

ЭЛЕКТРОННАЯ ЛЕКСИКОГРАФИЯ

ELECTRONIC LEXICOGRAPHY

Научная статья

УДК 81'23

doi: 10.17223/22274200/29/5

База данных RuTurkPsychLing как специализированный лексикографический источник: нормы оценок психофизиологического и эмоционального опыта

**Валерия Евгеньевна Владимирова¹, Ирина Сергеевна Коршунова²,
Зоя Ивановна Резанова³**

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

¹ picture_perfect@mail.ru

² korshunova-61818@mail.ru

³ rezanovazi@mail.ru

Аннотация. Представлен электронный специализированный словарь, организованный как база данных, отличающийся типом представляемой информации о единицах, принципами сбора данных, их предварительной обработки, организации и хранения; охарактеризованы его источниковедческие возможности. RuTurkPsychLing объединяет нормативные данные об оценках слов русского, хакасского и татарского языков по психолингвистическим параметрам знакомости, температуры, расположения в пространстве, размера, эмоциональности и манипулируемости.

Ключевые слова: психолингвистические базы данных, специализированные словари, электронные словари, семантика слова, нормы восприятия, тюркско-русский билингвизм

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

Для цитирования: Владимирова В.Е., Коршунова И.С., Резанова З.И. База данных RuTurkPsychLing как специализированный лексикографический источник: нормы оценок психофизиологического и эмоционального опыта // Вопросы лексикографии. 2023. № 29. С. 87–105. doi: 10.17223/22274200/29/5

Original article

doi: 10.17223/22274200/29/5

RuTurkPsychLing database as a specialized lexicographic source: Norms for evaluating psychophysiological and emotional experience

**Valeriia E. Vladimirova¹, Irina S. Korshunova²,
Zoya I. Rezanova³**

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

¹ picture_perfect@mail.ru

² korshunova-61818@mail.ru

³ rezanovazi@mail.ru

Abstract. The article presents a psycholinguistic database (DB) RuTurkPsychLing – a dictionary of a special type, which differs both in the type of information provided about lexemes and in the nature of collecting and organizing the information. The dictionary includes information about lexical units obtained using psycholinguistic methods and intended for use primarily in cognitive and psycholinguistic research. Dictionary data are organized using database technology to improve the efficiency of extracting and analyzing the dictionary data. The purpose of RuTurkPsychLing is to present normative data on the evaluation of words of the Russian language and two Turkic languages – Tatar and Khakass – according to the following parameters: “familiarity”, “temperature”, “location in space”, “size”, “emotionality” and “manipulability”. The DB correlates with the theoretical basis and methods of collecting material with local DBs created in the Tomsk State University Linguistic Anthropology Lab – RuWordPerception and TurkWordPerception. Combining the evaluations of Russian words given by Russian L1 speakers and Turkic-Russian bilinguals allows obtaining evidence of the influence of the bilingual mental lexicon on the perception of words; evaluation of equal concepts in native languages by three groups of participants will provide an empirical basis for the theoretical provisions of the hypothesis of linguistic relativity. The article describes the questionnaire procedure, methods and overall analysis of the collected data. The RuTurkPsychLing DB includes 364,155 ratings given by 100 native Russian speakers; 74,320 evaluations of Russian and Tatar words given by 16 Tatar-Russian bilinguals; 371,200 evaluations of Russian and Khakass words, given by 34 Khakass-Russian bilinguals. Participants rate the contribution of the parameters to the semantics as low, except for the relationship with emotionality (the mean score is from 3.78 to 4.06 out of 7). According to the data collected in the project, the contribution of perceptual components to the semantics of randomly sampled words is extremely small. Perceptual components of size,

position in space, temperature, emotionality and manipulability do not appear in the entire lexicon, but in specific semantic scopes. Comparison of ratings given by bilinguals in Russian and in their native language shows the common ratio. However, the native speakers of Tatar significantly differed in their evaluations of the Russian words in terms of the warm/cold parameter, the difference retained while evaluating the Tatar words, but manifested itself to a lesser extent. The materials of the psycholinguistic DB can be used to test existing hypotheses, to generate new hypotheses, to establish new patterns in the field of cognitive processing of units of the native and acquired languages, and to cross-validate the results of other studies, both psycholinguistic and proper linguistic.

Keywords: psycholinguistic databases, specialized dictionaries, electronic dictionaries, word semantics, perception norms, Turkic-Russian bilingualism

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

For citation: Vladimirova, V.E., Korshunova, I.S. & Rezanova, Z.I. (2023) RuTurkPsychLing database as a specialized lexicographic source: Norms for evaluating psychophysiological and emotional experience. *Voprosy leksikografii – Russian Journal of Lexicography*. 29. pp. 87–105. (In Russian). doi: 10.17223/22274200/29/5

В статье представлен словарь RuTurkPsychLing, созданный с использованием технологий баз данных (БД).

