

УДК 81'33

DOI: 10.37482/2687-1505-V206

КОЛМОГорова Анастасия Владимировна, доктор филологических наук, профессор, профессор департамента филологии Санкт-Петербургской школы гуманитарных наук и искусств Национального исследовательского университета «Высшая школа экономики». Автор 190 научных публикаций, в т. ч. 19 монографий, 7 учебных пособий*
ORCID: <https://orcid.org/0000-0002-6425-2050>

МАЛИКОВА Алина Вячеславовна, старший преподаватель кафедры романских языков и прикладной лингвистики Института филологии и языковой коммуникации Сибирского федерального университета. Автор 27 научных публикаций, в т. ч. двух монографий и одного учебного пособия**
ORCID: <https://orcid.org/0000-0002-3438-1839>

МОДЕЛИРОВАНИЕ СПЕЦИФИКИ ЭМОЦИОНАЛЬНОГО ВОСПРИЯТИЯ ИНТЕРНЕТ-ТЕКСТОВ НА РУССКОМ ЯЗЫКЕ ТУВИНСКО-РУССКИМИ БИЛИНГВАМИ В СОПОСТАВЛЕНИИ С РУССКИМИ МОНОЛИНГВАМИ: ОПЫТ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ¹

Данное исследование лежит в области применения технологии эмоционального анализа текстов. Целью исследования является статистическое моделирование специфики эмоционального восприятия интернет-текстов на русском языке тувинско-русскими билингвами. Проверяется гипотеза о значимости фактора организации ментального лексикона для конфигурации эмоциональной разметки текстов, производимой данной группой аннотаторов. Предполагается, что, ввиду сложности психолингвистической реальности билингвального языкового сознания тувинцев, при разметке будет проявляться специфика обработки эмоциональной информации, выраженной вербально на неродном, но коммуникативно значимом языке. Материалом исследования выступают 48 текстов из датасета постов из социальной сети «ВКонтакте» – по 6 текстов на каждый из 8 эмоциональных классов, выделяемых согласно модели эмоций «Куб Лёвхейма»: грусть, радость, отвращение, удивление, воодушевление, гнев, стыд, страх. Применяется разработанный и апробированный исследовательской группой компьютерный интерфейс для недискретной эмоциональной разметки текстов. В психолингвистическом эксперименте приняли участие 174 русских монолингва и 65 тувинско-русских билингвов. Перед тем как перейти к аннотированию, информанты-билингвы

*Адрес: 190068, Санкт-Петербург, наб. канала Грибоедова, 123, e-mail: nastiakol@mail.ru

**Адрес: 660041, г. Красноярск, просп. Свободный, д. 82, стр. 1; e-mail: malikovaav1304@gmail.com

¹Работа выполнена при финансовой поддержке Краевого фонда науки Красноярского края, проект № 629 «Компьютерное моделирование эмоционального восприятия интернет-текстов на русском языке тувинско-русскими билингвами»

Для цитирования: Колмогорова А.В., Маликова А.В. Моделирование специфики эмоционального восприятия интернет-текстов на русском языке тувинско-русскими билингвами в сопоставлении с русскими монолингвами: опыт экспериментального исследования // Вестн. Сев. (Арктич.) федер. ун-та. Сер.: Гуманит. и соц. науки. 2022. Т. 22, № 5. С. 39–50. DOI: 10.37482/2687-1505-V206

ответили на вопросы опросника Bilingual Language Profile для получения данных об их доминирующем языке. При анализе результатов разметки использовались t -критерий Стьюдента и сравнительно-сопоставительный метод. Обнаружена статистическая значимость отличий значений эмоциональных оценок, приписываемых текстам на русском языке русскими монолингвами и тувинско-русскими билингвами (возраст большинства информантов составляет 18–25 лет). Согласно результатам анализа, глубина этих отличий зависит от доминирующего языка билингва. Отмечается несовпадение вторичных эмоций при совпадении ведущей эмоции текстов при оценке данными группами, а также определенная нестабильность оценок информантов-билингвов при повторном оценивании. Результаты исследования вносят вклад в разработку проблемы восприятия эмоций в тексте билингвами.

Ключевые слова: *аффективные вычисления, билингвизм, доминантный язык, интернет-текст, психолингвистика, тувинский язык, эмоциональный анализ.*

Введение

В последнее десятилетие в междисциплинарном поле компьютерной лингвистики, психолингвистики и психологии активно развивается достаточно новое направление, ставшее известным как «парадигма аффективных вычислений». Оно объединяет исследования, сосредоточенные на процессах создания компьютерных систем, способных в процессе взаимодействия «чувствовать» и узнавать эмоции, отвечать на них, а также влиять на эмоциональное состояние других когнитивных агентов [1]. В качестве данных, на которых может строиться эмоциональное предсказание искусственного интеллекта, используются записи голоса, мимических движений человека, информация о его физиологических показателях (температура, пульс, давление), фиксация движений глаз. Но одним из самых важных типов данных является речевая/текстовая продукция человека, находящегося в том или ином эмоциональном состоянии.

На ее изучении строится технология эмоционального анализа текстов [2], результатом применения которой становится либо атрибуция текста, подаваемого на вход компьютерной системы, к одному из эмоциональных классов (дискретный эмоциональный анализ) (см., например, [3]), либо указание на его локализацию в многомерном эмоциональном пространстве (многомерный эмоциональный анализ) [4], например посредством приписывания «эмоциональных координат».

Подобные результаты становятся возможными благодаря тренировке компьютерных моделей на размеченных данных, в которых каждому объекту анализа (тексту) уже приписан его атрибут (эмоциональные координаты, эмоциональный класс). Качество такого датасета чрезвычайно важно для достижения высоких показателей точности и полноты классификации текстов программой.

