

ЭЛЕКТРОННАЯ ЛЕКСИКОГРАФИЯ

ELECTRONIC LEXICOGRAPHY

Научная статья

УДК 81'23

doi: 10.17223/22274200/26/5

Психолингвистическая база данных TurkWordPerception как лексикографический источник (оценки вклада модальностей восприятия в семантику)

**Зоя Ивановна Резанова¹, Валерия Евгеньевна Владимирова²,
Тимур Евгеньевич Машанло³**

^{1, 2, 3} Томский государственный университет, Томск, Россия

¹ rezanovazi@mail.ru

² picture_perfect@mail.ru

³ mashanlote@gmail.com

Аннотация. Представлена психолингвистическая база данных об оценках вклада пяти модальностей восприятия в формирование семантики существительных, прилагательных и глаголов татарского и хакасского языков. База данных характеризуется как лексикографический источник: представлены формирование словника, адаптация анкет для сбора данных, содержание и структура данных в базе. Характеризуются источниковедческие возможности базы данных в психолингвистических исследованиях и приводятся результаты первичного статистического анализа данных.

Ключевые слова: психолингвистические базы данных, модальности восприятия, татарский язык, хакасский язык, анкетирование, шкала Ликерта

Благодарности: Исследование выполнено при поддержке Программы развития Томского государственного университета (Приоритет-2030).

Для цитирования: Резанова З.И., Владимирова В.Е., Машанло Т.Е. Психолингвистическая база данных TurkWordPerception как лексикографический источник (оценки вклада модальностей восприятия в семантику)// Вопросы лексикографии. 2022. № 26. С. 96–114. doi: 10.17223/22274200/26/5

Original article

doi: 10.17223/22274200/26/5

Psycholinguistic database TurkWordPerception as a lexicographic source (assessments of the contribu- tion of perception modalities to semantics)

Zoya I. Rezanova¹, Valeria E. Vladimirova², Timur E. Mashanlo³

^{1, 2, 3} Tomsk State University, Tomsk, Russian Federation

¹ rezanovazi@mail.ru

² picture_perfect@mail.ru

³ mashanlote@gmail.com

Abstract. This article presents TurkWordPerception (<http://clingv.ru:3839/>), a psycholinguistic database (PBD) on the assessment of 600 words of the Khakas and Tatar languages. The PBD was created within the framework of the project for studying the linguistic diversity of Southern Siberia. The article substantiates the theoretical foundations, data structure, methods of collecting the material. Using the descriptive statistical data based on the database materials, source studies capabilities of the PBD are assessed. When creating the database, the authors employed works in the field of the theory of embodied cognition, the study of linguistic relativity and the theory of the bilingual mental lexicon. The database includes assessments of 200 nouns, 200 adjectives, and 200 verbs for both Tatar and Khakas, which are translated equivalents of RuWordPerception data words. Information about the assessment of the relationship between the words of the two Turkic languages and the modalities of perception was collected using the questionnaire method based on the seven-point Likert scale. Data collection was carried out using paper questionnaires or in the experiment created in the PsychoPy software. Metadata about the respondents were also collected based on the sociolinguistic questionnaire and the language experience questionnaire. Currently, the database includes assessments of 600 words of the Tatar language in 5 modalities of perception, received from 75 respondents – 142,225 assessments; 600 words of the Khakas language in 5 modalities of perception, received from 93 respondents – 188,370 assessments. The article uses descriptive statistics to monitor primary data sources, which are compared with

similar RuWordPerception data. Speakers of the Tatar and Khakas languages give assessments of strong connection with all modalities of perception and of average connection – above 4. Tatar speakers give lower assessments in all modalities of perception compared to Khakas speakers (the significance level $p=0.00$ according to the Mann-Whitney test). Comparison of these data with assessments of Russian translated equivalents given by speakers of the Tatar and Khakas languages revealed a high degree of generality of the level of assessments given by representatives of the two language communities, with an inverse relationship: the Tatar give higher assessments of Russian translation equivalents compared to the Khakas. Evaluating the words of the Russian language, the speakers of Tatar and Khakas give lower assessments to them than to the words of their native languages. TurkWordPerception data can serve as a basis for conducting a multilateral analysis with the data obtained in world psycholinguistics on the material of other languages, and in general will contribute to the introduction of the Tatar and Khakas languages into the paradigm of psycholinguistic research.

Keywords: psycholinguistic databases, modalities of perception, Tatar language, Khakas language, questioning, Likert scale

Acknowledgments: The study was supported by the Tomsk State University Development Programme (Priority-2030).

For citation: Rezanova, Z.I., Vladimirova, V.E. & Mashanlo, T.E. (2022) Psycholinguistic database TurkWordPerception as a lexicographic source (assessments of the contribution of perception modalities to semantics). *Voprosy leksikografii – Russian Journal of Lexicography*. 26. pp. 96–114. (In Russian). doi: 10.17223/22274200/26/5

Создание баз данных сегодня является одним из активно развивающихся прикладных направлений различных научных исследований, в том числе в лингвистике и психолингвистике. Данный тип электронных источников в настоящее время дополняет ресурсную базу проводимых исследований, представленную традиционными и электронными словарями и корпусами текстов.

