



Родные языки и культуры в современном изменяющемся мире

Электронный сетевой научный журнал

Native Languages
and Cultures in the Modern
Changing World

НОВЫЕ
ИССЛЕДОВАНИЯ
ПО РОДНЫМ
ЯЗЫКАМ
6-30

КОРПУСЫ И
БАЗЫ ДАННЫХ
НА РОДНЫХ
ЯЗЫКАХ.
КОРПУСНЫЕ
ИССЛЕДОВАНИЯ
31-84

НАУЧНАЯ ЖИЗНЬ

85-90

№ 1
2022

Хомушку Ольга Матпаевна, д. филос. н., ректор Тувинского государственного университета.

Абдурахмонова Нилуфар Зайнобиддин кизи, PhD, доцент кафедры прикладной лингвистики и лингводидактики Ташкентского государственного университета узбекского языка и литературы им. Алишера Навои, г. Ташкент, Узбекистан.

Бавуу-Сюрюн Мира Викторовна, д. филос. н., профессор кафедры тувинской филологии и общего языкознания Тувинского государственного университета.

Биткеева Айса Николаевна, д. филос. н., ведущий научный сотрудник научно-исследовательского центра по национально-языковым отношениям Института языкознания РАН.

Гатиатуллин Айрат Рафизович, к. техн. н., ведущий научный сотрудник Института прикладной семиотики АН РТ.

Дыбо Анна Владимировна, д. филос. н., профессор, член-корреспондент РАН, заведующий отделом урало-алтайских языков Института языкознания РАН.

Ергеш Бану Жантуганкызы, PhD, начальник отдела дистанционного обучения Евразийского национального университета им. Л.Н.Гумилева, г. Нур-Султан, Казахстан.

Ефремов Николай Николаевич, д. филос. н., главный научный сотрудник отдела якутского языка Института гуманитарных исследований и проблем малочисленных народов Севера СО РАН.

Ламажаа Чимиза Кудер-ооловна, д. филос. н., заместитель директора по научной работе Института фундаментальных и прикладных исследований Московского гуманитарного университета.

Монгуш Евгений Докурович, к. филос. н., проректор по научной работе Тувинского государственного университета.

Невская Ирина Анатольевна, д. филос. н., главный научный сотрудник сектора языков народов Сибири Института филологии СО РАН.

Озонова Айяна Алексеевна, к. филос. н., старший научный сотрудник сектора языков народов Сибири Института филологии СО РАН.

Ооржак Байлак Чаш-ооловна, д. филос. н., ведущий научный сотрудник научно-образовательного центра «Тюркология» Тувинского государственного университета (гл. редактор).

Салыхова Зугура Идрисовна, д. филос. н., профессор, заведующий кафедрой башкирской филологии Стерлитакамского филиала Башкирского государственного университета.

Сиразитдинов Зиннур Амирович, к. филос. н., ведущий научный сотрудник, заведующий лабораторией лингвистики и информационных технологий Института истории, языка и литературы Уфимского исследовательского центра РАН.

Сулейманов Джавдет Шевкетович, д. техн. н., научный руководитель Института прикладной семиотики Академии наук Республики Татарстан.

Чебоचाкова Ирина Максимовна, к. филос. н., заведующий сектором языка Хакасского научно-исследовательского института языка, литературы и истории.

Чертыкова Мария Дмитриевна, д. филос. н., ведущий научный сотрудник Научно-исследовательского института гуманитарных исследований и Саяно-Алтайской тюркологии Хакасского государственного университета им. Н.Ф. Катанова.

Шамина Людмила Алексеевна, д. филос. н., главный научный сотрудник сектора языков народов Сибири Института филологии СО РАН.

Шарипбай Алтынбек Амирович, д. техн. н., профессор кафедры информационной безопасности Евразийского национального университета им. Л. Н. Гумилева, г. Нур-Султан, Казахстан.

Юша Жанна Монгеевна, д. филос. н., ведущий научный сотрудник сектора фольклора народов Сибири Института филологии СО РАН.

Khomushku Olga Matpaevna, Doctor of Philosophy, Rector of Tuvan State University.

Abdurakhmonova Nilufar Zainobiddin kizi, PhD, Associate Professor of the Department of Applied Linguistics and Linguodidactics of the Tashkent State University of Uzbek Language and Literature named after Alisher Navoi, Tashkent, Uzbekistan.

Bavuu-Syuryun Mira Viktorovna, Doctor of Philology, Professor of the Department of Tuvan Philology and General Linguistics of Tuvan State University.