Однако оно чрезвычайно чувствительно ко множеству психолингвистических факторов, связанных со спецификой организации и функционирования речевого механизма аннотаторов (разметчиков), а именно: психотипу аннотатора, его способности интерпретировать эмоциональную составляющую текста; индивидуальному уровню эмпатии [5]; стратегиям интерпретации эмоции, к которым он прибегает с определенной регулярностью, поскольку они уже стали привычными паттернами поведения; специфике организации его ментального лексикона и степени владения языком и лингвокультурой в целом, например, чувствительность к культурным отсылкам и имплицитным культурным смыслам.

Проведя в рамках предыдущего проекта эмоциональное аннотирование с привлечением группы информантов – русских монолингвов [6], мы имели возможность наблюдать влияние всех перечисленных факторов, кроме последнего.

В связи с этим появилась исследовательская гипотеза о том, что фактор родного/неродного, доминантного/недоминантного языка будет так-

же влиять на общую конфигурацию эмоциональной разметки датасета. Следовательно, компьютерный анализатор эмоций в тексте на русском языке, обученный на датасете, размеченном русскими монолингвами, будет давать не вполне релевантные предсказания для восприятия билингов, чьим родным языком он не является.

Чтобы проверить данную гипотезу, мы провели сравнение результатов эмоционального аннотирования, полученных в двух группах аннотаторов: у русских монолингвов и тувинско-русских билингвов.

Актуальность исследования обусловлена тем, что информационное пространство социальной сети «ВКонтакте» стало на сегодняшний момент основной платформой социального взаимодействия в русскоязычном, но, одновременно, многонациональном и мультиэтническом российском социуме. На данной площадке выстраиваются как личные, так и внутригрупповые, межконфессиональные и институциональные отношения, например государства и граждан. Поскольку большинство сетевых текстов, в том числе предназначенных для двуязычного населения Республики Тыва, пишутся и публикуются на русском языке, важным становится предсказание эмоционального эффекта, производимого такими текстами на читателей-билингвов. От успешности подобного эмоционального прогнозирования зависит эффективность сетевой коммуникации, на которой зиждется стабильность в многонациональном российском обществе.

Новизна работы состоит в том, что экспериментальное выявление и моделирование специфики интерпретации эмоции в тексте на русском языке тувинско-русскими билингвами проводится впервые. Кроме того, на данном материале применяется разработанный исследовательской группой специальный компьютерный интерфейс для недискретной эмоциональной разметки текстов, позволяющий получать числовые данные, которые затем могут быть использованы для моделирования и визуализации.

Цель данной публикации – представить результаты статистического анализа отличий эмо-

циональных значений, приписываемых одним и тем же текстам на русском языке русскими монолингвами и тувинско-русскими билингвами.

Специфика билингвального языкового сознания: тувинско-русский билингвизм

Билингвальное языковое сознание в норме и при патологии является одним из наиболее востребованных объектов изучения в современной психолингвистике. Рассматриваются следующие вопросы:

- 1) специфика протекания процессов памяти [7];
- 2) особенности становления фонематического слуха [8];
- 3) проблема восстановления речевой деятельности после инсульта [9];
- 4) отличительные черты когнитивной обработки эмоциональных слов.

Последняя из упомянутых тематик находится в фокусе нашего интереса.

Существует несколько основных теорий, объясняющих то, как связаны эмоции и слова в билингвальном ментальном лексиконе. Так, согласно RHM-концепции (Revised Hierarchical Model [10]), ментальный лексикон для каждого языка имеет лексический и концептуальный слой. Для первого языка (L1) связь между смыслом (концептом) и его лексическим репрезентантом крепче, чем для второго (L2), в силу ее раннего формирования. Для L2 ее слабость компенсируется дополнительной связью между словом на L2 и его смыслом в ментальном лексиконе L1. Таким образом, когда речь идет об эмоциях, активизируемых словами в билингвальном языковом сознании, то, скорее всего, они «оживляют» одновременно связи и в ментальном лексиконе L1 и L2 [11]. В рамках концепции «контекста изучения языка» [12] исследователи констатируют, что слова L1 в большей степени способны активировать эмоции, чем слова L2, поскольку их усвоение происходило в заведомо более эмоциональной атмосфере – в общении с родственниками, родителями.

Экспериментальные исследования, выполняемые с использованием заданий на лексическое решение [13] и процедур измерения активности

мозга при помощи электроэнцефалографии [14] и функциональной магнитно-резонансной томографии [15], подтверждают тезис о влиянии статуса языка в психолингвистической реальности билингва (доминантный или нет, первый или нет) на характер когнитивной обработки эмоций и эмоциональных слов.

В исследовании мы опираемся на результаты, полученные на материале экспериментов с группой тувинско-русских билингов, поэтому кратко опишем языковую ситуацию в Республике Тува.

Современная этноязыковая ситуация в республике характеризуется как экзогlossная и несбалансированная: несмотря на то, что тувинский язык также является государственным, он уступает в коммуникативной мощности русскому и находится по отношению к нему «в положении функциональной дополнительности» [16, с. 194]. Как отмечает Т.Г. Боргоякова [16, с. 194], тувинский язык отличается от других тюркских языков (хакасского, алтайского и др.) высокой витальностью, что объясняется более поздним вхождением в состав СССР и, как следствие, тем, что Тува в наибольшей степени сохранила свою культурную и этническую самобытность, этнический язык.

Результаты ассоциативных экспериментов демонстрируют, что языковое сознание тувинских билингов строится преимущественно на основе тувинского языка, что указывает на его доминирующую роль в билингвальной психолингвистической реальности тувинцев [17, с. 21]. Тем не менее его коммуникативная мощностъ не равноценна в разных сферах социальной жизни: тувинцы проявляют наибольшую компетентность во владении устными формами тувинского языка и письменными формами русского языка [17, с. 21], который занимает лидирующие позиции в сфере управления и делопроизводства.

Одновременно исследователи признают, что «появились признаки ассимиляции тувинского языка и реального быстрого исчезновения в условиях глобализации» [18, с. 70].

Иными словами, психолингвистическая реальность билингвального языкового сознания

тувинцев представляет собой сложный и комплексный феномен, изучение которого посредством экспериментальной работы позволит получить новые сведения о нем. В частности, это касается специфики обработки эмоциональной информации, выраженной вербально на неродном, но коммуникативно достаточно сильном языке, т. е. в данном случае – русском.