Отбор единиц в базах данных и их параметризация определяются как общими особенностями предметной сферы, так и проблемными вопросами в частных научных направлениях, на решение которых они направлены. Вследствие этого состав базы данных и параметризация входящих в нее единиц в разной степени могут определяться частными теориями в рамках какой-либо науки (см. например, об этом [1]). При этом создаваемые базы данных могут быть представлены и в виде традиционных печатных изданий, как, например, психолингвисти-

ческая база данных о перцептивной и оценочной семантике русских прилагательных в словаре М.Г. Колбеновой, Ю.И. Александрова [2], так и, что гораздо более распространено, в виде электронных ресурсов, снабженных функцией автоматического поиска данных (например, [3–7] и др.).

Отметим, что для современного этапа развития психолингвистики характерным является создание двух типов баз данных. Создаются психолингвистические базы данных (ПБД), включающие оценки тысячами респондентов десятков, иногда сотен тысяч единиц и являющиеся результатом проведения мегаисследований международными коллективами (The English lexicon project [3], The French Lexicon Project [4], The Malay lexicon project [5] и др.). В качестве соотносительного с данными проектами следует отметить единственный проект на материале русского языка «StimulStat Project» («Лингвистическая база данных для слов русского языка»), в котором более 50 000 наиболее частотных слов русского языка и их форм описаны более чем по 97 различным параметрам [5, 7].

В то же время продолжают создаваться ПБД, ориентированные на проведение факторных исследований, направленные на проверку частных гипотез и теорий (см. обзор проектов данной весьма развитой традиции в [8], сайт, содержащий ссылки на мега-исследования, – в [9]). Одним из направлений развития таких проектов является вовлечение материалов все новых языков, создающих основу проведения сравнительных исследований. И все же количество языков, нормативные данные о психолингвистических параметрах единиц которых представлены в доступных для анализа ПБД, невелико. На материале русского языка и такие проекты, и исследования, выполняемые на их базе, единичны, при этом подчеркнем их значимость для включения данных о нормах оценок по психолингвистическим параметрам русского языка в общемировую парадигму и проведение собственно компаративных исследований [10–13]. Задача привлечения материалов других языков народов Российской Федерации при создании ПБД, которые бы могли послужить материалом проведения психолингвистических исследований, ждет своего решения. Нам известен только один проект ПБД, реализованный с вовлечением материала татарского языка. Это «Библиотека стимулов. Глаголы и существительные». В ней представлены соотнесенные данные об оценках единиц русского и татарского языков (696 существительных и 407 глаголов: данные по параметрам) [14].

В данной статье авторы представляют созданную в рамках проекта по исследованию языкового разнообразия Южной Сибири ПБД TurkWordPerception, включающую данные об оценках 600 слов русского и татарского языков по 5 модальностям восприятия. Представляемая база данных теоретически, по составу словника, методике сбора и обработки данных соотнесена с созданной ранее в проекте ПБД «RuWordPerception» [15]. «RuWordPerception» включает системы оценок 5 модальностей восприятия 200 существительных, 200 прилагательных, 200 глаголов русского языка носителями данного языка как родного и тюркско-русскими билингвами, являющимися носителями шорско-русского, хакасско-русского, татарско-русского билингвизма. Концепция психолингвистической базы данных, созданной на предшествующих этапах развития проекта RuWordPerception, представлена в статье З.И. Резановой, Т.Е. Машанло, А.А. Степаненко [16]. Кратко охарактеризуем объединяющие данные ПБД теоретические основания.

При создании БД мы основывались на работах в области теории воплощенного познания, прежде всего на работах Л. Барсолу [17, 18]. Основная идея теории воплощённого познания – познание/мышление основано на телесных механизмах и не может быть рассмотрено в отрыве от внешних условий его протекания. По теории Л. Барсолу, при использовании слов естественного языка, имеющих символьную природу, осуществляется активация соответствующих перцептивных символов и формирование симуляций, основывающихся на опыте восприятия мира органами чувств, что, как правило, не находится под контролем сознания. Основная направленность ПБД – представление данных о субъективных оценках носителями языков вклада перцептивного опыта, формируемого активацией пяти органов чувств, в семантику слов. Теоретической основой TurkWordPerception являются также положения гипотезы лингвистической относительности об обратном влиянии языка на восприятие и интерпретацию опыта физических контактов человека в «реальном мире». Подробнее об этом мы писали в [16], в данном случае еще раз подчеркнем, что, несмотря на то, что основные идеи данной гипотезы были высказаны почти столетием назад, вокруг нее не затихают споры. Новый всплеск интереса к гипотезе связан с развитием когнитивной науки, особенно той ее части, которая направлена на поиск экспериментальных подтверждений или опровержений положения об обусловленности когнитивных процессов языковой формой. Создаваемая база данных представляет эм-