Bitkeeva Aisa Nikolaevna, Doctor of Philology, Leading Researcher at the Research Center for National-Linguistic Relations of the Institute of Linguistics of the Russian Academy of Sciences.

Gatiatullin Airat Rafizovich, Candidate of Technical Sciences, Leading Researcher at the Institute of Applied Semiotics of the Academy of Sciences of the Republic of Tatarstan.

Dybo Anna Vladimirovna, Doctor of Philology, Professor, Corresponding Member of the Russian Academy of Sciences, Head of the Department of Ural-Altai Languages of the Institute of Linguistics of the Russian Academy of Sciences.

Ergesh Banu Zhantugankyzy, PhD, Head of Distance Learning Department of L.N.Gumilev Eurasian National University, Nur-Sultan, Kazakhstan.

Efremov Nikolay Nikolaevich, Doctor of Philology, Head of the Department of the Yakut Language of the Institute of Humanitarian Studies and Problems of Small Peoples of the North of the Siberian Branch of the Russian Academy of Sciences.

Lamazhaa Chimiza Kuder-oolovna, Doctor of Philosophy, deputy Director of Scientific Work at the Institute of Fundamental and Applied Research of the Moscow University for the Humanities.

Mongush Evgeny Dokurovich, Candidate of Philological Sciences, Vice-Rector for Scientific Work of Tuvan State University.

Nevskaya Irina Anatolyevna, Doctor of Philology, Chief Researcher at the Institute of Philology SB RAS.

Ozonova Ayyana Alekseevna, Candidate of Philology, Senior Researcher of the Siberian Languages Sector of the Institute of Philology SB RAS.

Oorzhak Baylak Chash-oolovna, Doctor of Philology, Leading researcher at the Scientific and Educational Center «Turkology» of Tuvan State University (editor-in-chief).

Salakhova Zugura Idrisovna, Doctor of Philology, Professor, Head of the Department of Bashkir Philology of the Sterlitamak branch of Bashkir State University.

Sirazitdinov Zinnur Amirovich, Candidate of Philological Sciences, Leading Researcher, Head of the Laboratory of Linguistics and Information Technologies.

Suleymanov Javdet Shevketovich, Doctor of Technical Sciences, Scientific Director of the Institute of Applied Semiotics of the Academy of Sciences of the Republic of Tatarstan.

Chebochakova Irina Maksimovna, Candidate of Philology, Head of the Language Sector of the Khakas Research Institute of Language, Literature and History.

Chertykova Maria Dmitrievna, Doctor of Philology, Leading Researcher at the Research Institute of Humanitarian Studies and Sayan-Altai Turkology of the N.F. Katanov Khakass State University.

Shamina Lyudmila Alekseevna, Doctor of Philology, Chief Researcher at the Institute of Philology SB RAS.

Sharipbai Altynbek Amirovich, Doctor of Technical Sciences, Professor of the Department of Information Security of the L. N. Gumilev Eurasian National University, Nur-Sultan, Kazakhstan.

Yusha Zhanna Mongeyevna, Doctor of Philology, Leading Researcher of the Folklore Sector of the Peoples of Siberia at the Institute of Philology SB RAS.

СЛОВО РЕКТОРА. Хомушку О. М.

СЛОВО РЕДАКТОРА. Ооржак Б. Ч.

НОВЫЕ ИССЛЕДОВАНИЯ ПО РОДНЫМ ЯЗЫКАМ

Сарбашева С.Б. Прилагательные кызыл, улан/улаан в топонимике Южной Сибири

Аннай Э.К. Лексема чүве ‘вещь, предмет, нечто’ как маркер экспрессивности в тувинском языке

Абдулманов А. А., Саматова Ж. М. Особенности развития казахской терминологии

Ооржак Б.Ч. «Отложенный» императив в тувинском языке.

КОРПУСЫ И БАЗЫ ДАННЫХ НА РОДНЫХ ЯЗЫКАХ. КОРПУСНЫЕ ИССЛЕДОВАНИЯ / Раздел содержит материалы IX Международной конференции по компьютерной обработке тюркских языков "TurkLang 2021" /

Лемская В. М. Проблемы междиалектного морфонологического варьирования при разметке в корпусе миноритарного языка (на примере чулымско-тюркского)

Норманская Ю.В. Первые тюркские кириллические книги на платформе ЛингвоДок

Ондар М.В. Электронные базы данных тувинских фольклорных текстов

Турсунов М. С. Разработка программного обеспечения для текстового корпуса на узбекском языке

Каршиев А., Турсунов М., Холмухамедов Б. Uzbekcorpora.uz: программное обеспечение корпуса узбекского языка