Дизайн эксперимента, материалы и методы

В рамках проекта по созданию компьютерного классификатора текстов по критерию доминирующей в них эмоции нами был сформирован аннотированный датасет постов из социальной сети «ВКонтакте» в объеме 3920 текстов из трех пабликов: «Подслушано», «Палата № 6» и «Карамель». Данные публичные группы были выбраны в силу того, что публикуемые в них посты имеют небольшой и достаточно стандартный объем в 80–100 слов и отличаются яркой и разнообразной эмоциональной палитрой.

Чтобы создать подобную выборку, нам нужно было разработать процедуру эмоционального аннотирования, в ходе которой информантам предъявляется текст и предлагается оценить эмоцию/эмоции, присутствующие в нем. Несмотря на разнообразие существующих классификаций эмоций, мы остановились на многомерной модели, известной как «Куб Лёвхейма» [19]. Трехмерное пространство Куба представляет собой эмоциональный континуум, а его 8 вершин – максимальные точки проявления базовых эмоций: грусть, радость, отвращение, удивление, воодушевление, гнев, стыд, страх (*рис. 1*).

Для удобства аннотаторов нами был разработан интерфейс для недискретной разметки, где сначала на экране информанта появляется инструкция, затем – текст (пост из «ВКонтакте» объемом 80–100 слов), далее 4 шкалы для оценивания. Они представляют собой внутренние диагонали Куба: «Стыд – Воодушевление», «Отвращение – Гнев», «Страх – Удивление», «Радость – Грусть». Середина шкалы – ноль эмоций, чем ближе к тому или иному полюсу информант передвигает ползунок на шкале, тем сильнее выражена та эмоция, которая является крайней точкой шкалы (*рис. 2*).

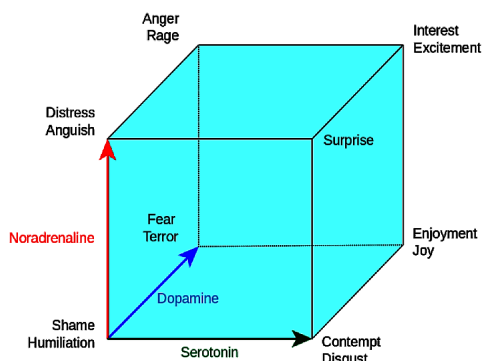


Рис. 1. Визуализация модели эмоций «Куб Лёвхейма» [19]
Fig. 1. Visualization of the Lövheim cube of emotion [19]

Дедушка пришел с фронта старшиной, весь в осколках — противотанковая артиллерия, наводчик орудия. В семнадцать лет убежал добровольцем, приписав себе год. Как-то сидим с ним, разговариваем, и я спросил его, сколько танков уничтожил, он ответил, что пять. А я в ответ: «А чего так мало?! Плохого кина нагляделся. Дедушка стал как бы оправдываться, что, мол, тяжело танк вывести из строя, он ведь тоже стреляет в ответ... Ну, и как-то тема сошла на нет. Было мне лет восемь или девять тогда. Сейчас мне 34 и за этот разговор мне больше всего в жизни стыдно. Зря тогда не извинился, деда уже нет 18 лет, вот теперь живу с этим.

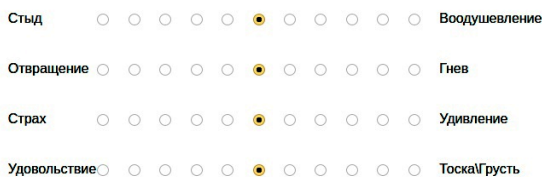


Рис. 2. Пользовательский экран интерфейса для не-дискретной эмоциональной разметки текстов
Fig. 2. User interface screen for non-discrete emotional markup of texts

Если аннотатор расставляет ползунки на всех шкалах, мы получаем числовые значения оценки на каждой шкале для каждого текста. Путем применения нескольких операций математического моделирования (4 значения превращаются в трехмерный усредняющий вектор, который проецируется на три оси Куба) мы можем конвертировать эти значения в эмоциональные координаты текста в пространстве Куба.

Тексты для аннотирования сначала были отобраны по тематическим эмоциогенным хештегам, а затем оценены с участием 2000 носителей русского языка, монолингвов на платформе

«Яндекс. Толока». Для проведения обсуждаемого в данной публикации эксперимента с билингвами мы использовали 48 текстов из данной выборки (по 6 на каждый из 8 эмоциональных классов), а в качестве аннотаторов выступили: а) 174 русских монолингва (ср. возраст = 21,1); б) 65 тувинско-русских билингвов (ср. возраст = 22,2). Но, поскольку не все информанты выполнили оценку до конца, в дальнейшем анализе мы будем опираться на данные от 46 информантов, оценивших все 48 текстов. В каждой группе более 70 % составляли студенты, обучающиеся в Сибирском федеральном университете. Прежде чем приступить к эмоциональному аннотированию респондентам было предложено пройти опросник, сформированный на базе опросника Bilingual Language Profile [20], для того, чтобы мы могли получить данные о доминирующем языке билингва. Обе группы информантов были сформированы из волонтеров посредством публикации соответствующего объявления в публичных группах в социальных сетях. Анкетирование и разметка проводились онлайн.

Результаты аннотирования по 4 шкалам анализировались нами при помощи статистической метрики, известной как *t*-критерий Стьюдента. Для качественного анализа использовался сравнительно-сопоставительный метод.

Исходная гипотеза эксперимента сформулирована следующим образом: если между эмоциональной интерпретацией текстов на русском языке в двух группах информантов (монолингвов и тувинско-русских билингвов) есть расхождение, то оценки 48 текстов на 4 шкалах в них будут статистически значимо отличаться.