пирический материал, который в соотнесении с данными, уже опубликованными БД RuWordPerception, позволит выявить наличие или отсутствие различий восприятия одних и тех же концептов в разных языках, что может быть использовано в качестве эмпирического материала для экспериментальной проверки данной гипотезы. Ранее проведенный на основе данных RuWordPerception сравнительный анализ оценок русских существительных по модальностям восприятия носителями русского языка как родного и тюркско-русскими билингвами выявил как наличие значительной общности в системе оценок, так и частных различий. Результаты кластерного анализа показали, что в группе носителей русского языка как родного проявляются связи между визуальной и тактильной модальностью, а в группе татарско-русских билингвов соответствующая связь устанавливается с аудиальной модальностью [19]. Обнаружение таких различий закономерно ставит вопрос об их истоках, которые, очевидно, базируются в различиях языкового маркирования основных перцепций в структурах материнских тюркских языков билингвов. Для выявления характера данных влияний и создавалась на следующем этапе развития проекта база данных оценок слов тюркского языка, репрезентирующих в материнских языках билингвов те же концепты, которые были отражены в лексическом составе русскоязычной базы данных. Таким образом, третьей базовой теорией при создании представляемой ПБД является теория ментального лексикона билингва, основная идея которой заключается в утверждении наличия в языковом сознании билингва двух систем, взаимодействующих при формировании трех компонентов ментальных репрезентаций: концептов – семантики – языковой формы [20].

При общности теоретических основ двух создаваемых баз данных мы естественным образом стремились также к сохранению в них общих контуров структуры данных. Однако в характеризующей ПБД по отношению к предшествующей были внесены частные изменения. TurkWordPerception включает только системы оценок слов по модальностям восприятия, в ней на данном этапе не представлены данные о субъективной частотности и возрасте усвоения, единицы не характеризуются по собственно лингвистическим характеристикам в структурах соответствующих языков, но данные позволяют проводить анализ языковых детерминаций на уровнях лексической и грамматической частеречной категоризации. Данные о субъективной ча-

стотности и возрасте усвоения татарских и хакасских слов вошли в третью создаваемую коллективом Лаборатории лингвистической антропологии ПБД RuTurkPsychLing.

Социолингвистические характеристики респондентов, оценки их языкового опыта собирались в соответствии с разработанной для БД RuWordPerception системой.

Словник базы данных формировался переводными эквивалентами русскоязычной базы данных RuWordPerception. В первичной ПБД состав формировался по принципу случайной выборки единиц, направленной на выявление некоего усредненного показателя связей слов трех базовых частей речи с модальностями восприятия без ориентации на отдельные лексико-семантические группы (ЛСГ). Однако материалы позволяют делать выборки по ЛСГ, которые могут собираться автоматически по семантическим тегам [15]. Переводы были осуществлены носителями языка, сотрудниками Лаборатории лингвистической антропологии, на татарский язык – Р. Дусеевым, на хакасский – А. Кольчиковой. При переводе мы руководствовались несколькими критериями: если слово русского языка имело несколько значений, выбиралось то, которое является более частотным; если слово имеет несколько вариантов перевода на татарский или хакасский, то для базы данных отбирался более частотный вариант; в состав словника базы данных включались переводы, состоящие из одного слова; если у слова было два варианта перевода (заимствование из русского слова и исконное тюркское слово), выбирался второй вариант. Выбор осуществлялся составителями базы данных на основе словарей татарского и хакаского языков и с учетом экспертных оценок носителей этих языков, членов Лаборатории лингвистической антропологии. Примеры соотнесенных словников представлены в табл. 1.

Для сбора оценок слов трех частей речи татарского и хакаского языков по вкладу в их семантику модальностей восприятия использован метод анкетирования по методике шкалирования (семичленная шкала Ликета). Содержательно инструкции в своей основе соответствовали представленным в RuWordPerception) (см.: [16]). Так как носители татарского и хакаского языков в совершенстве владеют русским языком, на котором получали образование и который в настоящее время является функционально доминирующим, вытесняя материнские языки в сферы семейного и обыденного общения, анкетирование проводилось на русском языке.

Т а б л и ц а 1

Примеры соотнесенных лексем в составе двух баз данных

Русский язык	Татарский язык	Хакасский язык
Прилагательные		
мыльный	сабынлы	сабынныҥ
рыжий	жирән	чигірен
трогательный	дулкынландыргыч	черексіністіг
червивый	кортлы	хурттыҥ
ягодный	жиләк	чистектіг
Существительные		
муха	чебен	сеек
нож	пычак	пычак
отец	әти	паба
сила	көч	күс
чайник	чәйнек	чейник
Глаголы		
смотреть	карга	көрерге
уколоть	чәнчергә	хазабызарға
храпеть	гырылдарга	харых тартарға
украсть	урларга	оғырлап аларға
хлопать	кагарга	хол сабарға

Уровень владения русским языком устанавливался по данным анкетирования (использовалась социолингвистическая анкета и анкета языкового опыта билингва). При этом в текст инструкции в качестве типовых примеров оценивания вклада конкретных модальностей восприятия были включены слова из родных для респондентов языков – хакасского и татарского. Данные слова – калибраторы – отбирались экспертами составителями инструкций на основе данных словарей и оценивались экспертами, носителями соответствующих языков, Р. Дусеевым и А. Кольчиковой. В качестве таких единиц включались слова, которые называли признаки, объекты, действия, имеющие высокий вклад конкретных модальностей (например, калибраторы оценки вклада аудиальной модальности на татарском языке: чиртеш «щелчок», сөйләм «разговорный», чакырырга «звать»). Оценки таких слов в дальнейшем использовались для контроля действий респондента: при наличии сильной связи с соответствующей модальностью восприятия в качестве допустимого принимался диапазон оценок от 5 до 7. В повторяющихся случаях неверных оценок калибраторов респондент не мог пройти основную часть анкетирования.

