



**РУССКИЙ ЯЗЫК ИЛИ РУССКИЕ ЯЗЫКИ В СТРАНАХ БАЛТИИ: К  
ПРОБЛЕМЕ СОЗДАНИЯ КОРПУСА**

**Таллиннский университет**

**Институт гуманитарных наук**

**Направление российских и восточноевропейских исследований**

**Таллинн, 16–17 сентября 2022 года**

**ТЕЗИСЫ ДОКЛАДОВ**

Составитель: И. В. Адамсон, PhD

**Таллинн 2022**

## КОНТАКТНЫЕ ЯВЛЕНИЯ В РУССКОМ ЯЗЫКЕ ЭСТОНИИ И ДАЛЬНЕЙШИЕ ПЕРСПЕКТИВЫ ИССЛЕДОВАНИЙ

**А. Вершик (Таллиннский университет, Таллинн, Эстония)**

Для того, чтобы развивать комплексные исследования вариантов русского языка в странах Балтии, следует решить, какие теоретические модели и подходы для этого уместны. Основная проблема состоит в чрезмерной дескриптивности. Дескриптивность, которая опирается на статические понятия вроде «заимствование», «калька» и пр., может привести к тому, что (1) исследование невозможно «перевести» на метаязык современной контактной лингвистики, а, значит, и сравнить с другими работами в этой области и (2) при отсутствии современного метаязыка, некоторые явления остаются без внимания (пример – т. н. смешанные копии (mixed copies)), и невозможно сказать, присутствуют ли они в материале. Возникает вопрос из области философии науки: как эмпирический материал и теоретические модели связаны между собой и как язык теоретических моделей влияет на сбор и интерпретацию данных.

В докладе предлагаются некоторые теоретические модели, которые применялись автором в исследовании влияния эстонского на русский в следующих жанрах: двуязычные телепередачи, устная речь на рынке, в магазине и т. д., и блоги. Эти модели не являются формальными (не основываются на ограничениях, а usage-based). Более общая модель описания контактной ситуации – это несколько модифицированный подход Thomason and Kaufman 1988 (сохранения языка или языковой сдвиг). Применение этой модели позволяет оценить, какое место на шкале занимает та или иная контактная ситуация (в данном случае, стадия небольших изменений в грамматике и семантике, копирование некоторых эстонских конструкций). Модели, которые позволяют описывать конкретные изменения – это модель копирования кодов (code-copying, Johanson 1993) и родственная модель PAT and MAT replication (Matras and Sakel 2007). Эти модели можно описать как функционально-когнитивные.

## ЛИТЕРАТУРА

Johanson, L. 1993. Code-copying in immigrant Turkish. In Guus Extra & Ludo Verhoeven (eds.), *Immigrant Languages in Europe*, 197–221. Clevedon/Philadelphia/Adelaide: Multilingual Matters.

Matras, Y. and Sakel, J. 2007. 7. Investigating the mechanism of pattern replication in language convergence. *Studies in Language* 31, 829–865.

Thomason, S.G. and Kaufman, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley & Los Angeles: University of California.

### **ЯЗЫК РУССКОЙ ДИАСПОРЫ: ЧЕТВЕРТЬ ВЕКА ИССЛЕДОВАНИЯ**

**И. П. Кюльмоя (Тартуский университет, Тарту, Эстония)**

В 1997 году в Тартуском университете состоялась первая конференция «Язык диаспоры: проблемы и перспективы», в которой участвовали исследователи из трех стран Балтии. На конференции были представлены в основном первые наблюдения над функционированием русского языка в этих странах (вне «метрополии»), а также обозначены перспективы дальнейших исследований. Позже конференции проводились также в Вильнюсе и Риге, а затем различным аспектам функционирования языка в диаспоре стали посвящать доклады и статьи на всем постсоветском пространстве, включая метрополию.

В представляемом докладе рассматривается основная тематика исследований русского языка диаспоры лингвистами Тарту и Таллинна, связи с подобными исследованиями в других странах, анализируется реализация намеченного за прошедшие 25 лет. Более пристальное внимание уделяется русскому языку старожильческого населения, в частности, диалекту староверов Эстонии, исследованиям говоров; созданным в Тартуском университете базам данных, а также проблеме сохранения русского языка в диаспоре.