Бурное развитие лексикографии в XX в. обусловило появление значительного количества лексикографических собраний, различающихся как типом единиц (слова, морфемы, словосочетания и т.д.), так и характером информации о единицах. Информация может быть получена с применением методологического аппарата разных наук: культурологии, политологии, психологии и под. RuTurkPsychLing включает данные о лексических единицах, полученные с использованием психолингвистических методик, и эта информация предназначена прежде всего для когнитивных и психолингвистических исследований.

Второе направление развития современной лексикографии связано с активным внедрением технологий автоматической обработки данных в практику подготовки лексикографических изданий. Современные электронные словари снабжаются автоматическими поисковыми системами, многократно повышающими эффективность их использования. Одним из технологических решений организации данных в электронных словарях являются технологии баз данных, применяемые в различных предметных сферах, общими признаками которых

мы вслед за Е.А. Мирошниченко считаем следующие: «1) любая БД хранится и обрабатывается в вычислительной системе; 2) БД является сильно структурированной системой; 3) БД включает схему данных» [1. С. 86]. База данных RuTurkPsychLing представляет психолингвистически релевантную информацию о единицах словаря в соответствии с этими принципами.

В статье характеризуются оба аспекта созданного лексикографического ресурса: 1) специфическая информация, представляемая в словаре (базе данных), ее место и востребованность в современной психолингвистической парадигме; 2) принципы сбора материала, их обработки и хранения на основе применения технологий БД.

Создание БД RuTurkPsychLing включено в решение актуальной задачи современной мировой психолингвистики – формирование баз данных, содержащих рейтинги слов по когнитивно значимым параметрам, влияющим на процессы их обработки. Ресурсы таких баз данных используются при проведении психолингвистических экспериментов, направленных на моделирование скрытых процессов в сознании носителя при восприятии слов естественных языков. При построении дизайна таких экспериментов исследователю необходимо учитывать то, какие параметры слова могут оказать потенциальное воздействие на скорость и эффективность распознавания и понимания слова. Экспериментаторы обычно учитывают как объективную частотность с опорой на корпусные данные, так и ее субъективные оценки, данные представителями определенной языковой группы, а также связи с другими субъективными факторами, как, например, возраст усвоения и концептуальная знакомость [2–4].

В настоящее время существует значительное количество словарей, объединяющих психолингвистические и собственно лингвистические характеристики слов в технологиях БД, представляющих нормативные данные о рейтингах психолингвистических параметров слов разных языков: модальности восприятия, доминирование, конкретность, образность, знакомость, возраст усвоения, ассоциированность с гендером, размером, расположением в пространстве, манипулируемость и др. [2–6].

Однако при этом мы отмечаем, что, во-первых, не все параметры слов, когнитивная значимость которых доказана в уже проведенных экспериментах, представлены в нормативных БД; во-вторых, наблюдается значительная диспропорция по представленности конкретных языков в уже реализованных проектах.

Описываемая в статье БД, полное зарегистрированное название которой «Психолингвистическая база данных RuTurkPsychLing: оценки слов русского, татарского, хакасского языков по параметрам “знакомость”, “температура”, “расположение в пространстве”, “размер”, “эмоциональность” и “манипулируемость”» [7], направлена на преодоление сложившейся диспропорции. Цель ее создания – представление нормативных данных об оценках слов русского и двух тюркских языков, татарского и хакасского, по обозначенным выше параметрам.

RuTurkPsychLing соотнесена по теоретическим основаниям, структуре и методам сбора материала с двумя БД, созданными ранее в лаборатории лингвистической антропологии ТГУ. База данных RuWordPerception включает оценки 200 прилагательных, 200 глаголов и 200 существительных русского языка, данные носителями русского языка как родного и тюркско-русскими билингвами по пяти модальностям восприятия (зрение, вкус, обоняние, слух, осязание), а также по параметрам возраста усвоения и субъективной частотности. База данных находится в открытом доступе на сайте <http://clingv.ru:3839/> [8]. База данных TurkWordPerception включает оценки переводных эквивалентов слов из RuWordPerception, данные носителями татарского и хакасского языков по пяти модальностям восприятия [9].

Далее охарактеризуем БД RuTurkPsychLing в соотнесении с ранее созданными БД и представим психолингвистически релевантные параметры слов, данные о которых собираются.

Своеобразие любого словаря во многом определяется составом словника и принципами его отбора. RuTurkPsychLing объединяет словник RuWordPerception и TurkWordPerception: в ней собраны оценки русских слов и их переводных аналогов в татарском и хакасском языках (195 существительных, 194 прилагательных и 200 глаголов). Сокращение словников произошло из-за того, что некоторые тюркские эквиваленты при переводе совпадали, в связи с этим дубликаты были удалены из словника. Приведем примеры таких словников в табл. 1, 2.

Объединяют представляемую БД RuTurkPsychLing с ранее созданными и основные теории, на которые мы опирались и которые были представлены в [8, 9].

Основное теоретическое положение, к которому мы обращаемся, – положение теории воплощенного познания, заключается в том, что по-

знание и языковая семантика укоренены в теле, языковая концептуализация основана на восприятии среды и пережитого опыта [10], вследствие этого семантику можно систематически описать с точки зрения телесного опыта и модальностей восприятия.

