-четыре примера из устного дискурса.
8. Какие есть типы содержательных элементарных дискурсивных единиц с точки зрения семантико-грамматической структуры? Как они соотносятся по частотности? Подберите и подробно проанализируйте три-четыре примера из устного дискурса.
9. Существует ли аналог предложения в устном дискурсе? Какие сведения нужно использовать, чтобы получить ответ на этот вопрос? Выберите образец устного дискурса и покажите, можно ли в этом материале выделить устные аналоги предложений.
10. Что такое мультимодальный (мультиканальный) дискурс? Как ведут себя прототипический говорящий и прототипический слушающий с точки зрения кинетического поведения? Выберите образец мультиканального дискурса и покажите, какие модальности и каналы задействованы в его структуре. Проанализируйте этот образец с точки зрения кинетического поведения говорящего и слушающего.
11. Какие аналогии можно провести между просодическим и жестовым поведением? В каких отношениях просодические знания помогают понять жестикуляцию, и наоборот? Есть ли в

жестикуляции аналоги вербального и просодического компонентов звучащей речи? Выберите образец мультимедийного дискурса и найдите примеры аналогичных явлений в просодии и в жестикуляции.

12. Какие существуют гипотезы относительно временного согласования между каналами? В чем проявляется такое согласование? Какими методами его можно проверить? Выберите образец мультимедийного дискурса и найдите примеры, показывающие согласования между каналами и его характеристики.

13. Важны ли для участника коммуникации жесты рук собеседника? Какие есть точки зрения по этому поводу? Каким образом это можно проверить? От каких факторов зависит то, какое внимание уделяется участниками коммуникации жестам собеседников? Насколько значительно индивидуальное варьирование в этом вопросе?

14. Что такое «портретная методология» при описании мультимедийного дискурса? Какие способы использования жестикуляционных портретов Вам известны? Какие факторы влияют на жестикуляционное поведение говорящего? Выберите образец мультимедийного дискурса и сформулируйте элементы жестикуляционного портрета одного или двух участников.

Теоретическая морфология и фонология

1. Лексическая фонология и цикличность
2. Автосегментная фонология и структура фонологического представления
3. Просодическая морфология и просодические ограничения на структуру словоформы
4. Распределенная морфология и раннее/позднее озвучивание
5. Алломорфия в распределенной морфологии
6. Теория оптимальности: основные принципы и теоретические допущения
7. Наносинтаксис: основные принципы и теоретические допущения

Психолингвистические модели

1. Междисциплинарность современной науки. Когнитивная наука как междисциплинарное образование. Психолингвистика как когнитивная наука. История психолингвистики по Лефельту (2013): дохомскианская эра. Традиционная история психолингвистики. Две психолингвистические традиции по Кларку (1992).
2. Краткая история изучения памяти в когнитивной психологии. Эксплицитная память. Рабочая память (Бэддели). Основные компоненты рабочей памяти. Вербальный компонент рабочей памяти. Тесты по определению объема рабочей памяти в психолингвистике: понимание vs. порождение речи. Тест Данеман и Карпентер 1980 г. Тесты по определению объема рабочей памяти на материале русского языка.
3. Изучение имплицитной памяти в психологии. Явление прайминга. Виды прайминга. Синтаксический прайминг. Повторение опорного слова. Синтаксический прайминг в экспериментах с детьми, билингвами и пациентами с языковой патологией. Функции и модели синтаксического прайминга. Индивидуальные различия испытуемых. Исследования с русскоязычными испытуемыми.

4. Сложноподчиненные предложения с придаточными времени. История исследований. Эксперименты с носителями разных языков. Понимание сложноподчиненных предложений с придаточными времени в исследованиях с детьми и пациентами с языковой патологией. Исследования с русскоязычными испытуемыми. Роль объема рабочей памяти.
5. Языковая неоднозначность. Значимость конструкций с синтаксической неоднозначностью для моделей синтаксического анализа. Локальная vs. глобальная синтаксическая неоднозначность. История исследований. Модель «садовой дорожки». Модель заблуждения (Фрэзер). Универсальность синтаксического парсера. Роль контекста. Разрешение синтаксической неоднозначности в разнотипных языках. Роль длины и структуры придаточного. Роль объема рабочей памяти. Просодическая гипотеза (Фодор).
6. Референциальный выбор: когнитивный многофакторный подход. Типология референциальных конфликтов. Влияние фактора синтаксической роли антецедента на разрешение референциального конфликта. Референциальный выбор в ситуации референциального конфликта. Модель референциального выбора. Факторы предотвращения референциального конфликта. Референциальные стратегии.
7. Вынесение vs. извлечение суждений о грамматической правильности или приемлемости предложений. Бесцветные зеленые идеи спят яростно. Экспериментальный синтаксис. Синтаксический эксперимент с точки зрения психолингвистики. Общие принципы работы с синтаксическим материалом.
8. Референциальная коммуникация: танграммы. История исследований. Исследование Кларка и Уилкес-Гиббс 1986 г. Совместная модель референции в диалоге. Общая позиция. Стратегии описания танграмм. Русские танграммы. «Взрослый» корпус. Корпус пациентов с афазией. «Детский» корпус. Мультимодальный корпус.
9. Порождение и понимание речи. Модель порождения речи Бок и Лефельта (1994). Эффект МакГурка. Эффект восстановления фонемы. Модель метрической сегментации. Логоген Мортон. Когортная модель В. Марслен–Вильсона. Коннекционистская модель TRACE.
10. Детская речь. Когнитивное развитие ребенка по Пиаже. Сенсомоторная стадия. Дооперациональная стадия. Стадия конкретных операций. Стадия формальных операций. Эгоцентризм Пиаже и Модель психического. Вопрос о врожденности языка. Этапы усвоения языка ребенком.
11. Нейролингвистика. История нейролингвистики. Патология речи. Афазиология. Модель А.Р. Лурии. Бостонская модель. Нарушения речи при шизофрении. Алалия. Нарушения речи при расстройствах аутистического спектра.
12. Нейролингвистика. Межполушарная асимметрия. Тест Вада. Расщепленный мозг. Прямая электрическая стимуляция коры мозга. Электросудорожная терапия. Метод дихотического прослушивания. Ген языка и специфические нарушения развития речи. Зеркальные нейроны.