## О ПЕРСПЕКТИВАХ СОЗДАНИЯ КОРПУСА ГОВОРА СТАРОВЕРОВ ПРИЧУДЬЯ

И. П. Кюльмоя, А. И. Рыко (Тартуский университет, Тарту, Эстония)

В русистике вопрос о создании диалектных баз данных был поставлен еще в 80-е годы прошлого века, а в течение последних десятилетий, с тех пор как появилась техническая возможность звуковой фиксации больших по объему текстов и инструменты для их автоматической обработки, разрабатываются проекты и создаются диалектные корпуса. С начала 2000-х годов Кафедра русского языка начала регулярный сбор материала по говору староверов Причудья и разработку баз данных по этому материалу.

В докладе пойдет речь о наиболее успешных проектах в области создания корпусов русской диалектной речи, а также о тех материалах по русским говорам старожильского населения Эстонии, которые собраны в течение последних 20 лет, и о возможностях представления этого материала в формате мультимедийного корпуса.

## К СОЗДАНИЮ КОРПУСА ВАРИАНТОВ РУССКОГО ЯЗЫКА (ТЕЗИСЫ К ОБСУЖДЕНИЮ)

М. Копотев (Хельсинкский университет, Хельсинки, Финляндия)

1. Редко когда научный проект совпадает с актуальными задачами деколонизации. Однако актуальность не должна заслонять научную перспективу, основу которой составляет объективность.
2. Одной из главных характеристик любого корпуса является **репрезентативность**, или максимально полное представление языка или его разновидностей. Создание корпуса вариантов русского языка должно базироваться на представлении о способах функционирования языка в определенных социальных условиях. Это значит, что до создания корпуса должна быть проведена социолингвистическая работа, которая позволит определить в каких сферах, в каких социальных стратах и в каких пропорциях русский язык используется в странах Балтии. Материал,

включенный в корпус, должен по возможности отражать реальные пропорции использования языка.

3. **Разметка корпуса** должна быть по возможности теоретически-нейтральной. Кроме стандартных лингвистических сведений (токенизация, лемматизация, морфологическая разметка и др.) разметка должна включать социолингвистические и территориальные характеристики. Параметры разметки следует обсудить и утвердить до сбора материала. Желательно, чтобы она была применима и к другим вариантам русского языка с целью дальнейшего пополнения корпуса.
4. **Публикация корпуса** в интернете требует технической подготовки на самых первых этапах. Результаты подготовки корпуса могут публиковаться постепенно, например, простые текстовые данные представляют ценность даже без лингвистической разметки, лемматизация данных важна и без морфологического аннотирования и так далее. Техническое планирование этих этапов позволит сэкономить ресурсы и время.
5. **Некоторые вопросы для обсуждения**
  - a. Корпус вариантов русского языка предполагает, что материалы корпуса созданы билингвами и/или носителям эритажного русского. В корпусе должно быть предусмотрено специфическое аннотирование такого рода: переключение кодов (code-switching), языковой бэкграунд, уровень владения русским языком и т. д. Эта часть аннотирования, видимо, не будет теоретически-нейтральной.
  - b. Разметка вариантов языкового употребления, например калек или лексических заимствований. С одной стороны, именно они составляют специфику корпуса, с другой, – разметка может превратиться в субъективное выделение уже известных явлений и игнорирование неизвестных. Кроме этого, существуют явления, например, орфографические варианты, которые трудно искать, если они не приведены к стандартной форме, то есть к Орфографическому словарю ИРЯ РАН, ориентация на который может вызвать сомнения в нейтральности.

**ПОДКОРПУС РЕЧИ УЧЕНИКОВ-БИЛИНГВОВ (РУССКО-ЭСТОНСКИЙ БИЛИНГВИЗМ) В  
КОРПУСЕ ДЕТСКОЙ РЕЧИ RUSLAPSED: ПРОБЛЕМЫ ОТРАЖЕНИЯ УСТНОЙ РЕЧИ И  
КВАЛИФИКАЦИИ ОШИБОК**

**О. Н. Бурдакова , Л. Д. Николаева, Е. Ю. Нымм (Нарвский колледж Тартуского  
университета, Нарва, Эстония)**

С 2019 г. в Нарвском колледже Тартуского университета ведется работа над созданием электронного корпуса русской речи детей дошкольного и школьного возраста, проживающих в Эстонии и воспитывающихся в семьях, где одним из языков домашнего общения является русский.

Корпус RusLAPSED многоцелевой и исследовательский, динамический (т. е. периодически обновляемый и наращивающий объем), полнотекстовый, смешанный по типу языковых данных (включающий расшифровки записей устной речи и письменные тексты) и размеченный (индексированный, аннотированный).