Таблица 1

**Фрагмент словника с совпадением
переводных эквивалентов на хакасском языке**

Хакасские существительные	Русские существительные
агын	течение
агырин полганы	замедление
агырыг	болезнь
агырыг	боль
азыранчаң чир	столовая

Таблица 2

**Фрагмент словника с совпадением переводных эквивалентов
на татарском языке**

Татарские существительные	Русские существительные
кайнау	кипение
канәфер	сирень
карау	просмотр
карау	уход
качкын	беглец

Модальности восприятия составляют базу исследований перцептивной лексической семантики [10, 11]. Однако, во-первых, перцептивные каналы – это не единственное, что может влиять на восприятие человека, во-вторых, в процессах восприятия перцептивная информация может вступать в разного рода интеракции с другими каналами восприятия и – отраженно – с другими свойствами слова. Состав психолингвистических параметров, по которым собираются оценки слов, составляет важнейшее отличительное свойство данной БД по отношению к ранее созданным.

В RuTurkPsychLing собраны оценки слов 1) по параметрам, являющимися частными по отношению к модальностям восприятия, представленным в ранее созданных БД, например, тактильная модальность конкретизирована через признак восприятия температуры, визуальная – через признак расположения в пространстве и параметр

размера; 2) по другим параметрам, которые находятся в сложных, неоднозначных отношениях с модальностями восприятия, например, параметр манипулируемости может относиться как к визуальной, так и тактильной модальностям восприятия.

Далее кратко представим исследования, в которых доказывалась релевантность включенных в БД параметров в процессах когнитивной обработки слов, а также существующие описания нормативных данных на материале разных языков, если таковые имеются.

В ряде экспериментов подтверждено перцептивное моделирование пространственных характеристик: значимость противопоставления высокого и низкого, а также ближнего и дальнего пространства, в том числе в процессах восприятия предложений [12–15]. Этот параметр представлен в ряде психолингвистических баз данных, в том числе на материале русского языка – для существительных [6]. Положение в пространстве соотносено с психолингвистическим параметром манипулируемости (так как возможность взаимодействия с объектом руками уменьшается по мере удаления объекта от субъекта, его воспринимающего) и размером (увеличение размера на определенном этапе уменьшает возможность манипулируемости). Взаимодействие с референтом слова при помощи рук рассматривается как отдельный перцептивный компонент [16]; нормативные данные о параметре представлены в БД [8, 9].

Перцептивный компонент «размер» исследуется в аспекте кросс-модального аудиовизуального взаимодействия, звукового символизма. Э. Сепир выявил кроссмодальные ассоциации между звуками и размером объекта [17], нормы для этого перцептивного параметра зафиксированы для английского языка [4].

Психолингвистический параметр соотносительности с температурой относится к осязательной модальности, при этом изучается он также в кроссмодальном аспекте (связь температуры и цвета) [18]. Психолингвистические нормы для данного параметра не собраны ни на одном языке. Рейтинги соотносительности перцептивной и interoцептивной семантики с эмоциональностью активно исследуются [19–21] и представлены в настоящее время также и в нормативных базах данных [5].

Концептуальная знакомость слова, будучи соотносительной с субъективной частотностью, представлена в нормативной базе на материале английского языка [4].

Таким образом, мы констатируем выявленную в психолингвистических исследованиях значимость данных параметров, и в то же вре-

мя они в разной степени представлены в современных БД. На материале русского, татарского и хакасского языков языка нормативные данные по перечисленным параметрам, кроме положения в пространстве для выборки русских существительных, собираются впервые. Более того, нам не известны БД, собирающие оценки слов русского языка, данные билингвами, хотя эта проблема является весьма актуальной в психолингвистике, так как установлено, что при оценке слов одного языка второй язык билингва может влиять в разной степени на когнитивные процессы, что, например, подчеркивается в работе Дж. Скотта и др. [4].

Объединение в одной базе данных оценок русских слов по психолингвистически значимым параметрам, данным носителями русского языка как родного и тюркско-русскими билингвами, позволяет получать эмпирические доказательства влияния особенностей ментального лексикона билингва на восприятие слов. Объединение в одной БД оценок слов, являющихся переводными эквивалентами в трех языках, даст эмпирическую основу для проверки теоретических положений гипотезы лингвистической относительности.

Охарактеризуем далее процедуру анкетирования и собранные данные.

При проектировании новой БД было соблюдено единство принципов создания анкет трех БД, использовался принцип ранжирования ответов на основе семичленной шкалы Ликерта (см. развернутую характеристику в [8, 9]).

Сбор данных был реализован в двух вариантах. На начальном этапе данные собирались в процедуре эксперимента, написанного на Python3 с использованием оболочки PsychoPy3 на платформе PsychoPy [22] в лаборатории лингвистической антропологии ТГУ.

Таким образом словник, принципы составления анкет и сбора данных были едиными для трех БД общего проекта лаборатории, варьирование было обусловлено потребностями адаптировать анкеты к особенностям исследуемых параметров и внешним условиям проведения анкетирования.