Результаты и обсуждение

Статистическое сравнение полученных в двух группах информантов «сырых» данных, т. е. значений оценок по 4 шкалам предложенного информантам интерфейса, с применением *t*-критерия Стьюдента показало, что по двум шкалам («Отвращение – Гнев» и «Страх – Удивление») во всей выборке данных наблюдается значимое расхождение: при $p < 0.05$ по шкале «Отвращение – Гнев» получено значение

0,000017094950741, а по шкале «Страх – Удивление» – 0,000003958679177 (табл. 1). Оба – в зоне значимости.

Иными словами, можно предположить, что интерпретация данных 4 эмоций в текстах имеет у монолингвов и билингвов свою специфику. Причем качественный анализ показывает, что тувинско-русские билингвы больше акцентируют внимание на эмоциях отвращения и удивления, нежели это делают монолингвы.

Кроме общего анализа значимости отличий между оценками эмоций, данными монолингвами и билингвами по всем текстам, мы также применили *t*-критерий Стьюдента для установ-

ления значимости отличий в оценках в группах текстов, где ведущей является та или иная эмоция. Подчеркнем, что 73 % текстов в обеих выборках попали в одну и ту же категорию по критерию ведущей эмоции. Так, в группах текстов, где доминирующей эмоцией стали гнев или удивление, значимые отличия в оценках монолингвов и билингвов были обнаружены на шкалах «Стыд – Воодушевление» и «Страх – Удивление»; в подгруппе грустных текстов – «Стыд – Воодушевление» и «Радость – Грусть»; в отвратительных – «Стыд – Воодушевление» и «Отвращение – Гнев»; в радостных и воодушевленных – «Стыд – Воодушевление»;

Таблица 1

***p*-ЗНАЧЕНИЯ (при $p < 0,05$), ПОЛУЧЕННЫЕ ПРИ СРАВНЕНИИ ДАННЫХ ОЦЕНКИ ЭМОЦИИ ТЕКСТОВ БИЛИНГВАМИ И МОНОЛИНГВАМИ С ПОМОЩЬЮ *t*-КРИТЕРИЯ СТЬЮДЕНТА**

***p*-VALUES (at $p < 0.05$) OBTAINED BY COMPARING THE EVALUATION DATA OF TEXTS' EMOTIONS BY BILINGUALS AND MONOLINGUALS USING STUDENT'S *t*-TEST**

№	Стыд – Воодушевление	Отвращение – Гнев	Страх – Удивление	Радость – Грусть
	Вся выборка			
1	0,168035836418516	0,000017094950741	0,000003958679177	0,650112519548620
	Гневные тексты			
2	0,024581956452828	0,143814898804548	0,000036464976001	0,721777101947431
	Грустные тексты			
3	0,018088925987058	0,834818530646044	0,155642855229764	0,000014442944112
	Отвратительные тексты			
4	0,044788024244891	0,000000000000001	0,072912310348732	0,965666584906091
	Радостные тексты			
5	0,023280080837632	0,083019252225977	0,158893920686528	0,293755614940468
	Воодушевленные тексты			
6	0,000000339789053	0,307239340153000	0,179120164999974	0,117777342739973
	Страшные тексты			
7	0,089459445498090	0,665581657253674	0,074643383944760	0,034271317708824
	Стыдные тексты			
8	0,835649518894447	0,351717329435228	0,918886773061636	0,834782760469545
	Удивительные тексты			
9	0,000078409080575	0,697143996460154	0,034617158909735	0,066264496473109

Примечание. Полужирным выделены статистически значимые показатели.

в страшных – «Радость – Грусть». В группе стыдных текстов значимых отличий в эмоциональной интерпретации текстов на русском языке не выявлено.

Таким образом, мы можем констатировать, что, в зависимости от ведущей эмоции, статистически значимые отличия между монолингвами и билингвами могут наблюдаться на всех 4 шкалах оценки. Причем интересно, что, если ответы всех информантов согласуются относительно выбора ведущей эмоции, различия в оценках второстепенных эмоций существенны. Например, следующий текст был оценен всеми информантами как гневный:

Муж работает в крупной продуктовой компании. Достали родственники, которые постоянно удивляются тому, что он не тащит домой продукты. Да, блин, мы зарабатываем прилично и можем себе позволить купить все, что надо. Какой смысл из-за пакета молока лишиться достойной работы и испортить репутацию.

На рис. 3 и 4 хорошо видно, что, несмотря на близкие и высокие значения гнева на шкале «Отвращение – Гнев» (disgust_rage), вторичные эмоции распределены по-разному: у монолингвов эмоция стыда выражена слабее, чем у билингвов (шкала shame_excitement); ни страх, ни удивление не выражены вовсе, а у билингвов хорошо проявляется эмоция удивления (диапазон от 0 до 2 по шкале fear_surprise).

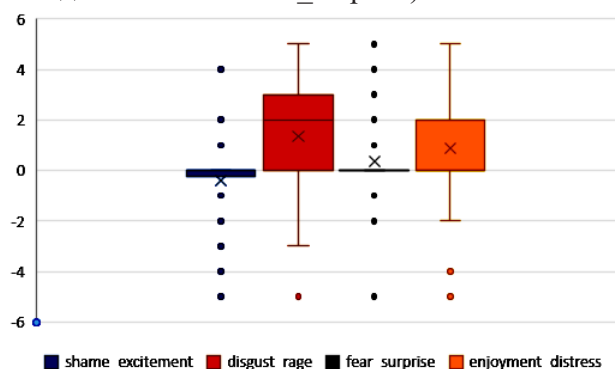


Рис. 3. Диаграмма размаха значений оценки текста № 1 русскими монолингвами

Fig. 3. Box plot of evaluations of text no. 1 by Russian monolinguals

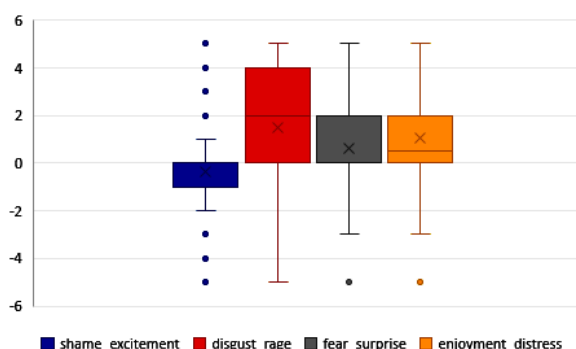


Рис. 4. Диаграмма размаха значений оценки текста № 1 тувинско-русскими билингвами

Fig. 4. Box plot of evaluations of text no. 1 by Tuvan-Russian bilinguals

Кроме того, при работе с данными, полученными только от билингвов, мы также увидели качественные и количественные отличия в зависимости от доминирующего языка билингва.