Корпус составляют записи русской речи детей, собираемые в Нарвском колледже ТУ начиная с 2005 г. С 2022 г. в рамках проекта «Компьютерная система по анализу ошибок в устной и письменной речи детей-билингвов и рекомендациям по их корректировке и пропедевтике (на материале корпуса детской речи RusLAPSED)» ведется целенаправленная работа по пополнению корпуса записями письменной и устной речи учеников, получающих образование на эстонском языке.

Принимая во внимание многоцелевой характер создаваемого корпуса (ориентация на две основные целевые группы — исследователи-лингвисты и учителя-методисты), разработчики корпуса взяли установку на точное воспроизведение текстов при «переводе» из устной формы или рукописной в электронную печатную (для изучения всех уровней текста, кроме фонетического) и на создание точной разметки ошибок, которая позволила бы в автоматическом режиме дать детальные и конструктивные рекомендации по предупреждению и исправлению конкретных ошибок в устной и письменной речи.

Эти установки потребовали, во-первых, выработки стандарта для передачи особенностей, свойственных речи билингва (иноязычные вкрапления, лексические заимствования,

кальки, фонетические особенности речи, зеркальное написание некоторых букв и т. д.), и для передачи этапов порождения речи в устной и письменной форме — самоперебивы, исправления, паузы хезитации (в устной речи) и самоисправления, вставки, пропуски (в письменных текстах).

Во-вторых, решение прикладной методической задачи (создание анализатора ошибок) требует теоретического осмысления важных проблем: 1) на какую норму ориентироваться в квалификации целого ряда лексических и грамматических явлений, которые характерны для регионального варианта русского языка в Эстонии, но не кодифицированы; 2) как отражать в корпусе подвижность языковой нормы и смену вариантов нормы. Решение этих проблем будет предложено в докладе.

## **ЭМПИРИЧЕСКИЙ МАТЕРИАЛ КАК ОСНОВА ДЛЯ КОРПУСА ТЕКСТОВ НА РУССКОМ ЯЗЫКЕ В ЭСТОНИИ**

**И. В. Адамсон (Таллиннский университет, Таллинн, Эстония)**

В сообщении речь пойдет о материале, который собирался автором на протяжении более, чем 10 лет, примерно с 2008 года. Это – примеры слов и словосочетаний с «локальным» значением, встретившимися в текстах русскоязычных СМИ Эстонии. Интерес к такому материалу возник по объективным причинам. В частности, при работе с русским языком неизбежно возникают вопросы об уместности употребления слова или словосочетания с точки зрения нормы и стилистики. Вопросы об этом задавались и задаются, например, журналистами, преподавателями, переводчиками.

Кроме того, существует и исследовательский интерес к такому объекту исследования. Например, в Тартуском университете многократно проводились конференции, посвященные языку диаспоры, где с различных сторон рассматривалась данная проблематика. В Таллиннском университете этот объект исследования находится в поле зрения социолингвистики.

Автором сообщения был написан ряд статей на тему о функционировании лексики, свойственной эстонскому русскоязычному дискурсу. Для получения объективных данных

использовались как печатные, так и электронные источники. На первом этапе удобным ресурсом оказалась система баз данных Интегрум, при помощи которой можно зафиксировать частоту употребления, а также время вхождения в тексты на основании содержащегося в системе материала. Например, Национальный корпус русского языка не дает должного объема информации об искомым единицах. Подобный поиск можно продолжить с использованием и других крупных ресурсов, данные которых пока еще необходимо проанализировать.

В результате проделанной работы был составлен перечень регулярно функционирующих слов в текстах на русском языке, опубликованных либо в самой Эстонии, либо относящихся к Эстонии. Его можно посмотреть здесь: [venekeelest.weebly.com](http://venekeelest.weebly.com). Также можно ознакомиться со списком публикаций автора по данной теме. Указанный веб-сайт сейчас находится в процессе обновления. Для усовершенствования интерпретации материала не хватает представительного корпуса, который позволил бы делать более глубокие наблюдения. Также требуется внесение примеров употребления, их обработка и систематизация.

## **ОПЫТ ВЫЯВЛЕНИЯ ЛЕКСИЧЕСКИХ ОСОБЕННОСТЕЙ РУССКОГО ЯЗЫКА СТРАН БАЛТИИ НА МАТЕРИАЛЕ ГИКРЯ**

**Я. А. Яковлева (Российский государственный гуманитарный университет, Москва,  
Россия)**

Мы исследуем современный русский язык (РЯ) стран Балтии средствами Генерального интернет-корпуса русского языка (ГИКРЯ), разрабатываемого с 2014 г. Это динамический мегакорпус (б. 20 млрд слов), созданный и размеченный автоматически; его тексты фиксируют и «старую» региональную специфику, и инновации.