Объем БД, организация данных, результаты описательной статистики

В настоящее время RuTurkPsychLing включает 364 155 оценок слов русского языка, данных 100 носителями русского языка как род-

ного (из них 7 мужчин), средний возраст 23 года; 74 320 оценок слов русского и татарского языков, данных 16 татарско-русскими билингвами (из них 6 мужчин), средний возраст 29 лет; 371 200 оценок слов русского и хакасского языков, данных 34 носителями хакасского языка как родного, хакасского-русскими билингвами (из них 14 мужчин), средний возраст 21 год.

База данных представлена в виде 8 файлов с расширением .csv, представленных в табличном формате, где каждая колонка является переменной, а в каждой строчке представлено значение данной переменной (см. пример на рис. 1). Основные файлы содержат оценки слов русского и тюркских языков по шкалам: связь слов с размером, расположением, температурой, эмоциональностью, связь слов с конкретным размером, конкретным расположением, конкретной температурой, конкретной эмоциональной оценкой, возможность взаимодействовать с объектом/выполнять действие руками, концептуальная знакомость.

participant_ID	word	translation_exp	pos	param	value
32	жанлек тиресе	пушнина	noun	t_lower	0
31	жанлек тиресе	пушнина	noun	space_upper	8
30	жанлек тиресе	пушнина	noun	size_upper	1
30	жанлек тиресе	пушнина	noun	know	2
32	жанлек тиресе	пушнина	noun	space_upper	1
32	жанлек тиресе	пушнина	noun	space_lower	8
30	жанлек тиресе	пушнина	noun	t_upper	1
32	жанлек тиресе	пушнина	noun	size_upper	5
30	жанлек тиресе	пушнина	noun	e_lower	0
30	жанлек тиресе	пушнина	noun	t_lower	0
32	жанлек тиресе	пушнина	noun	e_lower	0
31	жанлек тиресе	пушнина	noun	size_upper	5
30	жанлек тиресе	пушнина	noun	size_lower	0
30	жанлек тиресе	пушнина	noun	space_upper	1
32	жанлек тиресе	пушнина	noun	size_lower	0

Рис. 1. Структура tatar_full.csv, фрагмент таблицы

Объединение всех файлов в единую структуру позволяет автоматически извлекать данные в различных комбинациях и загружать в системы статистического анализа для выявления количественных закономерностей, проверки с помощью статистических методов теоретических гипотез.

Далее представим результаты первичной статистической обработки данных об оценках слов в представляемой БД. Анализ проводился в Rstudio с использованием языка R версии 4.0.5.

Мы рассчитали средние величины и другие базовые показатели для соотнесенных выборок. В таблице 3 представлены нормативные данные об оценках слов русского языка (L1-R), по психолингвистическим параметрам в сравнении с данными об оценках этих же слов носителями татарско-русского (L1-T) и хакасско-русского (L1-Kh) билингвизма. Во второй строке табл. 3, 4 отмечаются средние (Сред.), медиана (Мед.), минимальные (Мин.) и максимальные (Макс.) оценки в выборке.

Таблица 3

Описательные статистики оценок слов русского языка носителями русского как родного и тюркско-русскими билингвами

	L1-R				L1-T				L1-Kh			
	Сред.	Мед.	Мин.	Макс.	Сред.	Мед.	Мин.	Макс.	Сред.	Мед.	Мин.	Макс.
Эмоциональная оценка	4.06	4.00	1.63	6.43	3.86	3.75	1.00	7.00	3.78	3.67	1.53	6.20
Положительная/отрицательная оценка	2.44	1.94	1.20	5.80	2.13	1.83	1.00	6.50	2.54	2.27	1.46	5.51
Знакомость	6.86	6.95	4.43	7.00	6.87	7.00	2.00	7.00	6.92	6.98	5.16	7.00
Манипулируемость	2.81	2.33	1.10	6.56	4.05	4.00	1.00	7.00	2.19	1.74	1.12	6.33
Маленький/большой	4.15	4.17	1.75	6.68	3.87	4.00	1.00	7.00	4.22	4.25	1.98	6.32
Связь с размером	1.96	1.69	1.17	5.15	1.83	1.57	1.00	6.00	2.23	2.06	1.55	5.50
Низко/высоко	3.69	3.62	1.49	6.80	4.99	5.00	1.00	7.00	4.06	4.07	2.36	6.35
Связь с положением в пространстве	2.06	1.82	1.21	5.80	2.60	2.57	1.00	5.71	2.26	2.12	1.42	5.59
Холодно/тепло	4.10	4.14	1.67	6.79	2.20	2.00	1.00	6.33	4.46	4.50	2.17	6.48
Связь с температурой	1.65	1.36	1.03	6.22	1.59	1.25	1.00	6.00	2.10	1.92	1.57	6.23

Респонденты всех трех выборок (носители русского языка как родного и билингвы) дали высокие оценки слов по параметру знакомости; это свидетельствует о том, что по психолингвистическим параметрам оценивались слова, субъективно оцениваемые как знакомые в высокой степени. Сравнивая минимальные оценки, мы видим, что татарско-русские билингвы все же в меньшей степени знакомы с оцениваемыми словами (ср. минимальные значения по параметру L1-R – 4.43, L1-T – 2, L1-Kh – 5.16).

