

**ОТЗЫВ официального оппонента
о диссертации на соискание ученой степени
кандидата филологических наук Лемаева Владислава Игоревича
на тему: «Кросс-языковое распознавание эмоций в тексте и речи:
лингвистические аспекты»
по специальности 5.9.8. Теоретическая, прикладная и сравнительно-
сопоставительная лингвистика**

Диссертация В. И. Лемаева посвящена комплексному анализу влияния лингвистических признаков на качество межъязыкового переноса моделей для автоматического распознавания эмоций в письменном тексте и звучащей речи на русском языке.

Актуальность работы обусловлена рядом факторов. Во-первых, задача автоматического распознавания эмоций критически важна для создания адаптивных диалоговых систем, способных адекватно реагировать на эмоциональное состояние собеседника. Современное качество моделей распознавания требует дальнейшего совершенствования методов и подходов к решению этой задачи. Во-вторых, для многих языков мира наблюдается острая нехватка качественно размеченных данных, в частности, имеющих разметку по эмоциям, что делает межъязыковой перенос моделей с высокоресурсных языков весьма востребованной техникой. Эффективность данного подхода, однако, требует экспериментальной проверки в отношении каждого конкретного малоресурсного языка, включая русский, чтобы определить пределы его применимости. В-третьих, влияние собственно лингвистических свойств языков (синтаксис, лексика, графические способы выражения эмоций) на успешность такого переноса в задаче распознавания эмоций остается малоизученным, что образует лакуну, на восполнение которой направлена данная работа.

Представленная работа характеризуется высокой **степенью обоснованности положений, выносимых на защиту, научных выводов и рекомендаций, сформулированных в диссертации, которая обеспечивается следующим:**

- использованием комплексного методологического подхода, совмещающего методы теоретической и сопоставительной лингвистики (анализ типологических признаков языков, лексического сходства), корпусной лингвистики (работа с речевыми и текстовыми корпусами эмоциональных высказываний на 9 языках) и компьютерной лингвистики (применение и дообучение нейросетевых архитектур). Синтез методов позволяет получить верифицируемые данные о влиянии лингвистических факторов на качество классификации, недостижимые в рамках одного подхода;

- достаточно представительной теоретико-методологической базой исследования, включающей 128 источников, в том числе фундаментальные работы в области типологии языков, исследования по распознаванию эмоций, включая современные подходы на основе трансформеров, труды по созданию эмоциональных датасетов;

- корректно выстроенной методологией эксперимента, включающей набор современных архитектур (HuBERT, WavLM – для речи, XLM-RoBERTa – для текста) и многоязычные выборки данных. Использование сбалансированных сокращенных выборок для текстовых моделей, логика экспериментов с речевыми моделями, включающая сравнение результатов обучения моделей в нескольких конфигурациях (base, full, pretrain, pretrain_freeze, multi), позволяющих изолированно оценить влияние каждого фактора (объема данных, увеличения количества классов, предварительного обучения на больших речевых корпусах, мультязычного предобучения) на качество межъязыкового переноса, демонстрирует системный подход к проверке гипотез;

- репрезентативностью и объемом исследовательского материала: для оценки качества речевых моделей на русском языке привлечен крупнейший на данный момент корпус русской эмоциональной речи Dusha (300 000 записей, 350 часов), обеспечивающий надежность выводов о качестве распознавания отдельных эмоций; для оценки текстовых моделей

использованы русскоязычные выборки из корпусов CEDR (2000 примеров, 5 эмоций) и BRIGHTER (2400 примеров, 6 эмоций) со сбалансированным распределением обучающих и тестовых данных; для экспериментов по межъязыковому переносу привлечены четыре зарубежных речевых корпуса (IEMOCAP, nEMO, ESD, JVNv) с тестированием на выборке из корпуса Dusha, а также данные пяти языков из мультязычного корпуса BRIGHTER (немецкий, испанский, китайский, хинди, татарский) с тестированием на русскоязычной коллекции датасета. Объем материала, охват языков и сбалансированность выборок позволяет обеспечить корректность сопоставления результатов при оценке влияния лингвистических признаков на качество переноса моделей;