Чтобы определить доминирующий язык, мы воспользовались системой анализа, предложенной авторами опросника по билингвизму. Вопросам каждого из 4 модулей опросника присвоен различный максимальный балл, а каждому варианту ответа – своя ценность по шкале от 0 до 20 (модуль «Языковая история»), от 0 до 10 (модуль «Использование языка»), от 0 до 6 (модули «Языковая компетентность» и «Отношение к языкам»). Сумма ответов на все вопросы модуля о тувинском и русском языках в отдельности, помноженная на коэффициент данного модуля, дает общий балл за данный модуль. Разница между суммой баллов для тувинского и русского языков показывает индекс превалирования одного из языков в сознании билингва. На рис. 5 представлены итоговые значения оценки доминирующего языка, полученные путем суммирования баллов за все 4 модуля. В нашем случае разница со знаком «+» говорит о доминировании тувинского языка (Т), тогда как отрицательная – о доминировании русского языка (R). Мы также выделили паритетную (сбалансированную, В) группу респондентов, в которую попали значения от -7 до 9 (рис. 5).

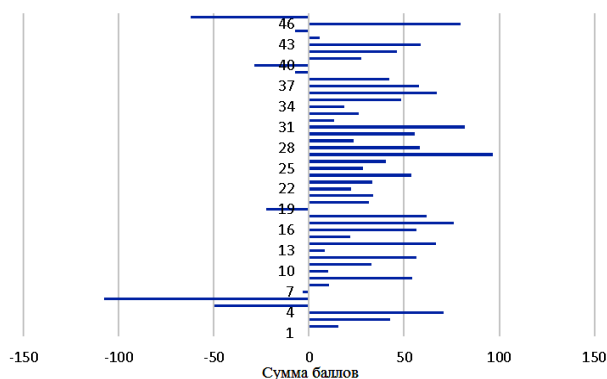


Рис. 5. Распределение 46 информантов-билингвов по параметру доминирующего языка

Fig. 5. Distribution of 46 bilingual informants by dominant language

По сумме баллов, полученных согласно опроснику по билингвизму, у 36 билингвов доминирующим языком является тувинский, у 5 – русский, 5 характеризуются сбалансированным билингвизмом.

При анализе по шкалам обнаружилось статистически значимые (при $p < 0.05$) отличия в оценках у билингвов с доминирующим русским языком, билингвов с доминирующим тувинским и сбалансированных билингвов (табл. 2).

Отсюда можно заключить, что доминирующий язык является одним из решающих факторов, которые влияют на результаты эмоциональной оценки, осуществляемой билингвами.

При сравнении с монолингвами большие статистические отличия проявляются у билингвов с доминирующим русским языком, меньшие – с доминирующим тувинским, еще менее выраженные – у информантов со сбалансированным билингвизмом. Однако во всех группах они статистически значимы.

Наибольшая «непохожесть» эмоциональной интерпретации монолингвов и билингвов с доминирующим русским языком может быть связана отчасти с небольшим количеством информантов среди вторых. Однако высокая значимость отличий может также свидетельствовать о том, что, даже в случае ведущей роли неродного языка в коммуникативном пространстве билингва, реальность его языкового сознания может быть иной, нежели чем у монолингва.

Интересен еще тот факт, что в тех случаях, когда информантам-билингвам приходилось несколько раз оценивать один и тот же текст из-за технических сбоев, связанных с неустойчивой интернет-сетью, они по-разному, иногда кардинально, оценивали текст в первый и во второй раз. Отметим, что технические сбои случались и при аннотировании текстов монолингвами,

Таблица 2

***p*-ЗНАЧЕНИЯ (при $p < 0,05$), ПОЛУЧЕННЫЕ ПРИ СРАВНЕНИИ ДАННЫХ ОЦЕНКИ ЭМОЦИИ ТЕКСТОВ БИЛИНГВАМИ ИЗ РАЗНЫХ ГРУПП ПО ДОМИНАНТНОМУ ЯЗЫКУ С ПОМОЩЬЮ *t*-КРИТЕРИЯ СТЬЮДЕНТА**

***p*-VALUES (at $p < 0.05$) OBTAINED BY COMPARING THE EVALUATION DATA OF TEXTS' EMOTIONS BY BILINGUALS FROM DIFFERENT DOMINANT LANGUAGE GROUPS USING STUDENT'S *t*-TEST**

Сравниваемые группы	Значение <i>p</i>
Т/Р-билингвы	0,013304970
Р/В-билингвы	0,003715299
Т/В-билингвы	0,032157913
Монолингвы/Т-билингвы	0,000947464
Монолингвы/В-билингвы	0,003171218
Монолингвы/Р-билингвы	0,000291929

Примечание. Полужирным выделены статистически релевантные значения.

однако последние при повторном оценивании практически точь-в-точь повторяли свой первый ответ.

Такая нестабильность эмоционального восприятия текстов у билингвов, очевидно, связана с тем, что в языковом сознании активируются одновременно эмоциональные составляющие русских и тувинских слов. От того, какая нейронная связь будет активирована сильнее в данный конкретный момент, зависит ответ информанта.

Заключение

Проведенная экспериментальная работа позволяет сделать несколько предварительных выводов, которые, тем не менее, нуждаются в дополнительной проверке в дальнейшем.

Во-первых, следует констатировать, что наша гипотеза подтвердилась: между числовыми оценками эмоциональности текстов, формулируемыми билингвами и монолингвами, есть статистически значимые отличия, а значит, есть и качественные отличия в интерпретации ими эмоциональности в тексте.

