

Е.А. Лютикова
lyutikova2008@gmail.com

I. Темы по корпусному исследованию синтаксиса русского языка

1. Селекция малых клауз со стороны лексических предикатов

Проблема: Несколько семантических групп матричных предикатов в русском языке (предикаты мнения, экспериенциальные предикаты, каузативные предикаты) допускают в качестве клаузального аргумента малую клаузу. В отличие от синтаксической валентности на инфинитивную клаузу, валентность на малую клаузу практически не обсуждается в литературе.

Исследовательский вопрос: Цель исследования – установить круг предикатов, обладающих актантной валентностью на малую клаузу, и определить, релевантно ли для таких конструкций противопоставление контроля и подъема.

Что предстоит делать: Провести корпусное исследование, составить базу примеров, дополнительно разметить примеры по релевантным параметрам, сформулировать обобщения относительно классов предикатов, обладающих селекцией на малую клаузу.

Уровень подготовки: Начиная со 2-го курса бакалавриата. Важны аккуратность, языковое чутье и интерес к синтаксису.

2. Ограничения на использование краткой формы прилагательного в малых клаузах

Проблема: В современном русском языке краткая форма прилагательного грамматична только в предикативной позиции. Помимо позиции финитного сказуемого, краткая форма может занимать позицию предиката в актантных и депиктивных малых клаузах. Однако дистрибуция краткой формы уже, чем дистрибуция актантных и депиктивных малых клауз.

Исследовательский вопрос: Цель исследования – установить дополнительные ограничения на использование краткой формы прилагательного в актантных и депиктивных малых клаузах.

Что предстоит делать: Провести корпусное исследование, составить базу примеров, дополнительно разметить примеры по релевантным параметрам, сформулировать обобщения относительно характеристик конструкций, в которых встречаются краткие прилагательные.

Уровень подготовки: Начиная со 2-го курса бакалавриата. Важны аккуратность, языковое чутье и интерес к синтаксису.

3. Варьирование позиции Вакернагелевых клитик в современном русском языке

Проблема: В современном русском языке клитики, тяготеющие к Вакернагелевой позиции (*ли, же, то* и нек. др.), демонстрируют синхронное варьирование в выборе хоста: они клитизируются к первому фонетическому слову либо к первой синтаксической группе своей клаузы. Неясно, одинакова ли дистрибуция этих вариантов для разных клитик.

Исследовательский вопрос: Цель исследования – установить, как устроена дистрибуция различных клаузальных клитик в отношении выбора хоста в современном русском языке.

Что предстоит делать: Провести корпусное исследование, составить базу примеров, дополнительно разметить примеры по релевантным параметрам, сформулировать обобщения относительно закономерностей выбора хоста.

Уровень подготовки: Начиная со 2-го курса бакалавриата. Важны аккуратность, языковое чутье, интерес к синтаксису и просодии.

II. Темы по синтаксису языков народов РФ

Темы по синтаксису тюркских, нахско-дагестанских, возможно уральских и некоторых индоевропейских языков по предварительному согласованию с научным руководителем. Обязательным требованием является доступ к работе с носителями исследуемого языка.

А.А. Герасимова, Е.А. Лютикова

1. Агент-ориентированное моделирование синтаксической вариативности

Проблема: Экспериментальные исследования синтаксической вариативности показывают, что носители отличаются по своим грамматическим предпочтениям. Часто выделяются группы носителей, которые предпочитают один из вариантов, а также группа носителей, не делающих различия между вариантами. Согласно теоретической концепции, связывающей формирование грамматики с процессом использования языка (т.н. *usage-based* подход), ожидается, что всякое взаимодействие носителей с использованием вариативной конструкции должно смещать их предпочтения в сторону использованного/услышанного варианта.

Исследовательский вопрос: Цель исследования – оценить динамику вариативного феномена в зависимости от характеристик групп носителей с помощью методов агент-ориентированного моделирования. В рамках модели симулируются взаимодействия между агентами, меняющие распределение вероятностей для грамматических вариантов. В зависимости от начальных характеристик агентов меняются итоги симуляций. Предполагается работа с моделью, в которой есть две группы носителей.

Что предстоит делать: Подготовить обзор литературы, использующей агент-ориентированное моделирование для исследования лингвистических задач; подобрать структуру модели (порядок симуляций), оптимальную для поставленной задачи; определить количественные метрики для оценки изменений; провести симуляции и обработать результаты. При успешном решении задачи и получении научно значимых результатов предполагается подготовка научной статьи в соавторстве с научным руководителем.