- применением комплексного качественно-количественного анализа материала: оценка значимости влияния лингвистических признаков (порядка слов, наличия отрицания, пунктуация) опирается на корректное применение статистического непараметрического теста PERMANOVA, подтверждающего объективность выявленных закономерностей, а также содержательный разбор конкретных примеров из корпусов;

- достаточным иллюстративным материалом, включающим подробные таблицы с количественными метриками качества (F-мера, Unweighted Accuracy) для всех эмоций и конфигураций обучения, графики, а также разбор конкретных примеров из корпусов, обеспечивающих прозрачность представленных выводов.

Достоверность научных положений, выводов и практических рекомендаций, сформулированных в диссертации В.И. Лемаева, подтверждается:

- использованием комплексного теоретико-методологического подхода, опирающегося на методы корпусной, сопоставительной и компьютерной лингвистики, а также машинного обучения;

- использованием в экспериментальной работе общепризнанных инструментов и методов исследования, включающих широко

апробированные архитектуры нейросетей, предобученные на крупных речевых и текстовых корпусах, метрики оценки качества обучения, используемые в задачах классификации, в частности, в области автоматического распознавания эмоций, что позволяет корректно сопоставлять результаты с данными других исследований;

- воспроизводимостью и верифицируемостью результатов, которые обеспечиваются общедоступностью используемых датасетов, четким описанием используемого набора гиперпараметров, применением тиражируемых процедур дообучения;

- репрезентативностью эмпирической базы, охватывающей сбалансированные по классам эмоций и полные (на отдельных этапах экспериментальной работы) выборки из речевых датасетов на 5 различных языках и многоязычного текстового датасета на 6 языках. Широкий охват языкового материала и его сбалансированность позволяют распространять выводы на широкий круг языков;

- апробацией результатов исследования в научном сообществе: основные положения диссертации докладывались на научном семинаре «Машинное обучение в автоматической обработке текстов» НИВЦ МГУ имени М.В. Ломоносова; результаты опубликованы в 5 рецензируемых научных работах в изданиях из перечня, рекомендованного для защиты в диссертационном совете МГУ им. М.В. Ломоносова.

Научная новизна исследования заключается в следующем:

- выявлено и систематизировано влияние некоторых лингвистических факторов (генетическая близость, сходство лексического состава, синтаксический тип языков и т.д.) на качество межъязыкового переноса моделей распознавания эмоций в устной речи на русском языке; показано, что ключевую роль играет сходство интонационных паттернов, тогда как лексика и синтаксис значимого влияния не оказывают;

- установлено влияние аналогичных лингвистических признаков на качество переноса моделей распознавания эмоций в письменных текстах;

доказано, что главным фактором является сходство лексических и графических маркеров, при этом тип письменности и базовый порядок слов значимого влияния не оказывают;

- экспериментально доказано, что набор лингвистических признаков, релевантных для успешного кросс-языкового переноса, различается для текстовых и речевых моделей распознавания эмоций.

Теоретическая значимость диссертации состоит в систематическом описании возможностей современных языковых моделей-трансформеров (HuBERT, WavLM, XLM-RoBERTa) для русского языка в задаче распознавания эмоций и при межъязыковом переносе. Автором выявлены и проранжированы по значимости лингвистические признаки, влияющие на успешность переноса, а также сформулированы проблемы, связанные с качеством и сбалансированностью существующих русскоязычных корпусов, ориентированных на задачу распознавания эмоций в тексте и речи. Работа вносит вклад в развитие теоретической, сравнительно-сопоставительной и прикладной лингвистики, прежде всего, в аспекте расширения системы знаний о лингвистических факторах, влияющих на качество автоматического извлечения эмоций.