На рис. 1 представлена инструкция, адресованная носителям татарского языка. Формат сбора данных претерпел некоторые изменения, что было сделано в результате критического анализа опыта сбора данных в первую ПБД. Мы отказались от отдельных анкет для каждой из модальностей, что было реализовано при сборе данных в RuWordPerception. Участники сразу оценивали слово по всем модальностям (рис. 1 с примером бумажной анкеты). Содержательно инструкции соответствовали представленным в RuWordPerception.

Разные слова могут быть связаны с разными ощущениями – зрением, слухом, вкусом, запахом и осязанием. Например, слово *табыдыг* связано в большей степени со вкусом, в то время как слово *хызарара* – со зрением; слово не обязательно связано только с одним ощущением, например, слово *торгы* может вызывать ассоциации как со зрением, так и с осязанием, однако у одного человека слово *торгы* теснее связано со зрением, а у другого – с осязанием.

Ваша задача – оценить то, как сильно связано указанное слово с различными ощущениями. Для оценивания используйте нижеприведенную таблицу, в которой 1 обозначает отсутствие у слова связи с конкретным ощущением, а 7 указывает на максимально тесную связь слова с данным ощущением. Пожалуйста, при оценивании слов используйте всю шкалу от 1 до 7.

1 Не связано	2 Очень слабо связано	3 Слабо связано	4 Связано ни слабо, ни тесно	5 Тесно связано	6 Очень тесно связано	7 Максимально тесно связано
ipiri	Зрение:					
	Осязание:					
	Вкус:					
	Запах:					
	Слух:					
sirpic	Зрение:					
	Осязание:					
	Вкус:					
	Запах:					
	Слух:					
хлорка	Зрение:					
	Осязание:					
	Вкус:					
	Запах:					
	Слух:					
ачын	Зрение:					
	Осязание:					
	Вкус:					
	Запах:					
	Слух:					

Рис. 1. Пример инструкции, использованной в исследовании

Мы отказались от применения онлайн-анкетирования с помощью гугл-анкет, так как обработка данных, полученных на предшествующем этапе, выявила значительное количество недобросовестно заполненных анкет респондентами. Мы вернулись к практике сбора данных с присутствием экспериментатора в двух вариантах – с использованием бумажных анкет и на основе созданного в программном обеспечении PsychoPy эксперимента, проводимого на компьютерах в лаборатории (рис. 2 с примером эксперимента в электронном виде). Анкетирование при помощи PsychoPy давало возможность рандомизировать порядок представления стимулов, бумажные анкеты были представ-

лены в двух вариантах на каждый язык с целью рандомизации представления стимулов. Процедура рандомизации задаваемых стимулов была направлена на устранение двух нежелательных эффектов при сборе данных с предъявлением списка слов: 1) уставание респондента, рассеивание внимания; 2) эффект праймирования соседними словами в списке.

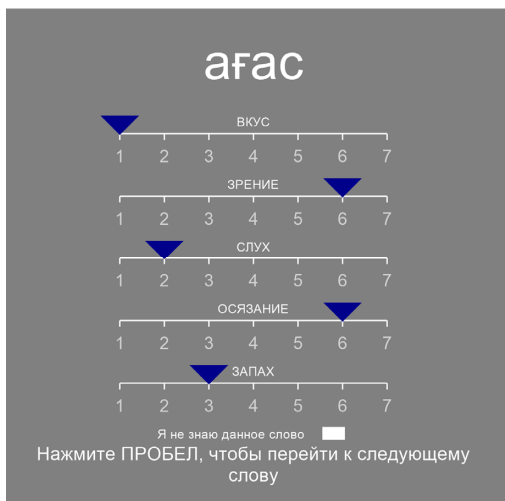


Рис. 2. Фрагмент эксперимента в программном обеспечении PsychoPy

Содержательно инструкции соответствовали представленным в RuWordPerception.

В настоящее время база данных включает оценки 600 слов татарского языка по 5 модальностям восприятия, полученные от 75 респондентов (среди которых 42 – мужчины), средний возраст респондентов – 27 лет, $SD = 12$, все участники – носители хакасского как родного, владеют русским в совершенстве, по их субъективным оценкам, называют русский родным языком наравне с татарским – 142 225 реакций; 600 слов хакасского языка по 5 модальностям восприятия, полученные от 93 респондентов (среди которых 52 – мужчины), средний возраст респондентов – 26 лет, $SD = 9$, все участники носители хакасского как родного, владеют русским в совершенстве – по их субъективным оценкам, называют русский родным языком наравне с хакасским – 188 370 реакций.

Объем собранных данных постоянно пополняется, но уже на данном этапе собранный материал позволяет проводить исследования влияния факторов этнокультурных и этноязыковых различий (тюркско-русское языковое взаимодействие) на восприятие эквивалентных лексических единиц.

Далее мы приведем результаты базового анализа данных в их соотношении с соответствующими данными «RuWordPerception» на примере фрагмента ПБД – оценок по основным модальностям восприятия слов. Статистический анализ был реализован на языке программирования R 3.3.1.

На начальном этапе данные были усреднены и затем были получены результаты описательной статистики оценок по 5 модальностям восприятия 600 слов, татарских и хакасских переводных аналогов, данных 2 группами респондентов – татарско-русскими билингвами, хакасско-русскими билингвами.