Стоит отметить, что оценки группы татарско-русских билингвов значительно отличаются от двух других групп. В таблице 3 представлены усредненные данные, и мы можем видеть, что все носители татарского языка имели тенденцию ставить крайние оценки.

Тем не менее, как можно видеть по данным, представленным в табл. 3, участники склонны оценивать соотношенность слов с большинством психолингвистических параметров ниже средней, ближе к её отсутствию, за исключением связи с эмоциональной оценкой, где связь оценивается от 3.78 до 4.06, при среднем 3.5 (оценки выше среднего отмечены в таблице полужирным шрифтом). В целом, проанализировав рейтинги вклада исследованных параметров в семантику случайной выборки слов, собранных в БД, мы можем сделать вывод о том, что вклад перцептивных компонентов в семантику слов крайне мал. Это является свидетельством в пользу того, что перцептивные компоненты соотношенности с размером, положением в пространстве, температурой, эмоциональностью и манипулируемостью проявляются не во всем лексиконе языка, но скорее в определенных семантических группах.

В тех группах слов, для которых такая связь устанавливается, распределение оценок показывает, что предметный мир, отраженный в существительных, глаголах и прилагательных, воспринимается как скорее теплый, нежели холодный (средние в группе L1-R – 4.10, L1-Kh – 4.46), расположенный скорее высоко, нежели низко (средние в группе L1-R – 3.69, L1-Kh – 4.06, L1-T – 4.99), скорее большой, нежели маленький (L1-R – 4.15, L1-Kh – 4.22, L1-T – 3.87) и, скорее положительный, нежели отрицательный (L1-R – 2.44, L1-Kh – 2.54, L1-T – 2.13). Отметим, что по последнему параметру была использована «перевернутая» шкала, на которой 1 – максимально положительная оценка, 7 – максимально отрицательная оценка).

Сравнение оценок слов, данных носителями русского языка как родного и билингвами, отражает ту же тенденцию, которая наблюдалась в

материалах БД RuWordPerception и TurkWordPerception – принципиальное совпадение в направленности оценок в группах носителей языка как родного и билингвов, носителей трех языков как родных. Случаи наибольшего несовпадения оценок выделены в таблице курсивом. К ним относятся оценки по параметру «холодный/теплый», где татарско-русские билингвы оценили слова представленной выборки в среднем как скорее холодные (2.20), а носители русского языка как родного и хакасско-русские билингвы – как скорее теплые (4.10 и 4.46 соответственно), также значимо отличаются средние татарско-русских билингвов по параметрам «высокий/низкий» (4.99 против 3.69 и 4.06) и «большой/маленький» (3.87 против 4.15 и 4.22).

Для визуализации на рис. 2, 3 мы выбрали следующие параметры: возможность взаимодействия с руками (*manip*), связь слов с конкретной температурой (*t_lower*), с конкретной эмоциональной оценкой (*emo_lower*), с конкретным размером (*size_lower*), с конкретным положением в пространстве (*space_lower*).

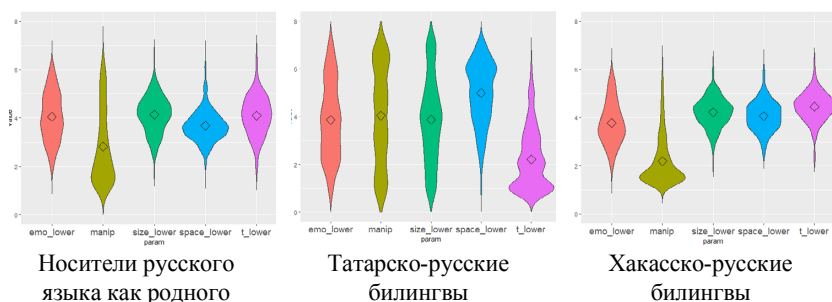


Рис. 2. Оценки слов русского языка как носителями как родного и билингвами

Как видим на рис. 2 и 3, наблюдается (1) большой разброс оценок в группе татарско-русских билингвов, (2) смещенность оценок по параметру температуры к нижним пунктам на шкале. Более сопоставимые распределения оценок наблюдаются в группах носителей русского языка как родного и хакасско-русских билингвов. Тем не менее оценки билингвами манипулируемости тяготеют к низким значениям, распределения оценок остальных параметров незначительно ниже в группе носителей русского языка как родного.

Далее в табл. 4 приведены сравнительные данные об оценках переводных эквивалентов в татарском и хакасском языках по тем же

параметрам и те же оценки русских слов носителями русского языка как родного, представленные выше в табл. 3.