Практическая значимость работы заключается в возможности использования полученных результатов для улучшения существующих и создания новых русскоязычных корпусов эмоциональных высказываний. Описанные закономерности влияния лингвистических признаков могут быть непосредственно применены при решении прикладных задач межъязыкового переноса для распознавания эмоций, прежде всего для русского языка, но также и для других целевых языков.

Личный вклад В. И. Лемаева состоит в самостоятельной постановке проблемы, выборе методов исследования, проведении аналитического обзора современной литературы по теме. Автором лично выполнены все практические эксперименты по обучению и тестированию речевых и текстовых моделей на русском и иностранных языках, включая выбор

наборов данных, настройку гиперпараметров, проведение дообучения и тестирования моделей, анализ результатов с привлечением языкового материала используемых текстовых наборов данных. Самостоятельно проведена оценка статистической значимости полученных результатов и их интерпретация.

К **достоинствам** выполненной работы следует отнести следующее:

1. Диссертант приводит логичное и последовательное изложение аналитических обзоров, охватывающих теоретические, методологические и прикладные аспекты автоматического распознавания эмоций. Диссертант использует релевантные современные источники, что позволяет ему систематизировать текущее положение дел в исследуемом вопросе, выявить существующие лакуны, и далее, с учетом этого, сформулировать исследовательские задачи и обосновать выбор методов для их решения. В частности:

а) диссертант предлагает довольно развернутую классификацию существующих мультимодальных наборов данных для распознавания эмоций (по степени естественности речи, по источнику текстов, по формальным и метапараметрам). Полагаем, что предложенная классификация (см. параграф 1.3) имеет самостоятельную ценность для научного сообщества, так как позволяет исследователям оперативно подбирать подходящих для определенных задач наборы данных с учетом жанра, языка, типа разметки и социально-демографических характеристик говорящих;

б) автор логично, вполне корректно и ясным языком представляет эволюцию подходов к распознаванию эмоций – от линейной регрессии и метода опорных векторов до современных трансформеров (см. параграф 1.5). Каждый метод сопровождается примерами его реального применения в задаче распознавания эмоций, что делает обзор специализированным, ориентированным на цели и предмет диссертационного исследования.

в) с опорой на литературные источники в главе 2 автор приходит к обоснованному выводу о неразработанности вопроса межъязыкового

переноса для русского языка как целевого, в частности, в контексте задачи распознавания эмоций. Это позволило диссертанту корректно поставить исследовательские задачи, определить дизайн эксперимента и сформировать методологическую базу, ориентированную на решение актуальных задач межъязыкового переноса на малоресурсные языки.

2. Заслуживает положительной оценки методическая продуманность экспериментальной работы: диссертант отдельно проводит серии экспериментов для речевых и для текстовых моделей. Для каждого типа моделей вначале оценивается качество их дообучения на русскоязычных данных, тем самым определяется baseline для последующей оценки качества с применением межъязыкового переноса. Затем реализуется оригинально построенный эксперимент по межъязыковому переносу, включающий предварительное обучение моделей на специализированной задаче распознавания эмоций с использованием данных различных языков, в том числе с варьированием режимов подготовки выборок данных для обучения. Такое построение позволяет не только оценить достижимый уровень качества для каждой модальности, но и, сопоставив полученные результаты, обоснованно выявить значимые лингвистические факторы, отличающие распознавание на базе речевых и текстовых данных.

3. Несомненным достоинством работы является попытка комплексного учета отдельных лингвистических факторов (синтаксической структуры высказывания, наличие лексико-грамматического маркера отрицания и графических элементов), потенциально влияющих на качество межъязыкового переноса. Диссертант предлагает и апробирует методику их анализа: по каждому из выделенных признаков последовательно осуществляется подготовка данных (разделение выборки по выбранному лингвистическому основанию), проводится дообучение, применяются апробированные статистические тесты для проверки значимости результатов, и на завершающем этапе автор, не ограничиваясь количественными данными, обращается к анализу конкретного языкового материала с целью

выявления лингвистических признаков, влияющих на качество. Такая структура анализа повышает достоверность и обоснованность выводов.