Данные описательной статистики представлены в табл. 2 (Средн. – среднее значение; Мин., Макс. – минимальное и максимальное значения; L1-T – данные об оценках носителями языка слов татарского языка; L1-Kh – данные об оценках носителями языка слов хакасского языка).

Т а б л и ц а 2

Описательные статистики оценок по 5 модальностям восприятия 600 слов хакасского и татарского языков татарско-русскими и хакасско-русскими билингвами

Модальности восприятия	Средн.		Мин.		Медиана		Макс.	
	L1-Kh	L1-T	L1-Kh	L1-T	L1-Kh	L1-T	L1-Kh	L1-T
Связь со зрением	4.12	4.06	3.59	3.36	4.13	4.05	4.58	5.03
Связь со слухом	4.09	3.95	3.66	3.11	4.09	3.94	4.61	4.86
Связь с обонянием	4.05	3.92	3.55	3.23	4.04	3.92	4.55	4.62
Связь со вкусом	4.04	3.91	3.41	3.31	4.04	3.90	4.49	4.78
Связь с осязанием	4.09	3.98	3.61	3.40	4.092	3.96	4.69	4.67

Как можно видеть, носители татарского и хакасского языков 1) дают высокие оценки связи со всеми модальностями восприятия, средние – выше 4 (при семичленной шкале оценок); 2) наблюдаются различия в оценках – носители татарского языка дают более низкие

оценки по всем модальностям восприятия по сравнению с носителями хакасского языка. Данное различие достигает уровня статистической значимости, что подтверждается сравнением двух групп по критерию Манна–Уитни (этот критерий применен, так как распределение данных не является нормальным), уровень значимости $p < 0,001$. Подчеркнем, что данные описательной статистики представляют результаты, усредненные относительно единиц языка, выбранных методом случайной выборки, данные относительно единиц отдельных частей речи исследователь может получить, извлекая из базы данных оценки слов отдельных лексико-семантических групп.

Сравним эти данные с оценками русских переводных эквивалентов, указанными носителями татарского и хакасского языков по материалам ПБД RuWordPerception, представленные в табл. 3 (система маркирования данных та же, что и в табл. 2).

Т а б л и ц а 3

Описательные статистики оценок по 5 модальностям восприятия 600 слов русского языка татарско-русскими и хакасско-русскими билингвами

Модальности восприятия	Средн.		Мин.		Медиана		Макс.	
	L1-T	L1-Kh	L1-T	L1-Kh	L1-T	L1-Kh	L1-T	L1-Kh
Связь со зрением	3.32	3.26	1.60	1.22	3.21	3.13	6.28	6.44
Связь со слухом	2.81	2.55	1.37	0.90	2.59	2.29	6.02	6.30
Связь с обонянием	2.45	2.14	1.34	1.09	2.06	1.77	5.88	6.00
Связь со вкусом	2.31	2.14	1.31	1.12	1.92	1.80	6.31	6.34
Связь с осязанием	2.98	2.66	1.46	1.23	2.85	2.47	5.87	5.78

Сравнение этих данных с оценками русских переводных эквивалентов, указанных носителями татарского и хакасского языков (по материалам ПБД RuWordPerception) свидетельствует о том, что: 1) в этом случае также проявляется значительная степень общности уровня оценок, данных представителями двух языковых сообществ, однако татары, в противоположность оценкам слов родного языка, дают более высокие оценки русским переводным эквивалентам по сравнению с хакасами; 2) оценивая слова русского языка, носители татарского и хакасского дают им менее высокие оценки, нежели словам

родных языков, – даже наиболее высокая связь со зрением получила средние оценки менее 4, другие модальности восприятия – менее 3.

Представим данные в диаграммах размаха. На рис. 3–6 в графиках на оси *X* маркированы цифрами оценки вклада модальностей восприятия: 1 – аудиальное, 2 – обонятельное, 3 – вкусовое, 4 – тактильное, 5 – зрительное; на оси *Y* – средние оценки по 7-балльной шкале.

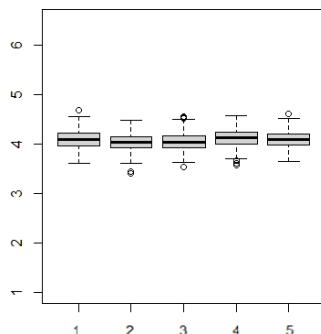


Рис. 3. Диаграмма размаха оценок связи с модальностями восприятия слов хакасского языка носителями хакасского языка

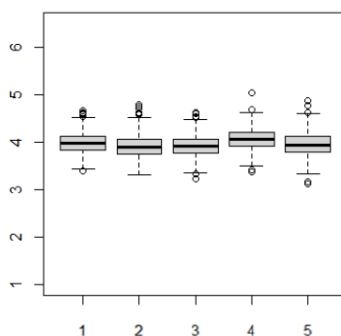


Рис. 4. Диаграмма размаха оценок связи с модальностями восприятия слов татарского языка носителями татарского языка

Далее представим диаграммы размаха оценок, данных носителями татарского и хакасского языков словам русского языка. Как можно видеть при сравнении диаграмм, при общей направленности распределения оценок наблюдаются значительные отличия.