Таблица 4

Описательные статистики оценок слов родного языка носителями русского, татарского, хакасского языков

	L1-T				L1-Kh				L1-R			
	Сред	Мед.	Мин.	Макс.	Сред	Мед.	Мин.	Макс.	Сред	Мед.	Мин.	Макс.
Эмоциональность	3.85	3.83	1.00	6.33	3.85	3.82	1.64	6.09	4.06	4.00	1.63	6.43
Положительная / отрицательная оценка	2.19	2.10	1.00	5.56	2.50	2.34	1.40	5.18	2.44	1.94	1.20	5.80
Знакомость	4.50	4.44	1.89	7.00	6.23	6.40	4.43	6.96	6.86	6.95	4.43	7.00
Манипулируемость	3.06	2.90	1.11	7.00	2.52	2.43	1.83	5.68	2.81	2.33	1.10	6.56
Маленький/большой	3.78	3.67	1.00	7.00	4.19	4.21	2.32	6.06	4.15	4.17	1.75	6.68
Связь с размером	2.15	2.00	1.00	5.11	2.20	2.06	1.60	5.68	1.96	1.69	1.17	5.15
Низко/высоко	3.63	3.50	1.00	7.00	4.10	4.11	2.31	5.92	3.69	3.62	1.49	6.80
Связь с положением	2.58	2.50	1.00	6.00	2.11	2.04	1.52	4.92	2.06	1.82	1.21	5.80
Холодно/тепло	3.72	3.75	1.00	7.00	4.40	4.40	2.21	6.00	4.10	4.14	1.67	6.79
Связь с температурой	1.74	1.62	1.00	5.00	2.04	1.94	1.42	5.89	1.65	1.36	1.03	6.22

Как было отмечено ранее, мы спрашивали респондентов о характере владения языками, и в анализ включались данные тех респондентов, которые владеют двумя языками, однако степень владения в двух выборках оказалась неравномерной, что отразилось в различии оценок по степени знакомости слов родных языков у татарских и хакасских билингвов (4.50 vs 6.23). При этом у выборки татарско-русских билингвов степень знакомости со словами татарского языка (переводными эквивалентами) оказалась ниже, чем знакомость соответствующих русских слов.

Мы можем предположить, что разный уровень знакомости слов мог оказать влияние на характер оценок: оценки слов на родном языке носителями русского языка и носителями хакасского языка (на русском и на хакасском соответственно) оказались ближе, нежели

оценки татарских слов носителями татарского как родного. Сравнение оценок, данных билингвами русским словам и словам родного языка (табл. 3, 4), показывает, что в целом общность их направленности сохраняется, однако наблюдаются и различия. Так, носители татарского языка значительно отличались в оценках русских слов по параметру «теплый/холодный», различие сохранилось и при оценивании слов родного языка, но проявилось в меньшей степени (3,72 – татарские слова против 2.20 – русские). Также татары значительно отличались от носителей русского языка как родного и хакасско-русских билингвов в оценках русских слов по параметру «большой/маленький», эта же тенденция сохранилась и при оценивании ими слов родного языка. Рейтинги параметра «высокий/низкий» в группе носителей татарского языка не совпали у слов родного языка и русского (3.63 против 4.99).

Далее также представим визуализацию распределения оценок в диаграммах на рис. 3.

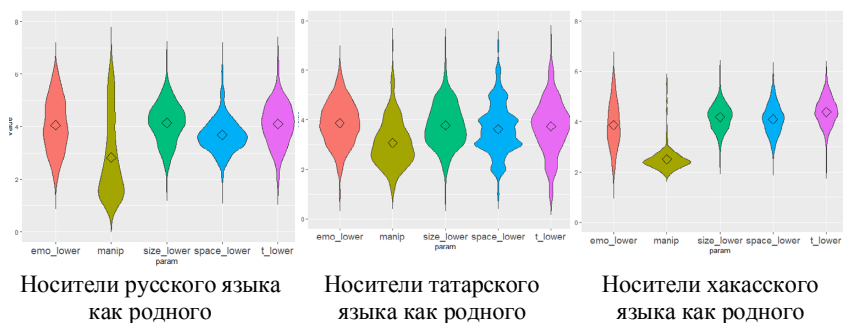


Рис. 3. Оценки слов родного языка носителями русского, татарского, хакасского языков

Как можно видеть на диаграммах, во всех трех группах наблюдается тенденция к низким (ниже среднего) оценкам по параметру «манипулируемость». Разброс рейтингов в группе татарско-русских билингвов несколько снизился у слов родного языка по сравнению с рейтингами этого параметра у слов русского языка. Существенно снизился разброс оценок, данных носителями хакасского языка, оценки более однородные и тяготеют к средним; выявлен низкий рейтинг манипулируемости.

Таким образом, в статье был охарактеризован электронный психолингвистический словарь, организованный как база данных, представлены его источниковедческие возможности в сфере психолингвистических и когнитивных исследований. Нормативные данные о психолингвистически релевантных параметрах 589 слов русского, татарского и хакасского языков могут использоваться при отборе стимульного материала в психолингвистических экспериментальных исследованиях процессов когнитивной обработки слов. Ресурсы электронного словаря, организованные с использованием технологий баз данных, будут способствовать включению лексиконов трех языков в парадигму сравнительных исследований в мировой экспериментальной психолингвистике. Материалы RuTurkPsychLing могут использоваться для проверки имеющихся гипотез на новом объективном материале, генерации новых гипотез, установления новых закономерностей в сфере когнитивной обработки единиц родного и осваиваемого языка, кросс-валидации результатов других исследований, как психолингвистических, так и собственно лингвистических.