Уровень подготовки: Начиная с 3-го курса бакалавриата. Важны усидчивость, готовность к самостоятельному решению технических задач и активному поиску информации. Обязательны навыки программирования. Предполагается, что студент будет самостоятельно разрабатывать программный код и заниматься его отладкой (с опорой на литературу и опубликованные *supplement materials*).

2. Экспериментальное исследование взаимодействия типологической фреквенталии и грамматики родного языка (на материале позиции падежных показателей)

Проблема: Множественные исследования с применением методологии усвоения искусственных языков (англ. *artificial language learning* (ALL)) показывают, что типологически более распространенная стратегия усваивается быстрее альтернативных вариантов. Из этого делается вывод о том, что типологические универсалии и фреквенталии неслучайны: более частые стратегии предпочтительны

ввиду того, что они эффективнее обрабатываются когнитивной системой. При этом открытым остается вопрос о том, насколько качество усвоения различных типологически допустимых стратегий зависит от родного языка.

Исследовательский вопрос: Наша цель – установить, как динамика усвоения различных значений определенного типологического параметра в искусственном языке зависит от значения этого параметра в родном языке. В качестве параметра мы рассмотрим позицию падежных показателей в существительных: в языках мира они существенно чаще оказываются суффиксами, а не префиксами (см. WALS, URL: <https://wals.info/chapter/51>). Первый этап исследования (обязательный на первый год) – проверка скорости и качества усвоения параметра носителями русского языка (NB. и суффиксы, и предлоги). Привлечение носителей других языков предполагается на следующих этапах.

Что предстоит делать: Подготовить обзор литературы, посвященной (а) исследованиям типологических обобщений в рамках методологии ALL с участием носителей разных языков, (б) аналитическим подходам, моделирующим позицию падежных показателей в существительном, (в) психолингвистическим исследованиям обработки падежных показателей; подготовить и запустить экспериментальное исследование, обработать результаты. При успешном решении задачи и получении научно значимых результатов предполагается подготовка научной статьи в соавторстве с научным руководителем.

Уровень подготовки: Начиная с 3-го курса бакалавриата. Преимуществом будет опыт экспериментальных исследований и навыки обработки данных в R.

О.В. Федорова
olgfedorova@gmail.com

Любые темы по психолингвистике, усвоению языка, нейролингвистике и билингвизму, (в том числе с использованием метода регистрации движений глаз), в частности:

1. Роль вербальной рабочей памяти в языковом поведении человека
2. Метод референциальной коммуникации «Танграммы»: исследование совместного решения когнитивной задачи взрослыми испытуемыми или детьми 7-8 лет
3. «Фильм о грушах» У. Чейфа: жанр репортажа
4. Дискурсивное и когнитивное развитие русскоязычного ребенка младшего (7-8 лет) и среднего (11-12 лет) школьного возраста
5. Синтаксическое и дискурсивное развитие русскоязычных дошкольников
6. Синтаксические и дискурсивные особенности речи людей с РАС (совм. с МЭГ-центром)
7. Экспериментальное исследование семантики (совм. с С.Г. Татевосовым)
8. Дискурсивные особенности речи людей с психическими расстройствами (совм. с коллегами-психологами)

9. Синтаксический прайминг
10. Исследования цветообозначений
11. Исследования синтаксической неоднозначности
12. Исследования дискурсивной неоднозначности
13. Исследования синтаксической сложности
14. Исследования понимания предложений с придаточными времени
15. Когнитивное преимущество билингов

Для студентов II курса бакалавриата возможны как экспериментальные курсовые работы, так и обзорные курсовые работы в форме реферата. Для студентов III и IV курсов бакалавриата и студентов магистратуры – только проведение собственных исследований.

К.А. Студеникина

xeanst@gmail.com

I. Исследования по сравнению языковой способности носителей языка и больших языковых моделей

Конкретная тема обсуждается с научным руководителем. Ниже приведены некоторые варианты:

1. Способность больших языковых моделей оценивать приемлемость по бинарной и градуальной шкале
2. Сопоставительный анализ времени реакции носителей и весов внимания больших языковых моделей
3. Наличие эффектов аттракции при согласовании для больших языковых моделей
4. Обработка предложений с эффектом «садовой дорожки» носителями языка и большими языковыми моделями
5. Исследование аналогов синтаксического прайминга у больших языковых моделей