Отметим общий характер соотношения оценок – более высокий уровень оценок связей слов со зрительной модальностью (5 из 7 на всех рисунках при максимально возможной 7) и наиболее низкий – с обонятельной и вкусовой (2 и 3 на всех рисунках, при минимальной 1), которые дают носители двух языков при оценке слов и родного и второго, русского, языка.

Однако при этом мы видим существенные различия в перепадах уровней различий. Оценивая по степени связи с обонятельной и вкусовой модальностью слова родного языка, хакасы и татары дают им значимо более высокие оценки, по сравнению с русскими переводными эквивалентами (ср. обоняние: хакасы 4.05 vs 2.11, татары 3.929 vs 3.93; вкус: хакасы 4.04 vs 2.14; татары 3.91 vs 2.31).

Однако и разброс оценок русских слов билингвами намного больше. Об этом свидетельствует значительное количество выбросов: это не случайные оценки или ошибки, от которых можно избавиться. Это показатель разнообразия выборки и оценок респондентов, с чем также связана ненормальность распределения.

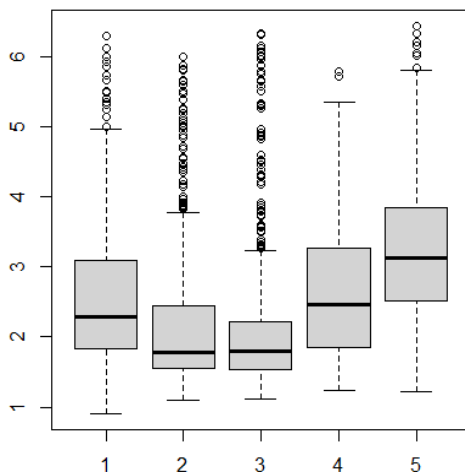


Рис. 5. Диаграмма размаха оценок носителями хакасского языка связи русских слов с модальностями восприятия

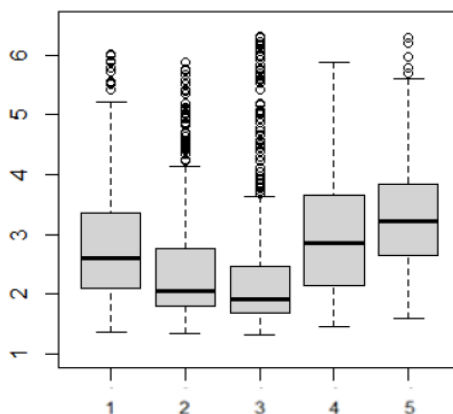


Рис. 6. Диаграмма размаха оценок носителями татарского языка связи русских слов с модальностями восприятия

Как следует из приведенного в таблицах и рисунках размаха данных, с одной стороны, наблюдаются общие закономерности, выявленные в материалах русскоязычной ПБД, а именно более высокие оценки зрительной и аудиальной модальностей, однако был обнаружен высокий уровень близости средних оценок полученных данных у татарско-русских и хакасско-русских билингвов и значительное отличие оценок, которые давали билингвы переводным эквивалентам – русским словам, где наблюдался значительно больший разброс оценок по разным типам модальностей восприятия. В результате проведенного первичного анализа мы делаем вывод, что одни и те же концепты, названные респондентам через посредство номинаций русского и родного языка, оценивались по связям с модальностями восприятия со значительной долей различий.

Таким образом, в статье была охарактеризована база данных как лексикографический источник, включающая информацию о психолингвистически релевантных параметрах 600 слов татарского и хакаского языков. Данные о вкладе модальностей восприятия учитываются при отборе стимульного материала в психолингвистических экспериментальных исследованиях процессов когнитивной обработки слов. Созданная база данных будет способствовать расширению эмпирической базы такого типа исследований в мировой психолингвистике. Данные TurkWordPerception могут послужить основой проведения их многостороннего сравнения с данными, полученными в ми-

ровой психолингвистике на материале других языков, и будут способствовать введению татарского и хакасского языков в парадигму психолингвистических исследований.

TurkWordPerception является постоянно пополняемой ПБД, но уже собранный материал позволяет сделать первичный вывод о том, что одни и те же концепты, поименованные переводными эквивалентами родного и второго, осваиваемого языка, оцениваются по вкладу модальностей восприятия по-разному, т.е. перцептивные компоненты занимают разное положение в структуре концептов в ментальном лексиконе билингвов, соотносимых с номинациями из системы родного и осваиваемого языков. Применение других методов статистического анализа к выборкам единиц, формируемым по заданным в базе данных основаниям, позволит выявить более сложные системы соотношения оценок вклада модальностей восприятия как внутри выделенных групп лексики, так и между группами.