Список источников

1. *Мирошниченко Е.А.* К формальному определению понятия «база данных» // Проблемы информатики. 2011. № 2 (10). С. 83–87.
2. *Soares A.P., Costa A.S., Machado J., Comesaña M., Oliveira H.M.* The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words // Behavior Research Methods. 2016. № 49 (3). P. 1065–1081. doi: 10.3758/s13428-016-0767-4
3. *Bonin P., Méot A., Bugaiska A.* Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times // Behavior Research Methods. 2018. № 50 (6). doi: 10.3758/s13428-018-1014-y
4. *Scott G.G., Keitel A., Becirspahic M., Yao B., Sereno S.C.* The Glasgow Norms: Ratings of 5,500 words on nine scales // Behavior Research Methods. 2018. № 51 (1, Pt. 2). doi: 10.3758/s13428-018-1099-3
5. *Speed L.J., Brybaert M.* Dutch sensory modality norms // Behavior Research Methods. 2022. № 54. P. 1306–1318. doi: 10.3758/s13428-021-01656-9
6. *Miklashevsky A.* Perceptual Experience Norms for 506 Russian Nouns: Modality Rating, Spatial Localization, Manipulability, Imageability and Other Variables // Journal of Psycholinguistic Research. 2018. № 47 (3). P. 641–661. doi: 10.1007/s10936-0179548-1
7. *Свидетельство о государственной регистрации базы данных № 2022622657* Российская Федерация. Психолингвистическая база данных RuTurkPsychLing: оценки слов русского, татарского, хакасского языков по параметрам «знакомость», «температура», «расположение в пространстве», «размер», «эмоциональ-

ность» и «манипулируемость»: № 2022622423 : заявл. 10.10.2022 : опубл. 27.10.2022 / В.Е. Владимирова, В.С. Диброва, А.С. Душейко [и др.] ; заявитель Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Томский государственный университет».

8. 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.

9. 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

10. Wilson M. Six views of embodied cognition // Psychonomic bulletin review. 2002. Vol. 9, № 4. P. 625–636.

11. 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

12. Dudschig C., Souman J., Lachmair M., de la Vega I., Kaup B. Reading “sun” and looking up: The influence of language on saccadic eye movements in the vertical dimension // PLoS One. 2013. № 8 (2) e56872. doi: 10.1371/journal.pone.0056872

13. Estes Z., Verges M., Barsalou L.W. Head up, foot down object words orient attention to the objects’ typical location // Psychological Science. 2008 № 19 (2). P. 93–97. doi: 10.1111/j.1467-9280.2008.02051.x

14. Cowey A., Small M., Ellis S. Left visuo-spatial neglect can be worse in far than in near space // Neuropsychologia. 1994. № 32 (9). P. 1059–1066. doi: 10.1016/0028-3932(94)90152-X

15. Winter B., Bergen B. Language comprehenders represent object distance both visually and auditorily // Language and Cognition. 2012. № 4 (1). P. 1–16.

16. Campanella F., Shallice T. Manipulability and object recognition: is manipulability a semantic feature? // Experimental Brain Research. 2011. Vol. 208, № 3. P. 369–383. doi: 10.1007/s00221-010-2489-7.

17. Sapir E. A study in phonetic symbolism // Journal of Experimental Psychology. 1929. № 12. P. 225–239.

18. Ho H.-N., Van Doorn G.H., Kawabe T., Watanabe J., Spence C. Colour-Temperature Correspondences: When Reactions to Thermal Stimuli Are Influenced by Colour // PLoS One. 2014. № 9 (3): e91854. doi: 10.1371/journal.pone.0091854

19. Larsson M., Willander J., Karlsson K., Arshamian A. Olfactory LOVER: Behavioral and neural correlates of autobiographical odor memory // Frontiers in Psychology. 2014. № 5. P. 1–5. doi: 10.3389/fpsyg.2014.00312.

20. Yeshurun Y., Sobel N. An odor is not worth a thousand words: From multidimensional odors to unidimensional odor objects // Annual Review of Psychology. 2010. № 61. P. 219–241. doi: 10.1146/annurev.psych.60.110707.163639

21. Winter B. Taste and smell words form an affectively loaded part of the English lexicon // Language, Cognition and Neuroscience. 2016. № 31 (8). P. 975–988.