II. Оценка больших языковых моделей с точки зрения безопасности генерируемых текстов

Конкретная тема обсуждается с научным руководителем. Ниже приведены некоторые варианты:

1. Анализ социально-культурных стереотипов в ответах больших языковых моделей
2. Тестирование безопасности больших языковых моделей на материале jailbreak-промптов с использованием прагматических маркеров (таких как косвенные речевые акты или ложные пресуппозиции)
3. Исследование устойчивости больших языковых моделей к лингвистическим состязательным атакам (на примере подстановки синонимов, грамматических ошибок или нестандартной пунктуации)
4. Влияние морфосинтаксической сложности запроса к большой языковой модели на безопасность генерации (на материале пассивных конструкций, номинализаций и др.)

III. Применимость мультязычных больших языковых моделей для обработки текстов на малоресурсных языках России

Требуется знакомство или желание познакомиться с грамматикой языка, выбранного для работы. В качестве задачи обработки текста может быть взята классификация, извлечение именованных сущностей и т. д. Конкретная задача зависит от выбранного языка. Итоговая тема обсуждается с научным руководителем. Ниже приведены некоторые варианты:

1. Эффективность различных техник промптинга мультязычных больших языковых моделей для обработки текстов на малоресурсном языке
 2. Влияние генетической и типологической близости языков на эффективность кросс-лингвистического переноса для обработки текстов
 3. Сравнение методов адаптации словаря для эффективной токенизации малоресурсных языков в предобученных мультязычных моделях
 4. Проецирование и анализ языковых представлений малоресурсных языков в многомерном пространстве мультязычных моделей
- Для всех тем необходимо умение программировать на Python или желание освоить этот навык.

П.В. Гращенков
pavelvalgra@gmail.com

1. Корпусное исследование сочинения в русском и других языках

Предполагаются исследования согласования по ближайшему конъюнкту и влияния таких факторов, как одушевленность, совпадение признаков, род, расстояние до мишени согласования

2. Корпусное исследование эффектов тяжести в русском языке

Нужно будет собирать и анализировать данные по склонности к правому выносу отдельных типов проекций (именные группы, клаузы и проч.) в разных синтаксических контекстах в зависимости от количества материала в них

3. Грамматическая категория X в нивхском языке

Исследование по грамматикам и текстам грамматической семантики и правил дистрибуции именных (падеж, число, классификация) или глагольных категорий (вид, время, финитность, наклонение...)

4. Синтаксическое явление X в нивхском языке

Где X – структура именной / глагольной проекции, согласование, инкорпорация, контроль нулевых единиц, ... Исследование будет проводиться по грамматикам и текстам.

5. Автоматизация лексикографической / корпусной работы

На материале нивхского и/ли других малых языков будут исследованы проблемы, возникающие при преобразовании отгlossированных текстов в корпуса и словари или словники. Предполагается наличие навыков программирования.

6. Типология и синтаксис сочиненных составляющих

Будут исследованы проблемы, возникающие при анализе нетривиального синтаксического поведения сочинения, в первую очередь, – согласования по ближайшему / последнему конъюнкту, но не только. Главная исследовательская проблема – как сочинительная стратегия в конкретном языке связана с решением о структуре сочиненных составляющих. Работа будет проводиться на широком типологическом материале.

Ю.В. Николаева

julianikk@gmail.com

Предлагаемые темы указывают возможные направления исследований. Окончательный вариант темы формулируется после согласования с научным руководителем.

Исследования в рамках предлагаемых тем также возможны с использованием данных участников с афазией.

Билингвизм

1. Интерференция: фонетика, просодия, морфосинтаксис, структура дискурса.
2. Речевые формулы и языковая компетенция (формулы вежливости, ритуалы, оформление высказывания) и их освоение в неродном языке.

Просодия

1. Просодические признаки и их соотношение с кинетическими признаками жестов.
2. Просодические признаки и прагматические значения.
3. Просодические признаки и структура дискурса.
4. Просодический портрет говорящего.
5. Просодическое оформление речевых сбоев и лексического поиска.

Жестикуляция

1. Жестикуляция и прагматические значения.
2. Жестикуляция и морфосинтаксические характеристики.
3. Невербальные маркеры структуры дискурса.
4. Немануальные жесты (движения головы, ног, плеч, корпуса, мимика).
5. Совместное конструирование дискурса.
6. Перформативные значения в невербальных каналах.
7. Кинетическое сопровождение речевых сбоев.
8. Лексический поиск, жестикуляция и просодия.
9. Кинетический портрет говорящего.