Список источников

1. Мишанкина Н.А. Базы данных в лингвистических исследованиях // Вопросы лексикографии. 2013. № 1 (3).
2. Колбенева М.Г., Александров Ю.И. Органы чувств, эмоции и прилагательные русского языка: Лингво-психологический словарь. М. : Языки славянских культур, 2010. 369 с.
3. Balota D.A., Yap M.J., Cortese M.J., Hutchison K.A., Kessler B., Loftus B., Neely J.H., Nelson D.L., Simpson G.B., Treiman R. The English lexicon project: A user's guide // Behav. Res. Methods. 2007. № 39. P. 445–459.
4. Yap M. J., Rickard Ferrand L., New B., Brysbaert M., Keuleers E., Bonin P., Méot A., Augustinova M., Pallier C. The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords // Behav. Res. Methods. 2010. № 42. P. 488–496.
5. Liow S.J., Jalil S. B., & Faizal S. S.B. The Malay lexicon project: A database of lexical statistics for 9,592 words // Behav. Res. Methods. 2010. № 42. P. 992–1003.
6. StimulStat Project / Лингвистическая база данных для слов русского языка. URL: http://stimul.cognitivestudies.ru/ru_stimul (дата обращения: 06.10.2019).
7. Alexeeva S., Slioussar N., Chernova D. StimulStat: A lexical database for Russian // Behav. Res. Methods. 2018. № 50. P. 2305–2315. doi: 10.3758/s13428-017-0994-3
8. Keuleers E., Balota D.A. Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments // The Quarterly Journal of Experimental Psychology. 2015. № 68(8). P. 1457–1468. doi: 10.1080/17470218.2015.1051065
9. Word megastudy data available. URL: <http://crr.ugent.be/programs-data/megastudy-data-available> (дата обращения: 13.03.2022).

10. Miklashevsky A. Perceptual Experience Norms for 506 Russian Nouns: Modality Rating, Spatial Localization, Manipulability, Imageability and Other Variables // Journal of psycholinguistic research. 2018. Vol. 47. № 3. P. 641–661. doi: 10.1007/s10936-017-9548-1

11. Grigoriev A., Oshhepkov I. Objective age of acquisition norms for a set of 286 words in Russian: relationships with other psycholinguistic variables // Behav. Res. Methods. 2013. № 45 (4). P. 1208–1217. doi: 10.3758/s13428-013-0319-0. PMID: 23435657

12. Akinina Y., Malyutina S., Ivanova M., Iskra E., Mannova E., Dragoy O. Russian normative data for 375 action pictures and verbs // Behav. Res. Methods. 2015. № 47 (3). P. 69–707. doi:10.3758/s13428-014-0492-9

13. Резанова З.И., Миклашевский А.А. Моделирование образно-перцептивного компонента языковой семантики при помощи психолингвистической базы данных // Вестник Томского государственного университета. Филология. 2016. № 5 (43). С. 71–92. doi: 10.17223/19986645/43/6

14. Акинина Ю.С., Грабовская М.А., Вечкаева А.М., Игнатъев Г.А., Исав Д.Ю., Ханова А.Ф. Библиотека психолингвистических стимулов: новые данные для русского и татарского языка // Седьмая международная конференция по когнитивной науке : тез. докл. Светлогорск, 20–24 июня 2016 г. М. : Ин-т психологии РАН, 2016. С. 93–95.

15. Психолингвистическая база данных RuWordReception. URL: <http://clingv.ru:3839/> (дата обращения: 06.03.2022).

16. Резанова З.И., Машанло Т.Е., Степаненко А.А. Перцептивный компонент семантики существительных, прилагательных, глаголов русского языка в билингвальной перспективе (психолингвистическая база данных RuWordPerception) // Вестник Томского государственного университета. 2020. № 450. С. 49–57. doi: 10.17223/15617793/450/6

17. Barsalou L.W. Grounded cognition // Annu. Rev. Psychol. 2008. Vol. 59. P. 617–645. doi: 10.1146/annurev.psych.59.103006.093639

18. Barsalou L.W., Simmons W.K., Barbey A.K., Wilson C.D. Grounding conceptual knowledge in modality-specific systems // Trends in cognitive sciences. 2003. Vol. 7. № 2. P. 84–91. doi: 10.1016/S13646613(02)000293

19. Резанова З.И., Степаненко А.А. Перцептивный компонент семантики имен существительных в восприятии носителей русского языка как родного и тюрко-русских билингвов (на основе психолингвистической базы данных RuWordPerception) // Вестник Томского государственного университета. 2020. № 455. С. 32–40. doi: 10.17223/15617793/455/5

20. Pavlenko A. The Bilingual Mental Lexicon: Interdisciplinary Approaches. Bristol : Multilingual Matters, 2009. 272 p.

References

1. Mishankina, N.A.(2013) Databases in Linguistic Research. *Voprosy leksikografii – Russian Journal of Lexicography*. 1 (3). (In Russian).