Блок II: Лингвистические аспекты новых информационных технологий
Компьютерная морфология и синтаксис

1. Понятие языковой модели. Сглаживание, аддитивное сглаживание.
2. Интерполяция и откат в языковой модели. Метод Уиттена-Белла.
3. Проблемы энграммных моделей, модификации энграммных моделей.
4. Скрытая марковская модель. Алгоритм Витерби декодирования наиболее вероятной последовательности.
5. Особенности применения скрытых марковских моделей к морфологическому анализу. Вычисление вероятностей перехода и лексических вероятностей.
6. Условные случайные поля. Основная идея. Отличие от скрытых марковских моделей.
7. Алгоритм Нивре синтаксического анализа. Основные операции. Вариант алгоритма для непроективных деревьев.
8. Графовые методы автоматического синтаксического анализа. Алгоритм Чу-Лю-Эдмондса.
9. Векторные представления слов, различные варианты, их свойства. Обучение векторных представлений слов.
10. Нейронные сети в задачах обработки текста. Входное представление текста. Обучаемые векторные представления слов. Архитектура сети для текстовой классификации.
11. Свёрточные нейронные сети. Задачи, где они используются.
12. Рекуррентные нейронные сети: основные определения. Задачи, где они используются. Языковые модели на основе рекуррентных сетей.

Машинный перевод

1. Архитектура энкодер-декодер, задачи, где она используется.
2. Механизм внимания для задачи машинного перевода, основная идея и формулы.
3. Трансформерный слой, архитектура Трансформер, задачи, где она используется.
4. Предобучение больших языковых моделей: основная идеология. Модель BERT.
5. Предобучение больших языковых моделей: основная идеология. Модель GPT.
6. Предобучение больших языковых моделей вида энкодер-декодер: модель BART.
7. Декодирование в моделях энкодер-декодер. Поиск по лучу, его модификации.
8. Автоматическое исправление грамматических ошибок: основные методы.

Компьютерная семантика

1. Извлечение устойчивых словосочетаний. Типы ассоциативных мер.
2. Мера взаимной информации. Ее особенности. Варианты меры.
3. Термины и их свойства. Типы терминологических ресурсов. Различия в критериях отбора терминов. Признаки для автоматического извлечения терминов.
4. Тезаурусы, их структура и используемые отношения. Методы нахождения семантического сходства по тезаурусу.
5. Методы автоматического извлечения отношений гипоним-гипероним из текстов. Методы оценки качества извлечения этих отношений
6. Задача автоматического разрешения неоднозначности. Простые методы. Методы, основанные на знаниях. Методы машинного обучения в задаче автоматического разрешения неоднозначности.
7. Извлечение информации из текстов. Основные подзадачи. Подход, основанный на словарях и правилах. Этапы работы. Проблемы.
8. Извлечение именованных сущностей на основе машинного обучения. Подготовка обучающей коллекции. Применяемые методы. Признаки.
9. Задачи автоматического анализа тональности, варианты постановки задачи. Проблемы в

задаче анализа тональности.

10. Особенности словарей оценочной лексики. Подходы к автоматическому извлечению словаря оценочной лексики по корпусу. Машинное обучение в задаче анализа тональности.

2. Критерии оценки результатов сдачи итогового государственного экзамена

Каждый вопрос экзаменационного билета оценивается по десятибалльной шкале, где «10 баллов» соответствует отличному ответу на вопрос, «0 баллов» - отсутствию ответа на вопрос. Для получения итогового балла баллы за каждый вопрос суммируются.

Максимальное количество баллов у одного экзаменуемого составляет 20 баллов.

Критерии перевода баллов в оценку приведены в таблице:

17-20 баллов	«отлично»
13-16 баллов	«хорошо»
9-12 баллов	«удовлетворительно»
0-8 баллов	«неудовлетворительно»

В ряде случаев по решению председателя и членов ГЭК результат в 16 баллов может соответствовать оценке «отлично».

6. Литература

Основная и дополнительная литература для подготовки к экзамену соответствует спискам, приводимым в программах по соответствующим дисциплинам.

Программу разработал

Коллектив кафедры ТиПЛ филологического факультета МГУ имени М. В. Ломоносова под руководством заведующего кафедрой доктора филологических наук С. Г. Татевосова