С.А. Бурлак
svetlana.burlak@bk.ru

Темы по сравнительно-историческому языкознанию, языковым изменениям или старославянскому языку (по согласованию с преподавателем).

И.М. Кобозева
kobozeva@list.ru

Темы по лексической и грамматической семантике (предпочтительно по служебной лексике: коннекторам, частицам, вводным словам), прагматике, языку СМИ, лингвистическому анализу художественного текста (по согласованию с преподавателем).

А.А. Кибрик

Просьба связываться по электронной почте: aakibrik@gmail.com

I. Темы по **референции** в английском языке (на основе имеющегося **корпуса** газетных текстов) (желателен интерес к программированию и/или математическому моделированию и наличие базовых знаний в этих областях).

1. Статистические методы моделирования референциального выбора.
2. Выбор между вариантами полных именных групп.
3. Типы кореферентности.

II. Темы по исследованию **верхнекускоквимского** языка (атабаскская языковая семья, Аляска) (с использованием имеющихся словарных и текстовых материалов)

1. Оценка лексических различий между верхнекускоквимским языком и соседними атабаскскими языками.
2. Фонемная структура морфемы.
3. Просодические исследования в области верхнекускоквимского языка.
4. Порядок слов в клаузе.
5. Инвентарь и употребление послелогов.
6. Типы клауз в дискурсе.

III. Темы по аляскинскому русскому (реликтовый вариант русского языка, сформировавшийся на Аляске в XIX веке) (с использованием имеющихся словарных и текстовых материалов)

1. Лексические и грамматические особенности аляскинского русского.
2. Диалектные источники явлений, обнаруживаемых в аляскинском русском.
3. Просодические явления аляскинского русского.

IV. Темы по **устной** русской речи (на основе имеющихся корпусов устного дискурса и параллельного устно-письменного **корпуса** рассказов)

1. Просодические явления в устном дискурсе (возможен целый набор разных тем по различным типам явлений, например, по дискурсивным акцентам, двумеровости, тоновым регистрам, темпу, удлинению гласных, парентетической интонации...).
2. Порядок слов в устном дискурсе (возможен целый набор разных тем по порядку различных составляющих, например относительное расположение главного и зависимого слова, порядок в сочиненных группах, расположение частиц и т.д.)
3. Дискурсивные маркеры – их функции и дистрибуция.
4. Дискурсивные маркеры и риторические отношения – принципы соответствия (на примере одного или нескольких риторических отношений, например каузальных отношений).
5. Референциальный выбор (полные именные группы vs. местоимения, нулевые выражения)
6. Исследования в области мультимедийной коммуникации (например, типы клауз и соответствующих им жестов).

V. Темы по русскому **жестовому языку** глухих (РЖЯ)

1. Социолингвистическое исследование РЖЯ (региональные варианты, двуязычие, переключение кодов и т.д.)
2. Кивки головой, их грамматические и дискурсивные функции в РЖЯ.

3. Исследования в области различных грамматических и лексических аспектов РЖЯ.

VI. Темы в области социолингвистического исследования «Языки Москвы»

1. Лингвистическая демография Москвы.
2. История языкового разнообразия Москвы.
3. География распространения языков в Москве.
4. Использование языков в различных социальных сферах.

VII. Другие темы

1. Выражение определенности/неопределенности в русском языке (*можно в сопоставлении с артиклевými языками*)
2. Многопоточная коммуникация (одновременное использование телефона, электронной почты, мессенджеров, соцсетей)
3. Подлежащая референция в русском языке (*корпусное исследование; возможны варианты темы – чисто синхронный подход; исследование из истории русского языка; сопоставление русского материала с другими славянскими языками*)
4. Исследование нулевой анафоры в устном (или письменном) русском дискурсе (*факторы использования нулевой анафоры, в отличие от местоимений; различие между контекстами наст. и прош. времени; сопоставление с др. славянскими языками и т.д.*)
5. Исследование глагольной анафоры в русском языке. (В частности: взаимодействие глагольной и именной анафоры: *Я проголосовал за себя, и Петя тоже*).
6. Порядок слов SOV в русском языке (типа *Я его не видел*) (*можно в диахроническом плане, можно в типологическом и т.п.*)
7. Заимствованные словообразовательные модели в русском языке (*заимствования/калькирование из латинского, немецкого и других языков*)
8. История морфосинтаксической модели типа *рок-музыка* в русском языке