Абжалова М. А. Синонимайзер (синонимизатор) в образовательном корпусе узбекского языка

НАУЧНАЯ ЖИЗНЬ

Тарыма А. К., Ооржак Б. Ч., Салчак А. Я., Хертек А. Б., Бавуу-Сюрюн М. В., Ондар В. С., Монгуш Ч. М. Цифровизация культурного наследия Тувы: состояние, проблемы и перспективы

3 SPEECH OF THE RECTOR. Khomushku O. M

4 SPEECH OF THE EDITOR-IN-CHIEF Oorzhak B. Ch.

NEW RESEARCH ON NATIVE LANGUAGES

Sarbasheva S. B. Adjectives kyzyl, ulan/ulaan in the toponymy of Southern Siberia

11 **Annaj E. K.** The Word чүве ‘thing, object, something’ as Marker of Expressiveness in Tuvan

19 **Abdulmanov A. A., Samatova Zh. M.** Features of the development of Kazakh terminology

25 **Oorzhak B. Ch.** The «deferred» Imperative in the Tuvan language.

31 **CORPUSES AND DATABASES IN NATIVE LANGUAGES. CORPUS STUDIES** The section contains materials of the IX International Conference on Computer Processing of Turkic Languages «TurkLang 2021»

31 **Lemskaya V. M.** Problems of Interdialect Morphological Variation when Processing Markup in the Corpus of a Minority Language (an Example of Chulym-Turkic)

43 **Normanskaja J. V.** First Turkic cyrillic books on the LingvoDoc platform.

58 **Ondar M. V.** Electronic databases of Tuvan folklore texts.

62 **Tursunov M. S.** Development of software for a Text corpus in Uzbek.

71 **Karshiev A. B., Tursunov M. S., Kholmukhamedov B.** UZBEKCORPORA.UZ: software of the Uzbek language corpus.

79 **Abjalova M. A.** Synonymizer in the educational corpus of the Uzbek language

SCIENTIFIC LIFE

85 **Taryma A. K., Oorzhak B. Ch., Salchak A. Ja., Hertek A. B., Bavuu-Sjurjun M. V., Ondar V. S., Mongush Ch. M.** Digitalization of Tuva's cultural heritage: status, problems and prospects



Дорогие друзья!

*М*ы рады представить вашему вниманию первый номер электронного сетевого научного журнала «Родные языки и культуры в современном изменяющемся мире», учредителем которого является Тувинский государственный университет. Это событие является знаковым, поскольку Указом Президента Российской Федерации В. В. Путина 2022 г. объявлен Годом культурного наследия народов России.

Исследование своеобразия языков и культур, основанных на уникальности историко-культурного наследия народов, взаимопроникновение, сплетение богатых национальных культур, традиций народов, живущих в многонациональном государстве – одна из самых актуальных задач, способствующих сохранению и развитию коренных народов.

Мы надеемся, что наш журнал станет площадкой для обсуждения актуальных вопросов сохранения и развития родных языков и культур, обмена опытом по цифровизации культурного наследия, объединения и распространения новых знаний в гуманитарной области!

*Ректор Тувинского государственного университета,
доктор философских наук О.М. Хомушку*

Уважаемые читатели и авторы!

Основными целями электронного сетевого научного журнала «Родные языки и культуры в современном изменяющемся мире» являются: привлечение внимания научного сообщества к актуальным проблемам исследования и сохранения родных языков и культур; обмен информацией о фундаментальных и прикладных научных исследованиях в данной области; популяризация родных языков и культур; освещение вопросов преподавания родных языков и культур; сохранение языкового и этнокультурного разнообразия.

В журнале планируется обсуждение актуальных проблем языков и культур коренных народов: новые исследования, создание корпусов и баз данных, электронных ресурсов по родным языкам и культурам; преподавание и обучение языкам коренных народов; корпусные исследования; документирование нематериального культурного наследия народов, их языков и диалектов.

Электронный сетевой научный журнал «Родные языки и культуры в современном изменяющемся мире» отражает результаты оригинальных исследований, освещение научной жизни, представляет сведения об исследователях России и зарубежья, занимающихся проблемами родных языков и культур коренных народов мира.

Журнал публикует статьи на русском и английском языках. Языки материалов: русский, английский, тувинский и др. родные языки.

Сетевой научный журнал выходит 4 раза в год и доступен на сайте: rodnoyuazyk.ru.

В первом номере представлены статьи, посвященные родным языкам тюркских народов. Он содержит результаты исследований по лексике и грамматике этих языков, анализа проблем создания и развития национальных корпусов.