Во-вторых, подчеркнем, что при общем совпадении мнений относительно ведущей эмоции, вербализованной в тексте, информанты-монолингвы и информанты-билингвы выбирали различные вторичные эмоции.

В-третьих, неодинакова глубина отличий между монолингвами и билингвами: она зави-

сит от доминирующего языка билингва. Наибольшие отличия показывает, как ни парадоксально, сравнение монолингвов и билингвов с доминирующим русским языком.

Наконец, важно, что билингвы демонстрируют некоторую волатильность или нестабильность оценок при первом и повторном оценивании одного и того же текста.

Таким образом, представляется, что подобного рода экспериментальная работа вносит определенный вклад в разработку проблемы восприятия и интерпретации эмоции в тексте билингвами, поскольку предложены новая методология и инструментарий для дальнейшего анализа.

Перспективы работы имеют технологический и социальный векторы: во-первых, ее результаты могут стать основой для разработки компьютерного приложения для генерации эмоциональных текстов, предназначенных именно для билингвов; во-вторых, наши наблюдения за «эмоциональным скольжением» в восприятии текстов на русском языке станут основой для создания методических рекомендаций для специалистов по связям с общественностью, журналистов, SMM-менеджеров по написанию экологичных в отношении эмоционального воздействия текстов, адресатами которых являются билингвы.

Список литературы

1. *Daily S.B., James M.T., Cherry D., Porter J.J. III, Darnell S.S., Isaac J., Roy T.* Affective Computing: Historical Foundations, Current Applications, and Future Trends // *Emotions and Affect in Human Factors and Human-Computer Interaction* / ed. by M. Jeon. Houghton: Academic Press, 2017. P. 213–231. DOI: [10.1016/B978-0-12-801851-4.00009-4](https://doi.org/10.1016/B978-0-12-801851-4.00009-4)
2. *Alm C.O., Rot D., Sproat R.* Emotions from Text: Machine Learning for Text-Based Emotion Prediction // *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing* / ed. by R. Mooney, C. Brew, L.-F. Chien, K. Kirchhoff. Vancouver: Association for Computational Linguistics, 2005. P. 579–586. URL: <https://aclanthology.org/H05-1073.pdf> (дата обращения: 26.04.2022).
3. *Volkova E.P., Mohler B.J., Meurers W.D., Gerdemann D., Bülhoff H.H.* Emotional Perception of Fairy Tales: Achieving Agreement in Emotion Annotation of Text // *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* / ed. by D. Inkpen, C. Strapparava. Los Angeles: Association for Computational Linguistics, 2010. P. 98–106. URL: <https://aclanthology.org/W10-0212.pdf> (дата обращения: 26.04.2022).

4. Gunes H., Pantic M. Automatic, Dimensional and Continuous Emotion Recognition // *Int. J. Synth. Emot.* 2010. Vol. 1, № 1. P. 68–99. DOI: [10.4018/jse.2010101605](https://doi.org/10.4018/jse.2010101605)
5. Kolmogorova A., Kalinin A., Malikova A. Semiotic Function of Empathy in Text Emotion Assessment // *Biosemiotics.* 2021. Vol. 14, № 2. P. 329–344. DOI: [10.1007/s12304-021-09434-y](https://doi.org/10.1007/s12304-021-09434-y)
6. Kolmogorova A.V., Kalinin A.A., Malikova A.V. Non-Discrete Sentiment Dataset Annotation: Case Study for Lövheim Cube Emotional Model // *Digital Transformation and Global Society. DTGS 2020. Communications in Computer and Information Science* / ed. by D.A. Alexandrov, A.V. Boukhanovsky, A.V. Chugunov, Y. Kabanov, O. Koltsova, I. Musabirov. Cham, 2020. Vol. 1242. P. 154–164. DOI: [10.1007/978-3-030-65218-0_12](https://doi.org/10.1007/978-3-030-65218-0_12)
7. Kazanas S.A., Wilck A.M., Altarriba J. Adaptive Memory: Greater Memory Advantages in Bilinguals' First Language // *Int. J. Biling.* 2022. Vol. 26, № 1. P. 49–64. DOI: [10.1177/13670069211022856](https://doi.org/10.1177/13670069211022856)
8. Pan L., Ke H., Styles S.J. Early Linguistic Experience Shapes Bilingual Adults' Hearing for Phonemes in Both Languages // *Sci. Rep.* 2022. Vol. 12. Art. № 4703. DOI: [10.1038/s41598-022-08557-7](https://doi.org/10.1038/s41598-022-08557-7)
9. Mechelli A., Crinion J.T., Noppeney U., O'Doherty J., Ashburner J., Frackowiak R.S., Price C.J. Structural Plasticity in the Bilingual Brain: Proficiency in a Second Language and Age at Acquisition Affect Grey-Matter Density // *Nature.* 2004. Vol. 431. P. 757. DOI: [10.1038/431757a](https://doi.org/10.1038/431757a)
10. Kroll J.F., Stewart E. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations // *J. Mem. Lang.* 1994. Vol. 33, № 2. P. 149–174. DOI: [10.1006/jmla.1994.1008](https://doi.org/10.1006/jmla.1994.1008)
11. Eilola T.M., Havelka J., Sharma D. Emotional Activation in the First and Second Language // *Cogn. Emot.* 2007. Vol. 21, № 5. P. 1064–1076. DOI: [10.1080/02699930601054109](https://doi.org/10.1080/02699930601054109)
12. Harris C.L., Gleason J.B., Ayçiçeği A. When Is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers // *Bilingual Minds: Emotional Experience, Expression, and Representation* / ed. by Pavlenko A. Clevedon: Multilingual Matters, 2006. P. 257–283. URL: <https://www.bu.edu/psych/charris/papers/HarrisEmotionBiling.pdf> (дата обращения: 26.04.2022).
13. Kazanas S.A., Altarriba J. Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals // *J. Psycholinguist. Res.* 2016. Vol. 45, № 2. P. 395–406. DOI: [10.1007/s10936-015-9357-3](https://doi.org/10.1007/s10936-015-9357-3)
14. Vélez-Uribe I., Rosselli M. Electrophysiological Correlates of Emotion Word Processing in Spanish–English Bilinguals // *Biling. Lang. Cogn.* 2021. Vol. 24, № 1. P. 31–55. DOI: [10.1017/S136672892000036X](https://doi.org/10.1017/S136672892000036X)
15. Chen P., Lin J., Chen B., Lu C., Guo T. Processing Emotional Words in Two Languages with One Brain: ERP and fMRI Evidence from Chinese–English Bilinguals // *Cortex.* 2015. Vol. 71. P. 34–48. DOI: [10.1016/j.cortex.2015.06.002](https://doi.org/10.1016/j.cortex.2015.06.002)
16. Боргоякова Т.Г. Социолингвистические процессы в республиках Южной Сибири: моногр. Абакан: Хакас. гос. ун-т им. Н.Ф. Катанова, 2002. 264 с.
17. Цыбенова Ч.С. О языковой компетенции тувинцев (результаты социологического анкетирования) // *Учен. зап. ЗабГГПУ.* 2013. № 2(49). С. 148–155.
18. Бавуу-Сюрюн М.В. Тувинский язык на современном этапе: образовательный аспект // *Новые исследования Тувы.* 2010. № 3. С. 57–71. URL: <https://nit.tuva.asia/nit/article/view/500/876> (дата обращения: 26.04.2022).
19. Lövheim H. A New Three-Dimensional Model for Emotions and Monoamine Neurotransmitters // *Med. Hypotheses.* 2012. Vol. 78, № 2. P. 341–348. DOI: [10.1016/j.mehy.2011.11.016](https://doi.org/10.1016/j.mehy.2011.11.016)
20. Birdsong D., Gertken L.M., Amengual M. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin. Web. 20 Jan. 2012. URL: <https://sites.la.utexas.edu/bilingual/> (дата обращения: 26.04.2022).