22. Peirce J.W., Gray J.R., Simpson S., MacAskill M.R., Höchenberger R., Sogo H., Kastman E., Lindeløv J. PsychoPy2: experiments in behavior made easy // *Behavior Research Methods*. 2019. № 51 (1). doi: 10.3758/s13428-018- 01193-y

References

1. Mirosnichenko, E.A. (2011) K formal'nomu opredeleniyu ponyatiya "baza dannykh" [On the formal definition of the concept of "database"]. *Problemy informatiki*. 2 (10). pp. 83–87.
2. Soares, A.P. et al. (2016) The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*. 49 (3). pp. 1065–1081. doi: 10.3758/s13428-016-0767-4
3. Bonin, P., Méot, A. & Bugaiska, A. (2018) Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times. *Behavior Research Methods*. 50 (6). doi: 10.3758/s13428-018-1014-y
4. Scott, G.G. et al. (2018) The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*. 51 (1-2). doi: 10.3758/s13428-018-1099-3
5. Speed, L.J. & Brybaert, M. (2022) Dutch sensory modality norms. *Behavior Research Methods*. 54. pp. 1306–1318. doi: 10.3758/s13428-021-01656-9
6. 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/s10936-0179548-1
7. Vladimirova, V.E. et al. (2022) *Certificate of state registration of the database No. 2022622657 Russian Federation. Psycholinguistic database RuTurkPsychLing: assessments of the words of the Russian, Tatar, Khakass languages in terms of "familiarity", "temperature", "location in space", "size", "emotionality" and "manipulation"*. No. 2022622423. Application of October 10, 2022. Published on October 27, 2022. Applicant: Federal State Autonomous Educational Institution of Higher Education. National Research Tomsk State University. (In Russian).
8. 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
9. Rezanova, Z.I., Vladimirova, V.E. & Mashanlo, T.E. (2022) Psikholingvisticheskaya baza dannykh TurkWordPerception kak leksikograficheskii istochnik (otsenki vklada modal'nostey vospriyatiya v semantiku) [Psycholinguistic database TurkWordPerception as a lexicographic source (assessing the contribution of perception modalities to semantics)]. *Voprosy leksikografii – Russian Journal of Lexicography*. 26. pp. 96–114. doi: 10.17223/22274200/26/5
10. Wilson, M. (2002) Six views of embodied cognition. *Psychonomic Bulletin Review*. 4 (9). pp. 625–636.

11. Barsalou, L.W. et al. (2003) Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*. 2 (7). pp. 84–91. doi: 10.1016/S13646613(02)000293
12. Dudschig, C. et al. (2013) Reading “sun” and looking up: The influence of language on saccadic eye movements in the vertical dimension. *PLoS ONE*. 8 (2) e56872. doi: 10.1371/journal.pone.0056872
13. Estes, Z., Verges, M. & Barsalou, L.W. (2008) Head up, foot down object words orient attention to the objects’ typical location. *Psychological Science*. 19 (2). pp. 93–97. doi: 10.1111/j.1467-9280.2008.02051.x
14. Cowey, A., Small, M. & Ellis, S. (1994) Left visuo-spatial neglect can be worse in far than in near space. *Neuropsychologia*. 32(9). pp. 1059–1066. doi: 10.1016/0028-3932(94)90152-X
15. Winter, B. & Bergen, B. (2012) Language comprehenders represent object distance both visually and auditorily. *Language and Cognition*. 4 (1). pp. 1–16.
16. Campanella, F. & Shallice, T. (2011) Manipulability and object recognition: is manipulability a semantic feature? *Experimental Brain Research*. 3 (208). pp. 369–383. doi:10.1007/s00221-010-2489-7
17. Sapir, E. (1929) A study in phonetic symbolism. *Journal of Experimental Psychology*. 12. pp. 225–239.
18. Ho, H-N. et al. (2014) Colour-Temperature Correspondences: When Reactions to Thermal Stimuli Are Influenced by Colour. *PLoS ONE*. 9 (3). e91854. doi: 10.1371/journal.pone.0091854
19. Larsson, M. et al. (2014) Olfactory LOVER: Behavioral and neural correlates of autobiographical odor memory. *Frontiers in Psychology*. 5. pp. 1–5. doi: 10.3389/fpsyg.2014.00312
20. Yeshurun, Y. & Sobel, N. (2010) An odor is not worth a thousand words: From multidimensional odors to unidimensional odor objects. *Annual Review of Psychology*. 61. pp. 219–241. doi: 10.1146/annurev.psych.60.110707.163639
21. Winter, B. (2016) Taste and smell words form an affectively loaded part of the English lexicon. *Language, Cognition and Neuroscience*. 31 (8). pp. 975–988.
22. Peirce, J.W. et al. (2019) PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. 51 (1). doi: 10.3758/s13428-018- 01193-y

Сведения об авторах:

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

Коршунова Ирина Сергеевна – аспирант, младший научный сотрудник лаборатории лингвистической антропологии Томского государственного университета (Томск, Россия). E-mail: korshunova-61818@mail.ru

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

Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

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

Irina S. Korshunova, postgraduate student, junior researcher, Linguistic Anthropology Laboratory, Tomsk State University (Tomsk, Russian Federation). E-mail: korshunova-61818@mail.ru

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

The authors declare no conflicts of interests.

*Статья поступила в редакцию 20.04.2023;
одобрена после рецензирования 04.06.2023; принята к публикации 27.06.2023.*

*The article was submitted 20.04.2023;
approved after reviewing 04.06.2023; accepted for publication 27.06.2023.*