Дифференциальный характер ГИКРЯ позволяет учитывать возрастную (ВКонтакте и Живой Журнал), гендерную (только ВК) и региональную специфику авторов. Объем текстов из стран Балтии в сегментах ВК и ЖЖ — ок. 9 и 117 млн слов соответственно.

Выявляется как лексика, специфичная для отдельных стран, — *тариба* ‘совет, государственный орган’, *шакотис* ‘торт’ (Литва), *бунжа* ‘емкость’, *ринда* ‘очередь’ (Латвия),



*коолитус* ‘курсы’, *хайгекасса* ‘больничная касса’ (Эстония), так и «межгосударственная» — *рядный дом* ‘дом на несколько квартир с отдельным входом для каждой’, (Латвия, Эстония).

РЯ стран Балтии исторически связан с вариантами РЯ соседних территорий бывшего СССР, ср. *файно* ‘хорошо, отлично’ (Литва, Латвия, Белоруссия, Украина; полонизм), *ижно* ‘даже’ (Латвия, Псковская обл.), *хабарик* ‘окурок’ (Эстония, Петербург). Есть и давние номинации, характерные лишь для одной страны: *битловка* ‘водолазка’ — только Латвия (ср. *гольф* в том же знач.: Литва, Белоруссия, Украина).

В ГИКРЯ имеются десятки и сотни вхождений слов, характерных для Балтии. Эта лексика встречается и в текстах из России, но специфически балтийские единицы выявляются не только тематически (тексты о странах Балтии), но и статистически.

Подчеркнем, что анализ материалов ГИКРЯ по существу лишь начинается и можно смело прогнозировать приращение знаний о специфике вариантов РЯ Балтии.

## **ФОРМИРОВАНИЕ РЕГИОНАЛЬНОГО ВАРИАНТА РУССКОГО ЯЗЫКА ЛАТВИИ: ПРОБЛЕМА ИСТОЧНИКОВ<sup>1</sup>.**

**И. Кошкин (Латвийский университет, Рига, Латвия), Т. Стойкова (Вентспилсская высшая школа, Вентспилс, Латвия)**

1. Для региональных лингвистических исследований важное значение имеют источники, содержащие необходимую информацию для описания особенностей местного варианта русского языка. Выявление таких источников и их анализ становятся самостоятельной научной проблемой, так как актуальным является поиск языкового материала, который можно считать релевантным для характеристики именно регионального варианта языка. С проблемой источников связана и проблема создания корпуса. Источниковедческую базу для выявления и описания особенностей русского языка Латвии с точки зрения их формирования составляют следующие типы источников: а) лексика латышского языка как контактирующего языка (латышский литературный язык, латышская разговорная и

---

<sup>1</sup> This work is supported by state research programme “Letonika – Fostering a Latvian and European Society” project “Diversity of Latvian Language in Time and Space” (№ VPP-LETONIKA-2021/4-0003).

диалектная речь, язык латышской литературы, латышская лексикография), б) тексты русскоязычных публицистических изданий, в) региональные лексикографические источники, включающие материал русского языка.

2. Лексика латышского языка во всем ее историческом и социально-функциональном многообразии дает представление о роли и особенностях русского языка в Латвии в разные периоды контактов русского и латышского языков. Лексические заимствования разных видов (собственно заимствования, кальки) позволяют сделать вывод о характере языковой ситуации и направлениях языковых контактов (языковая ситуация социального билингвизма, двунаправленный характер контактирования). Особенно большое значение в этой связи имеют просторечно-разговорная и диалектная лексика, а также язык художественных произведений (например, язык и стиль изданного в первые десятилетия XX в. романа „Rīga” («Рига»), автором которого является латышский писатель Август Деглавс). Сопоставление параллельных контекстов латышских и русских периодических изданий (например, выходивших в 30-е годы XX века латышских газет „Jaunākās Ziņas” («Последние новости»), „Rīts” («Утро») и русской газеты «Сегодня») позволяет выявить явления интерференции и другие контактные изменения в региональном варианте русского языка.

3. Тексты русскоязычных публицистических изданий отражают тенденции, приводящие к формированию специфических особенностей регионального варианта русского языка Латвии. Особенно это характерно для языка русской публицистики XX века, так как возникли определенные исторические условия для углубления противопоставления языка метрополии и местного варианта. Например, такое противопоставление отчетливо проявляется при анализе образования и употребления имен существительных русского языка с общим значением лица. В русском языке Латвии 1918–1940 гг., судя по данным языка газет (газета «Сегодня» и др.), наблюдается варьирование словообразовательных синонимов-дуплетов, один из которых пришел из языка метрополии (*спортсмен, конторщик* и др.), второй же в языке метрополии неизвестен и является новообразованием в местном варианте (*спортист, конторист* и др.).