4. Следует отметить детальность и наглядность презентации экспериментальных данных. Все результаты дообучения и переноса моделей представлены в подробных таблицах и графиках, что позволяет верифицировать полученные результаты и проверить выводы диссертанта с опорой на фактические данные.

Вместе с тем при знакомстве с работой возникает ряд **замечаний и вопросов**, требующих дополнительного разъяснения:

1. В положении 1 (о качестве переноса речевых моделей) и положении 3 (о качестве переноса текстовых моделей) ключевой причиной высокого качества называется «близость языков». Далее в тексте работы автор указывает, что речевые модели показывают лучшее качество при переносе с польского языка, так как он «близкий» к русскому, при этом также делается вывод, что качество переноса речевых моделей определяется интонацией больше, чем лексикой или структурными свойствами языка. В текстовых моделях лучшее общее качество показывает модель, предварительно дообученная на испанском языке, а более высокое качество автор связывает с лексическим компонентом. В связи с этим, у нас есть предположение, что термин «близость» в положении 1 и положении 3 имеет различное содержание. Полагаем, что необходимо дать пояснения, в каком смысле используется термин «близость языков» в каждом из положений? Подразумевается ли генетическая близость языков? синтаксический тип языков? фактор «пересечения» лексики? сходство выражения эмоций через идентичные интонационные паттерны? или имеются в виду какие-то другие признаки, определяющие степень близости к русскому каждого из рассматриваемых языков?

2. В перечне методов исследования заявлен «лингвокультурологический метод (анализ культурных особенностей, свойственных отдельным языкам, и как они выражаются в лексике)» (с. 6).

Однако в тексте работы мы не нашли систематического применения этого метода: насколько можно судить, культурные особенности носителей языков практически не анализируются, а выводы о способах выражения эмоций строятся на распределении лексических маркеров, синтаксической организации высказываний и пунктуации без обращения к культурным контекстам. Просим диссертанта пояснить, как использовался данный метод.

3. Анализируя результаты дообучения моделей на русскоязычном речевом корпусе, автор заключает, что «размерность и количество параметров модели имеет значительное влияние на качество работы» (с. 60). Однако различия по метрике WA составляют лишь тысячные доли (например, для WavLM_large и WavLM_base соответственно 0,878 и 0,872 – см. табл. 1). Насколько правомерно говорить о значительном влиянии в данном случае, и не могут ли эти различия быть следствием статистической погрешности?

4. В эксперименте по влиянию порядка слов автор делает вывод о том, что модели придают больший вес финальной позиции глагола (с. 96). Однако в приведенных диссертантом примерах присутствуют и другие маркеры, которые могут сигнализировать о выражаемой эмоции (лексика, восклицательные знаки и т.д.). Поэтому полагаем, что потенциально не только финальная позиция глагола в предложении могла повлиять на решение моделей. Автор не приводит количественных данных о регулярности этого наблюдения. В связи с этим считаем, что вывод о влиянии на качество переноса нахождения глагольной единицы в финальной позиции требует дополнительного обоснования.

5. В таблице 2 (с. 62) меньшая по размеру модель rubert-tiny2 демонстрирует на некоторых эмоциях и наборах данных незначительно более высокие результаты, чем более крупная RuBERT. Представляется, что это противоречит тезису о том, что большее количество параметров позволяет улучшать качество модели (с. 49). Как автор объясняет этот феномен?

6. В работе отсутствует прямое сопоставление результатов межъязыкового переноса (представленные в главе 4) с результатами обучения моделей на русском языке (показанные в главе 3). Для текстовых данных языковые модели, обученные непосредственно на русских корпусах, демонстрируют качество классификации 0,81–0,86 (табл. 2), в то время как при обучении на том же наборе данных с межъязыковым переносом с любого из языков-источников наилучшее общее качество не превышает 0,63 (табл. 8), а для отдельных эмоций падает до нуля. Это ставит вопрос: не свидетельствует ли столь значительное падение качества о том, что предварительное дообучение на целевую задачу распознавания эмоций на других языках не помогает, а скорее ухудшает результаты по сравнению с простым обучением на пусть небольшом, но сразу русскоязычном корпусе текстов? Хотелось бы услышать пояснение автора, как он оценивает целесообразность применения переноса для русского языка в текстовой модальности с учетом полученных результатов экспериментов.