2. Kolbeneva, M.G. & Aleksandrov, Yu.I. (2010) *Organy chuvstv, emotsii i prilagatel'nye russkogo yazyka: Lingvo-psikhologicheskii slovar'* [Sense organs, emotions and adjectives of the Russian language: Linguistic and psychological dictionary]. Moscow: Yazyki slavyanskikh kul'tur.
3. Balota, D.A. et al. (2007) The English lexicon project: A user's guide. *Behav. Res. Methods*. 39. pp. 445–459.
4. Yap, M.J. et al. (2010) The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behav. Res. Methods*. 42. pp. 488–496.
5. Liow, S.J., Jalil, S.B. & Faizal S.S.B. (2010) The Malay lexicon project: A database of lexical statistics for 9,592 words. *Behav. Res. Methods*. 42. pp. 992–1003.
6. *StimulStat Project | Lingvisticheskaya baza dannyykh dlya slov russkogo yazyka* [StimulStat Project | Lexical database for Russian language]. [Online] Available from: http://stimul.cognitivestudies.ru/ru_stimul (Accessed: 06.10.2019).
7. Alexeeva, S., Slioussar, N. & Chernova, D. (2018) StimulStat: A lexical database for Russian. *Behav. Res. Methods*. 50. pp. 2305–2315. doi: 10.3758/s13428-017-0994-3
8. Keuleers, E. & Balota, D.A. (2015) Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *The Quarterly Journal of Experimental Psychology*. 68(8). pp. 1457–1468. doi: 10.1080/17470218.2015.1051065
9. Crr.ugent.be. (2022) *Word megastudy data available*. [Online] Available from: <http://crr.ugent.be/programs-data/megastudy-data-available> (Accessed: 13.03.2022).
10. Miklashevsky, A. (2018) Perceptual Experience Norms for 506 Russian Nouns: Modality Rating, Spatial Localization, Manipulability, Imageability and Other Variables. *Journal of Psycholinguistic Research*. 47 (3). pp. 641–661. doi: 10.1007/s10936-017-9548-1.
11. Grigoriev, A. & Oshhepkov, I. (2013) Objective age of acquisition norms for a set of 286 words in Russian: relationships with other psycholinguistic variables. *Behav. Res. Methods*. 45 (4). pp. 1208–1217. doi: 10.3758/s13428-013-0319-0. PMID: 23435657
12. Akinina, Y. et al. (2015) Russian normative data for 375 action pictures and verbs. *Behav. Res. Methods*. 47 (3). pp. 69–707. doi:10.3758/s13428-014-0492-9
13. Rezanova, Z.I. & Miklashevskiy, A.A. (2016) Modeling of the perceptual-based component of language semantics using a psycholinguistic database. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya – Tomsk State University Journal of Philology*. 5 (43). pp. 71–92. (In Russian). doi: 10.17223/19986645/43/6
14. Akinina, Yu.S. et al. (2016) [Library of psycholinguistic stimuli: new data for the Russian and Tatar languages]. *Abstracts of the Seventh International Conference on Cognitive Science*. Svetlogorsk. 20–24 June 2016. Moscow: Institute of Psychology of the Russian Academy of Sciences. pp. 93–95. (In Russian).
15. RuWordReception Psycholinguistic Database. [Online] Available from: <http://clingv.ru:3839/> (Accessed: 06.03.2022).
16. Rezanova, Z.I., Mashanlo, T.E. & Stepanenko, A.A. (2020) A Bilingual Perspective on the Perceptual Component of Word Semantics in Russian: The RuWordPerception Psycholinguistic Database. *Vestnik Tomskogo gosudarstvennogo*

universiteta – Tomsk State University Journal. 450. pp. 49–57. (In Russian). doi: 10.17223/15617793/450/6

17. Barsalou, L.W. (2008) Grounded cognition. *Annu. Rev. Psychol.* 59. pp. 617–645. doi: 10.1146/annurev.psych.59.103006.093639

18. Barsalou, L.W. et al. (2003) Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*. 7 (2). pp. 84–91. doi: 10.1016/S13646613(02)000293

19. Rezanova, Z.I. & Stepanenko, A.A. (2020) The Perceptual Component of Noun Semantics in the Perception of Native Russian Speakers and Turkic-Russian Bilinguals (Based on the RuWordPerception Psycholinguistic Database). *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 455. pp. 32–40. (In Russian). doi: 10.17223/15617793/455/5

20. Pavlenko, A. (2009) *The Bilingual Mental Lexicon: Interdisciplinary Approaches*. Bristol: Multilingual Matters.

Сведения об авторах:

Резанова Зоя Ивановна – д-р филол. наук, зав. кафедрой общей, компьютерной и когнитивной лингвистики ТГУ, заместитель зав. Лабораторией лингвистической антропологии Томского государственного университета (Томск, Россия). E-mail: rezanovazi@mail.ru

Владимирова Валерия Евгеньевна – канд. филол. наук, младший научный сотрудник Лаборатории лингвистической антропологии Томского государственного университета (Томск, Россия). E-mail: picture_perfect@mail.ru

Машанло Тимур Евгеньевич – канд. филол. наук, младший научный сотрудник Лаборатории лингвистической антропологии Томского государственного университета (Томск, Россия). E-mail: mashanlote@gmail.com

Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Zoya I. Rezanova, Dr. Sci. (Philology), head of the Department of General, Computational and Cognitive Linguistics, deputy head of the Linguistic Anthropology Laboratory, Tomsk State University (Tomsk, Russian Federation). E-mail: rezanovazi@mail.ru

Valeria E. Vladimirova, Cand. Sci. (Philology), junior researcher, Linguistic Anthropology Laboratory, Tomsk State University (Tomsk, Russian Federation). E-mail: picture_perfect@mail.ru

Timur E. Mashanlo, Cand. Sci. (Philology), junior researcher, Linguistic Anthropology Laboratory, Tomsk State University (Tomsk, Russian Federation). E-mail: mashanlote@gmail.com

The authors declare no conflicts of interests.

*Статья поступила в редакцию 14.03.2022;
одобрена после рецензирования 06.09.2022; принята к публикации 02.12.2022.*

*The article was submitted 14.03.2022;
approved after reviewing 06.09.2022; accepted for publication 02.12.2022.*