4. Важным источником по изучению особенностей регионального русского языка являются лексикографические и краеведческие материалы – как те, которые относятся к периоду XIX в. – первой половины XX в., так и те, которые собраны и опубликованы исследователями в последние десятилетия. Анализ этих источников позволяет выявить слова-регионализмы (в отличие от диалектизмов и архаизмов русского языка, встречающихся на более широкой территории распространения русского языка), а также лексические элементы, свойственные языку определенной социально-этнографической группы русскоязычного населения Латвии (имеется в виду язык староверов).

## **РУССКИЙ ЯЗЫК В ЛИТВЕ: К ВОПРОСУ О КОДИФИКАЦИИ**

**Е. Коницкая, А. Беловодская (Вильнюсский университет, Вильнюс, Литва)**

Как и всякий язык, существующий вне пределов метрополии, русский язык Литвы отличается некоторыми особенностями (напр., в области написания литовских имен собственных, учитывающем специфику фонетической системы литовского языка: *Чюрленис, Маркучяй* и т. п.), что неоднократно отмечалось литовскими русистами (в частности, в исследованиях Б. М. Синочкиной).

Восстановление государственной независимости Литвы и появление новых реалий в жизни меняющегося общества активизировали процессы расхождения русского языка Литвы и нормативного литературного русского языка. Усилению подобной дивергенции способствует влияние инонационального окружения со стороны как государственного, так и родственных славянских языков (польского, белорусского, украинского), причем присутствие последних в Литве стало намного более заметным в результате интенсификации миграционных процессов. Сказанное выше позволяет рассматривать «литовский русский» язык как один из вариантов плюрицентрического русского языка, в котором, в силу особенностей языкового ландшафта нашей страны и специфики функционирования русского языка в Литве, стихийно складываются узловые нормы, требующие компетентной лингвистической оценки и кодификации. Соответственно, актуальным становится решение следующих задач:

- унификация принципов транслитерации литовских имен собственных;
- унификация наименований литовских общественных институтов, государственных учреждений и т. п.;
- фиксация появившихся в русском языке Литвы политкорректных наименований, а также заимствований и калек с литовского языка;
- фиксация прецедентных феноменов, активно используемых в текстах СМИ и отражающих специфику местного культурного фона;
- фиксация морфологических и синтаксических особенностей русского языка в Литве.

Кодификация русского языка Литвы необходима для успешного взаимодействия различных социальных институтов, образовательных учреждений и СМИ, в сферу действия которых попадает работа с русскоязычным населением Литвы. В свою очередь, нормализация использования русского языка в публичном пространстве позволит сделать более качественным использование русского языка на уровне повседневной коммуникации жителей Литвы.

### **ЛИНГВИСТИЧЕСКАЯ БИОГРАФИЯ СТУДЕНТОВ-БИЛИНГВОВ**

**А. Бергманн (Гумбольдтский университет, Берлин, Германия), Е. Протасова  
(Хельсинкский университет, Хельсинки, Финляндия)**

Эксперты в области овладения вторым языком рекомендуют преподавателям знакомиться с лингвистическими биографиями обучаемых, т. е. их семейным положением, контекстами встреч с изучаемым языком и другими языками, используемыми в повседневной жизни, а также изучать их языковые установки и учитывать тип сообщества, в котором живут студенты. Преподаватели должны замечать и анализировать культурные различия, которые могут разделить сообщества, а также возможные социокультурные конфликты. Это должно быть признано в учебной программе, что влечет за собой неизбежную осторожность при подходе к реальному обучению. Даже если учащийся вырос в другой культуре, любой педагог должен выяснить, какие обстоятельства сформировали

индивидуальную языковую идентичность. Важно знать, чем отличается владение русским языком у обучаемых от владения другими языками, как студенты выбирают разные языки в разных предметных областях и в каких контекстах использование языковых разновидностей может быть причиной непонимания и коммуникативных неудач.

Настоящее исследование построено на корпусе сочинений двуязычных студентов, изучающих русский язык в Германии и Финляндии. Помимо особенностей русской речи, изучаются причины, по которым обучаемые выбирают этот предмет, особенности самоанализа и место русского языка как домашнего среди других языков студента. Обращается внимание на прогресс в написании сочинений, индивидуальные особенности письменной речи и персеверирующие проблемы. В контрастивном аспекте рассматриваются ошибки студентов, для которых другим повседневно используемым языком является немецкий или финский. Делается попытка объяснить наблюдаемые различия.