References

1. Daily S.B., James M.T., Cherry D., Porter J.J. III, Darnell S.S., Isaac J., Roy T. Affective Computing: Historical Foundations, Current Applications, and Future Trends. Jeon M. (ed.). *Emotions and Affect in Human Factors and Human-Computer Interaction.* Houghton, 2017, pp. 213–231. DOI: [10.1016/B978-0-12-801851-4.00009-4](https://doi.org/10.1016/B978-0-12-801851-4.00009-4)
2. Alm C.O., Rot D., Sproat R. Emotions from Text: Machine Learning for Text-Based Emotion Prediction. Mooney R., Brew C., Chien L.-F., Kirchoff K. (eds.). *Proceedings of Human Language Technology Conference and Conference*

on *Empirical Methods in Natural Language Processing*. Vancouver, 2005, pp. 579–586. Available at: <https://aclanthology.org/H05-1073.pdf> (accessed: 26 April 2022).

3. Volkova E.P., Mohler B.J., Meurers W.D., Gerdemann D., Bühlhoff H.H. Emotional Perception of Fairy Tales: Achieving Agreement in Emotion Annotation of Text. Inkpen D., Strapparava C. (eds.). *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Los Angeles, 2010, pp. 98–106. Available at: <https://aclanthology.org/W10-0212.pdf> (accessed: 26 April 2022).

4. Gunes H., Pantic M. Automatic, Dimensional and Continuous Emotion Recognition. *Int. J. Synth. Emot.*, 2010, vol. 1, no. 1, pp. 68–99. DOI: [10.4018/jse.2010101605](https://doi.org/10.4018/jse.2010101605)

5. Kolmogorova A., Kalinin A., Malikova A. Semiotic Function of Empathy in Text Emotion Assessment. *Biosemiotics*, 2021, vol. 14, no. 2, pp. 329–344. DOI: [10.1007/s12304-021-09434-y](https://doi.org/10.1007/s12304-021-09434-y)

6. Kolmogorova A.V., Kalinin A.A., Malikova A.V. Non-Discrete Sentiment Dataset Annotation: Case Study for Lövheim Cube Emotional Model. Alexandrov D.A., Boukhanovsky A.V., Chugunov A.V., Kabanov Y., Koltsova O., Musabirov I. (eds.). *Digital Transformation and Global Society. DTGS 2020. Communications in Computer and Information Science*. Cham, 2020. Vol. 1242, pp. 154–164. DOI: [10.1007/978-3-030-65218-0_12](https://doi.org/10.1007/978-3-030-65218-0_12)

7. Kazanas S.A., Wilck A.M., Altarriba J. Adaptive Memory: Greater Memory Advantages in Bilinguals' First Language. *Int. J. Biling.*, 2022, vol. 26, no. 1, pp. 49–64. DOI: [10.1177/13670069211022856](https://doi.org/10.1177/13670069211022856)

8. Pan L., Ke H., Styles S.J. Early Linguistic Experience Shapes Bilingual Adults' Hearing for Phonemes in Both Languages. *Sci. Rep.*, 2022, vol. 12. Art. no. 4703. DOI: [10.1038/s41598-022-08557-7](https://doi.org/10.1038/s41598-022-08557-7)

9. Mechelli A., Crinion J.T., Noppeney U., O'Doherty J., Ashburner J., Frackowiak R.S., Price C.J. Structural Plasticity in the Bilingual Brain: Proficiency in a Second Language and Age at Acquisition Affect Grey-Matter Density. *Nature*, 2004, vol. 431, p. 757. DOI: [10.1038/431757a](https://doi.org/10.1038/431757a)

10. Kroll J.F., Stewart E. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *J. Mem. Lang.*, 1994, vol. 33, no. 2, pp. 149–174. DOI: [10.1006/jmla.1994.1008](https://doi.org/10.1006/jmla.1994.1008)

11. Eilola T.M., Havelka J., Sharma D. Emotional Activation in the First and Second Language. *Cogn. Emot.*, 2007, vol. 21, no. 5, pp. 1064–1076. DOI: [10.1080/02699930601054109](https://doi.org/10.1080/02699930601054109)

12. Harris C.L., Gleason J.B., Ayçiçeği A. When Is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. Pavlenko A. (ed.). *Bilingual Minds: Emotional Experience, Expression, and Representation*. Clevedon, 2006, pp. 257–283. Available at: <https://www.bu.edu/psych/charris/papers/HarrisEmotionBiling.pdf> (accessed: 26 April 2022).