П. О. Россяйкин

petrossyaykin@gmail.com

Описательные и аналитические исследования семантики и семантико-синтаксического интерфейса на материале любых языков по следующим направлениям:

1. Неопределенные местоимения и связанные темы

Примеры тем:

- Семантика и дистрибуция серии неопределенных местоимений *У* в языке *Х*
- Диахроническое исследование дистрибуции серии местоимений *У* в русском языке
- Дистрибуция *либо*-местоимений в отрицательных контекстах

2. Модальные показатели

Примеры тем:

- Система модальности в языке *Х*
- Семантика показателей слабой необходимости в русском языке
- Свойства модальных глаголов *стоит*, *следовать* в отрицательных контекстах

3. Отрицательные показатели и конструкции

Примеры тем:

- Отрицание кванторов в русском языке
- Сфера действия конструкции *ни ... ни ...* в русском языке
- Возможность прочтений с разорванной сферой действия в предложениях с отрицательными местоимениями в русском языке

4. Скалярные фокусные показатели

Примеры тем:

- Семантика и дистрибуция скалярного фокусного показателя *У* в языке *Х*
- Дистрибуция показателя *почти*
- Свойства аддитивного следствия в предложениях с *даже* и *вообще*

5. Дизъюнктивные союзы

Примеры тем:

- Семантика и дистрибуция дизъюнкции *хоть ... хоть ...* в русском языке
- Сфера действия показателей дизъюнкции в отрицательных контекстах в русском языке
- Интерпретация дизъюнкции *то ли ... то ли ...* в предложениях с кванторами

Ю. В. Синицына

jv.sinitsyna@yandex.ru

По согласованию с руководителем возможны исследования в рамках следующих направлений:

1. Синтаксис и семантика именных категорий / полипредикативных конструкций в финно-угорских языках в типологической перспективе
2. Синтаксис и семантика сравнительных конструкций, полисемия сравнительных показателей
3. Социолингвистические исследования

А. А. Сорокин

alexey.sorokin@list.ru

- 1) исправление опечаток, грамматических ошибок и другие родственные задачи;
- 2) приложения компьютерной лингвистики в образовании;
- 3) контекстное обучение больших языковых моделей (прежде всего YandexGPT).
- 4) компьютерная морфология и синтаксис, в т. ч. автоматическое словоизменение, морфологический анализ и др.

По желанию студента возможен выбор любой другой разумной темы в рамках компьютерной лингвистики, инициатива студента приветствуется.

Во всех задачах требуется умение программировать на языке Python, а при необходимости — расширение этого умения. Обычно курсовая второго курса включает в себя работу с языковыми моделями с помощью промпт-инжиниринга, а третьего и более старших курсов — обучение нейронных сетей. Желательно также знание математики в рамках стандартного курса ОТиПЛа и стремление расширить это знание при необходимости.

С.Г. Татевосов

tatevosov@gmail.com

В 2024 году предлагаются работы, подпадающие под один из следующих научных сюжетов.

1. Деривационная морфология русского языка
2. Семантика и синтаксис «среднего залога»
3. Семантика дискурсивных частиц
4. Индискриминативы в русском языке
5. Семантика и синтаксис реципрока (в языке X)
6. Эвиденциальность (в языке X)
7. Отрицательно-полярные единицы (в русском языке / в языке X)
8. Семантика и синтаксис каузатива (в языке X)
9. Семантика и синтаксис антикаузатива (в языке X)
10. Видо-временная система (языка X)
11. Скалярные импликатуры
12. Семантика параметрических прилагательных

Примечание: «Язык X» — любой язык, кроме русского и английского; такие темы предполагают работу с носителями.

О.И. Беляев

belyaev@ossetic-studies.org

Темы по морфологии и синтаксису иранских языков (по согласованию с руководителем).

Н.В. Лукашевич

louk_nat@mail.ru

Курсовые связаны с задачами автоматической семантической обработки текстов. Понадобится программирование, знакомство (или желание разобраться) с методами машинного обучения

Основные тематические направления:

- разработка датасетов и методов для информационного поиска, включая информационный поиск для больших языковых моделей (retrieval augmented generation, RAG);
- обучение и тестирование больших языковых моделей с точки зрения учета культурных и правовых особенностей России;
- методы анализа тональности и аргументации, автоматизированное пополнение ресурсов для анализа тональности текстов;
- автоматические методы извлечения информации, знаний из текстов, включая извлечение словосочетаний, терминов, отношений;
- семантический анализ текстов, включая методы разрешения лексической неоднозначности.