В итоге рассмотрения сочинений выясняется, что многоязычие молодых людей носит отпечаток того, что в детстве с ними говорили по-русски. Культура – это то, что притягивает и отталкивает. Хотя на общение с родными потрачено много времени, результат не всегда такой, какого они ждут, и разочарование может отрицательно сказываться на их установках.

#### **ЧТО ВЛИЯЕТ НА ЭТНИЧЕСКУЮ САМОИДЕНТИФИКАЦИЮ ДВУЯЗЫЧНЫХ УЧАЩИХСЯ: НЕКОТОРЫЕ НАБЛЮДЕНИЯ**

**Н. В. Мальцева-Замковая, И. М. Моисеенко (Таллиннский университет, Таллинн,  
Эстония)**

В докладе будут рассмотрены факторы, оказывающие влияние на отождествление себя с определенной национальной группой. В психологии «этническая идентификация является динамическим, многоуровневым конструктом, который соотносится с самоопределением, или чувством осознания себя как члена определенной этнической группы» (Phinney J.S.

2000). В социологии «национальный идентитет» – это соотнесение своего «я», самоотождествление с определенной группой (Parsons T. 1977).

В анкетировании в 2008 и 2018 годах приняло участие по 25 человек. На вопрос, кем они считают себя по национальности, были получены следующие ответы:

Этническая принадлежность	Количество ответов	
	2008	2018
эстонец	7	8
русский	7	7
русский эстонец	4	1
эстонец и русский	0	5
славянин	1	0
украинец	1	1
осетин	1	0
немец, эстонец, русский	0	2
не отнесли себя ни к одной национальности	3	1
не ответил на вопрос	1	0

Полученные в 2018 году данные свидетельствуют о том, что определенная часть учащихся относит себя одновременно к двум этническим группам, чего не наблюдалось десять лет назад. Все респонденты, кроме одного, относят себя к какой-либо национальности, хотя в 2008 году трое опрошенных не смогли этого сделать.

В задачи анкетирования также входило определение факторов, влияющих на самоидентификацию:

- а) родной язык (по собственной оценке);
- б) владение языками;
- в) язык в семье;
- г) родина родителей;
- д) самооценка владения русским языком;
- е) владение русским и эстонским языком (в сравнении);
- ж) язык обучения в школе

Наиболее значимыми из факторов по данным исследования 2008 года оказались следующие:

- язык обучения в школе,
- домашний язык,
- степень владения русским или эстонским языком.

Значимость перечисленных факторов, влияющих на этническую самоидентификацию учащихся-билингвов в разные годы, в сравнительном аспекте будет рассмотрена в сообщении.

## **РУССКОЯЗЫЧНЫЕ МЕДИА ЭСТОНИИ В АСПЕКТЕ ЛИНГВОЭКОЛОГИИ**

**Е. С. Михеева (Тартуский университет, Тарту, Эстония)**

Исследования медиа в различных аспектах представляются всегда актуальными, что обусловлено несколькими причинами. Прежде всего, отметим, что медиа представляют собой такой материал, который невозможно изучить сколь бы то ни было полно в современных условиях цифровизации и глобализации. Кроме того, в медиа при сохранении основного конструктивного принципа сочетания экспрессии и стандарта (В. Г. Костомаров) отражаются различные процессы, происходящие в современном коммуникативном пространстве метрополии и диаспоры, которые, в свою очередь,

вызывают различного рода видоизменения и трансформации медиа, «перестраивая» их для наиболее эффективного выполнения коммуникативных целей в определенном коммуникативном пространстве.

Сегодня мы являемся свидетелями становления языковых вариантов, все больше расходящихся с кодифицированной нормой метрополии в силу различных социальных условий, при этом кодифицированный «язык метрополии» – литературный русский язык продолжает являться образцом и неким «мерилом» в исследованиях языка диаспоры, особенно в нормативном аспекте. Возникает исследовательский вопрос – а правильно ли это и всегда ли нужно измерять нормативность того или иного языкового факта наличием или отсутствием его в кодифицированном русском литературном языке метрополии?

В докладе мы постараемся дать обоснованный ответ на этот вопрос. Для этого мы определяем понятие «лингвоэкология» и отделяем его от понятий «норма» и «кодификация», выделяем случаи неосознанного и осознанного нарушения норм в русскоязычных медиа Эстонии с точки зрения русского литературного языка метрополии, классифицируем случаи неосознанного нарушения нормы на типы «ошибок» и предпринимаем попытку понять, а все ли выделенные нами типы нарушения нормы являются «ошибками».