7. Утверждение автора о существовании «принятого» набора из 6 базовых «архетипичных» эмоций представляется слишком категоричным (с. 12). В психологии существуют разные подходы к определению перечня базовых эмоций и их количества (К. Изард – 10, П. Экман – 6, Р. Плутчик – 8 и т.д.). Также, при подготовке наборов данных для прикладных задач разметка проводится с опорой на различные инвентари базовых эмоций (например, в указанный автором набор данных SSEC, используемый в SemEval-2016 Task 4, включена разметка по 8 эмоциям, а не по 6 категориям). Кроме того, описание пространственной модели эмоций как модели, содержащей всего две размерности, представляется нам неполным: в психологии пространственная модель эмоций А. Рассела включает три размерности: валентность – интенсивность – доминирование, и эта концепция также в полном объеме используется при разметке наборов данных и составлении лексиконов.

8. Автор в ряде случаев использует терминологию и формулировки, которые представляются нам не вполне точными. Например,

а) вводятся термины «эмоциональные слова» (с. 9) – видимо, имеется в виду лексика с эмоционально-оценочным компонентом значения, «эксплицитная лексика» (с. 85) – видимо, предполагается эксплицитное выражение эмоций посредством использования слов с эмоционально-оценочным компонентом;

б) при обсуждении классификации эмоций автор указывает, что «выделяют около 300 различных эмоций, причем среди них встречается и множество схожих по значению» (с. 12). Строго говоря, значением обладают не эмоции, а лексические единицы, их номинирующие;

в) в автореферате, при пояснении примера *«никаких постоянных увлечений нет, совсем.»* автор указывает, что здесь «грусть выражается имплицитно через контекст» (с. 21 автореферата). Полагаем, что речь идет не о контексте (экстралингвистическом или лингвистическом), с учетом которого следует трактовать высказывание, а об описании положения дел, которое ассоциативно связано с эмоцией грусти.

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования. Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В.Ломоносова к работам подобного рода. Содержание диссертации соответствует специальности 5.9.8 Теоретическая, прикладная и сравнительно-сопоставительная лингвистика (по филологическим наукам), а также критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В.Ломоносова. Диссертационное исследование оформлено согласно требованиям Положения о совете по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук Московского государственного университета имени М.В.Ломоносова.

Таким образом, соискатель Лемаев Владислав Игоревич заслуживает присуждения ученой степени кандидата филологических наук по специальности 5.9.8 Теоретическая, прикладная и сравнительно-сопоставительная лингвистика.

Официальный оппонент:

кандидат филологических наук,
заведующий кафедрой «Лингвистика и перевод»
института лингвистики и международных коммуникаций
ФГАОУ ВО «Южно-Уральский государственный университет
(национальный исследовательский университет)»

БАБИНА Ольга Ивановна

08.07.2026 г.

Контактные данные:

тел.: +7(351)267-93-94, e-mail: babinaoi@susu.ru
Специальность, по которой официальным оппонентом
защищена диссертация:
10.02.21 – Прикладная и математическая лингвистика

Адрес места работы:

454080, г. Челябинск, ул. проспект Ленина, д. 76,
ФГАОУ ВО «Южно-Уральский государственный университет (национальный
исследовательский университет)», институт лингвистики и международных
коммуникаций, кафедра «Лингвистика и перевод»
Тел.: +7(351)267-99-00; e-mail: info@susu.ru

Подпись сотрудника ФГАОУ ВО «Южно-Уральский государственный университет
(национальный исследовательский университет)» О.И. Бабиной удостоверяю:

Начальник службы делопроизводства

Н.Е. Циулина