13. Kazanas S.A., Altarriba J. Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *J. Psycholinguist. Res.*, 2016, vol. 45, no. 2, pp. 395–406. DOI: [10.1007/s10936-015-9357-3](https://doi.org/10.1007/s10936-015-9357-3)

14. Vélez-Urbe I., Rosselli M. Electrophysiological Correlates of Emotion Word Processing in Spanish–English Bilinguals. *Biling. Lang. Cogn.*, 2021, vol. 24, no. 1, pp. 31–55. DOI: [10.1017/S136672892000036X](https://doi.org/10.1017/S136672892000036X)

15. Chen P., Lin J., Chen B., Lu C., Guo T. Processing Emotional Words in Two Languages with One Brain: ERP and fMRI Evidence from Chinese–English Bilinguals. *Cortex*, 2015, vol. 71, pp. 34–48. DOI: [10.1016/j.cortex.2015.06.002](https://doi.org/10.1016/j.cortex.2015.06.002)

16. Borgoyakova T.G. *Sotsiolingvisticheskie protsessy v respublikakh Yuzhnoy Sibiri* [Sociolinguistic Processes in the Republics of Southern Siberia]. Abakan, 2002. 264 p.

17. Tsybenova Ch.S. O yazykovoy kompetentsii tuvintsev (rezul'taty sotsiologicheskogo anketirovaniya) [On Language Competence of the Tuvans (Results of Sociolinguistic Survey)]. *Uchenye zapiski ZabGGPU*, 2013, no. 2, pp. 148–155.

18. Bavuu-Syuryun M.V. Tuvinskiy yazyk na sovremennom etape: obrazovatel'nyy aspekt [Tuvan Language on Modern Stage: Educational Aspect]. *Novye issledovaniya Tuvy*, 2010, no. 3, pp. 57–71. Available at: <https://nit.tuva.asia/nit/article/view/500/876> (accessed: 26 April 2022).

19. Lövheim H. A New Three-Dimensional Model for Emotions and Monoamine Neurotransmitters. *Med. Hypotheses*, 2012, vol. 78, no. 2, pp. 341–348. DOI: [10.1016/j.mehy.2011.11.016](https://doi.org/10.1016/j.mehy.2011.11.016)

20. Birdsong D., Gertken L.M., Amengual M. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. COERLL, University of Texas at Austin. Web. 20 January 2012. Available at: <https://sites.la.utexas.edu/bilingual/> (accessed: 26 April 2022).

DOI: 10.37482/2687-1505-V206

Anastasiya V. Kolmogorova

National Research University Higher School of Economics;
nab. kanala Griboedova 123, St. Petersburg, 190068, Russian Federation;
ORCID: <https://orcid.org/0000-0002-6425-2050> e-mail: nastiakol@mail.ru

Alina V. Malikova

Siberian Federal University;
prosp. Svobodnyy 82, str. 1, Krasnoyarsk, 660041, Russian Federation;
ORCID: <https://orcid.org/0000-0002-3438-1839> e-mail: malikovaav1304@gmail.com

MODELLING THE EMOTIONAL PERCEPTION OF RUSSIAN INTERNET TEXTS BY TUVAN-RUSSIAN BILINGUALS COMPARED TO RUSSIAN MONOLINGUALS: A PILOT STUDY

This study was carried out in the field of emotional text analysis. The aim of the paper was statistical modelling of the emotional perception of Internet texts in the Russian language by Tuvan-Russian bilinguals. A hypothesis about the significance of the factor of mental lexicon organization in the emotional markup of texts by this group was tested. We assumed that, due to the complexity of the psycholinguistic reality of the bilingual linguistic consciousness of the Tuvans, the markup would reflect some specific features in their processing of emotional information expressed verbally in a non-native, but communicatively meaningful language. The research material included 48 texts from the dataset of posts from the social network VKontakte: 6 texts for each emotional class according to Lövheim's three-dimensional model for emotions, i.e. distress, enjoyment, disgust, surprise, excitement, anger, shame, and fear. A computer interface developed and tested by the research group was used for non-discrete emotional markup of texts. The psycholinguistic experiment involved 174 Russian monolinguals and 65 Tuvan-Russian bilinguals. Before moving on to annotation, bilingual informants completed the Bilingual Language Profile questionnaire for us to obtain data on their dominant language. To analyse the results, we applied the Student's *t*-test and the comparative method. We found a statistical significance of differences in the affective evaluations attributed to texts in the Russian language by Russian monolinguals and Tuvan-Russian bilinguals (the majority of them were 18–25 years old). According to the results, the magnitude of these differences depends on the bilingual's dominant language. There is a discrepancy in secondary emotions between Russian monolinguals and Tuvan-Russian bilinguals, while the texts' leading emotion is common for both groups. In addition, we observed a certain instability in bilinguals' evaluations when they were asked to do the task again due to a technical error. The results of this research contribute to the study of bilinguals' perception of emotions in texts.

Keywords: *affective computing, bilingualism, dominant language, Internet text, psycholinguistics, Tuvan language, emotion analysis.*

Поступила 02.05.2022
Принята 30.09.2022
Опубликована 14.10.2022

Received 2 May 2022
Accepted 30 September 2022
Published 14 October 2022

For citation: Kolmogorova A.V., Malikova A.V. Modelling the Emotional Perception of Russian Internet Texts by Tuvan-Russian Bilinguals Compared to Russian Monolinguals: A Pilot Study. *Vestnik Severnogo (Arkticheskogo) federal'nogo universiteta. Ser.: Gumanitarnye i sotsial'nye nauki*, 2022, vol. 22, no. 5, pp. 39–50. DOI: 10.37482/2687-1505-V206