## **ЯЗЫКОВАЯ РЕФЛЕКСИЯ ДИАСПОРЫ: ПРОБЛЕМЫ САМОИДЕНТИФИКАЦИИ**

**Е. И. Костанди (Тартуский университет, Тарту, Эстония)**

Доклад продолжает анализ особенностей речи русской диаспоры Эстонии с точки зрения проявления в них черт языковой картины мира, формирующейся у носителей русского языка вне языковой метрополии в ситуации постоянных языковых и межкультурных контактов. К числу таких особенностей относится и языковая рефлексия, которая, сопровождая любую речевую деятельность, активизируется в условиях диаспоры.

Языковая рефлексия диаспоры косвенно затрагивалась во многих работах, посвященных особенностям русской речи в Эстонии, однако предметом специального анализа стала



сравнительно недавно – в последнее десятилетие. В фокусе внимания исследователей были, например, языковая рефлексия в конкретных функциональных стилях и сферах (разговорная речь, язык медиа, русские говоры Эстонии, художественная литература и др.) или ее отдельные аспекты и функции, например, сопоставительный аспект, прагматические функции. В проанализированном к настоящему времени материале обнаружены многочисленные и разнообразные метаязыковые единицы, посредством которых комментируются, характеризуются, оцениваются, сопоставляются разные аспекты языка и речи. Среди них частотны единицы, обусловленные функционированием русского языка в диаспоре. Некоторые результаты анализа последних уже представлены в ряде публикаций как автора доклада, так и других исследователей. В то же время ряд проявлений языковой рефлексии диаспоры не описан и не проанализирован.

В докладе будут рассмотрены метаязыковые единицы в устной и письменной русской речи в Эстонии, репрезентирующие проблемные аспекты языковой – тем самым и более общей – самоидентификации конкретных говорящих/пишущих, отдельных групп и диаспоры в целом. Эти аспекты имеют как достаточно очевидные проявления (комментирование и оценка знания языков, переключения кода, перевода и др.) и относительно однозначную интерпретацию, так и менее явные, порой не выраженные формально, реализующиеся в последствиях имплицитной языковой рефлексии (выбор языка общения, этикетных формул, коммуникативного поведения и др.).

**МЕТАЯЗЫКОВАЯ ФУНКЦИЯ ПЕРЕКЛЮЧЕНИЯ КОДА КАК ПРОЯВЛЕНИЕ  
САМОИДЕНТИФИКАЦИИ ГОВОРЯЩИХ (НА ОСНОВЕ ИНТЕРНЕТ-ОБЩЕНИЯ  
РУССКОЯЗЫЧНЫХ ЖИТЕЛЕЙ ИТАЛИИ И ЭСТОНИИ)**

**А. Деци (Тартуский университет, Тарту, Эстония)**

Самая очевидная характеристика письменной и устной речи в многоязычных условиях – переключение кода (ПК). Под ПК, при широком подходе, понимается любой переход с одного языка на другой в ходе общения. Он отражает языковую рефлексия говорящих, которые постоянно сопоставляют и выбирают языковые средства разных языков, выполняющие разнообразные прагматические задачи. Эти задачи включают, например,

выражение оценки языковых фактов и маркирование (языковой) идентичности. Это говорит о том, что метаязыковая функция ПК тесно связана с самоидентификации говорящих.

Языковая рефлексия, в том числе вопросы ПК и самоидентификации, в речи русскоязычной диаспоры Эстонии частично были описаны на основе разного материала, в меньшей степени в интернет-общении, где функции самовыражения и воздействия на других пользователей первичны. Изучение данного явления в интернет-дискурсе русскоязычной диаспоры Италии только начинается.

В центре внимания в докладе – сопоставление метаязыковой функции ПК в аспекте самоидентификации в интернет-дискурсе русскоязычных жителей Италии и Эстонии. Материал для анализа состоит из двух отдельных корпусов приблизительно в 600 комментариев, собранных на разных форумах и группах общения Фейсбука. Думается, что сопоставление языковых явлений и, в частности, функции ПК метаязыкового характера разных диаспор способствует лучшему осмыслению процессов, определяющих идентичность говорящих и дальнейшее языковое развитие диаспоры.

## **ОБ ОБЩИХ НАПРАВЛЕНИЯХ В РАБОТЕ НАД ЯЗЫКАМИ НАСЛЕДИЯ**

**М. Полинская (Мэрилендский университет, Мэриленд, США), К. Фридман (Бар-Иланский университет, Рамат-Ган, Израиль), Н. Меир (Бар-Иланский университет, Рамат-Ган, Израиль)**

В сфере билингвизма появляется все больше исследований, которые показывают, что грамматика языка наследия (heritage language, e. g., Polinsky 2018; Montrul & Polinsky 2021) отличается от грамматики базового языка (англ. baseline). Среди главных причин, которые ведут к грамматической дивергенции, можно выделить две: влияние доминантного языка (англ. cross-linguistic influence/transfer) и недостаточность входных данных (недостаточный инпут, ср. англ. diminished input), вследствие которой происходит внутреннее переструктурирование языка наследия. В качестве примера влияния доминантного языка можно привести позицию некоторых наречий в русском языке наследия с доминантными

романскими языками или греческим: так же, как в романских языках/греческом, появляется тенденция ставить наречия образа действия сразу после глагола, перед дополнением (*Маша быстро сварила кашу* → *Маша сварила быстро кашу*). Пример внутреннего переструктурирования при недостатке инпута в русском языке наследия— обобщение окончания [-ах] как показателя косвенных падежей мн. числа; такое развитие единого показателя, как кажется, не зависит от доминантного языка.

При исследовании грамматической дивергенции важно определить, какая из названных причин лежит в основе различий между базовым языком и языком наследия. Это предполагает знание того, как носители обоих вариантов трактуют данное конкретное явление, а также сравнение одного и того же языка наследия в контексте разных доминантных языков. Такое сравнение лишь относительно недавно стало применяться в работах по языкам наследия, и его методические принципы будут освещены в докладе. В качестве примера мы рассмотрим русский язык в Израиле (доминантный язык—иврит) и русский язык в США (доминантный язык — американский английский). В нашем исследовании мы сравнили порождение речи у разных русскоговорящих групп, живущих в Израиле и США. Экспериментальный контекст предполагал порождение согласования между именем и прилагательным (*золотой зонт, золотая чашка*), использование винительного падежа (*вижу жирафа*) и порождение числовых групп (*пять чашек*). Результаты показали эффект доминантного языка и некоторую роль беглости (беглость владения русским языком уменьшает эффект доминантного языка). Предложенный подход типологического сравнения будет полезен в работе над языками наследия, и мы надеемся на его более широкое применение в будущем.

## ЛИТЕРАТУРА

Montrul, S., & Polinsky, M., eds. 2021. *The Cambridge handbook of heritage languages and linguistics*. Cambridge: CUP; Polinsky, M. 2018. *Heritage languages and their speakers*. Cambridge: CUP.

## ОПЫТ СОЗДАНИЯ ВЕБ-КОРПУСА ТЕКСТОВ РУССКОЯЗЫЧНЫХ САЙТОВ ЭСТОНИИ ЗА ПЕРИОД С 2017 ПО 2021 ГОД

Елена Каллас (Институт эстонского языка, Таллинн, Эстония)

В докладе рассматриваются основные вопросы создания веб-корпуса текстов русскоязычных сайтов Эстонии за период с 2017 по 2021 год (Russian sites in Estonian Web 2017-2022), поисковый интерфейс которого базируется на корпусном менеджере Sketch Engine ([sketchengine.eu](http://sketchengine.eu)). Корпус был создан компанией-разработчиком Lexical Computing Ltd. в сотрудничестве с Институтом эстонского языка (eki.ee). Специфика корпуса заключается в том, что он состоит только из русскоязычных текстов Интернета, опубликованных на веб-страницах, зарегистрированных на домене .ee. Временной охват корпуса начинается в 2017 году и доходит до 2021 года. Корпус пополняется каждые два года. На данный момент объём корпуса составляет 300 миллионов словоупотреблений, которые распределены между пятнадцатью тематическими подкорпусами (форумы, блоги, спорт, здоровье, онлайн-торговля и др.). Разметка корпуса была осуществлена при помощи теггеров RFTagger and TreeTagger.

Веб-корпус текстов русскоязычных сайтов Эстонии позволяет отследить языковые изменения, происходящие в русском языке в Эстонии, и, в частности, используется при составлении русской части Большого словаря Института эстонского языка (EKI Ühendsõnastik). На примере конкретных слов и словосочетаний (*реновация, макра, семейный врач и др.*), мы продемонстрируем возможности статистической обработки корпусных данных, позволяющих отслеживать изменение частот и контекстов в различные периоды времени, а также осуществлять анализ лексико-синтаксической сочетаемости и парадигматических (синонимических и антонимических) связей слова.

### **КРУГЛЫЙ СТОЛ. Создание корпуса вариантов русского языка в различных странах: теоретические и практические проблемы и пути решения**

Ведущие: Капитолина Федорова, Наталия Чуйкина, Инна Адамсон