Обсуждаются пути формирования топонимической системы в Южной Сибири. Топонимы рассматриваются с точки зрения способов включения и адаптации разноязычных элементов на данной территории, типичности/нетипичности, регулярности/нерегулярности образования того или иного географического названия. Исследователями ставятся вопросы о статусе неизученных и малоизученных грамматических показателей.

Во втором разделе представлены статьи участников IX Международной конференции по компьютерной обработке тюркских языков «TurkLang 2021» (21–23 сентября 2021 г., Кызыл). Они отражают круг вопросов, посвященных актуальным вопросам создания и развития национальных корпусов, баз данных и электронных ресурсов на тюркских языках. Освещается проблема разработки методики организации материала в корпусе бесписьменного чулымско-тюркского языка, который имеет высокую степень языкового междиалектного варьирования. Поднимаются вопросы разработки программного обеспечения для текстовых корпусов узбекского языка. Описывается практическая значимость уже разработанных языковых корпусов в образовательной деятельности и корпусных исследованиях древних и современных тюркских языков. Задачи исследования лексики и грамматики, документирования родных языков и создания корпусов для них имеют особую значимость в свете глобализации и угрозы исчезновения малых языков и культур. Эти и некоторые другие вопросы будут рассмотрены в данном выпуске.

Главный редактор, доктор филологических наук, Б. Ч. Ооржак

ПРИЛАГАТЕЛЬНЫЕ КЫЗЫЛ, УЛАН / УЛААН В ТОПОНИМИКЕ ЮЖНОЙ СИБИРИ

ADJECTIVES KYZYL, ULAN / ULAAN IN THE TOPONYMY OF SOUTHERN SIBERIA



Сарбашева Сурна Борисовна,

кандидат филологических наук, доцент, декан факультета алтаистики и тюркологии, ФГБОУ ВО «Горно-Алтайский государственный университет».

Sarbasheva Surna Borisovna,

Candidate of Philological Sciences, Associate Professor, Dean of the Faculty of Altaic Studies and Turkology, Gorno-Altai State University.

surnasarbasheva@mail.ru

В статье представлены результаты исследования путей формирования топонимической системы в Южной Сибири, способов включения и адаптации разноязычных элементов на данной территории, выявление многослойности топонимов, типичности/нетипичности, регулярности/ нерегулярности образования того или иного географического названия. Рассматриваются распространенные в Южной Сибири топонимы с компонентами со значением «красный». Это тюркская лексема кызыл и монгольская *улан/улаан*. Анализируются материалы по алтайскому, тувинскому, хакасскому, шорскому языкам. Отмечается, что характерными чертами номинации географических объектов Сибири и Монголии являются учет физико-географических признаков местности, в частности, состав почвы и горных пород. На исследуемой территории обнаруживают каштановые почвы и их цвет сыграли главную роль в наименовании природных объектов. Для рассматриваемых языков характерно при наименовании географических объектов использование апеллятивов-геоморфологических терминов, характеризующих рельеф местности, а также гидронимических апеллятивов и почвоведческих терминов, поэтому в топонимах Южной Сибири типичными являются наименования, состоящие из сочетаний прилагательных *кызыл, улан/улаан* и апеллятивов. Отмечается, что с большей частотой встречаются наименования с компонентом *кызыл*, меньше – с компонентом *улаан*. В редких случаях наблюдается многослойность топонимики: верхний слой представлен топонимом с тюркским компонентом *кызыл*, нижний – с монгольским компонентом *улан/улаан*. В работе использовались топонимические словари, исследования по геологии и краеведению, а также полевой материал. В работе применялись методы сравнения, описания и этимологического анализа.

Ключевые слова: *топонимика, этимология, номинация, мотивированность, прилагательные кызыл и улан/улаан 'красный', аппелятивы.*

The article presents the results of the study of the ways of formation of the toponymic system in Southern Siberia, the ways of inclusion and adaptation of multilingual elements in this territory, the identification of multilayered toponyms, typicality/atypicality, regularity/irregularity of the formation of a geographical name. The toponyms common in Southern Siberia with components with the meaning "red" are considered. These are the Turkic token *Kyzyl* and the Mongolian *ulan/ulaan*. The materials on the Altai, Tuvan, Khakass, and Shor languages are analyzed. It is noted that the characteristic features of the nomination of geographical objects of Siberia and Mongolia are taking into account the physical and geographical features of the area, in particular, the composition of soil and rocks. Chestnut soils are found on the studied territory and their color played a major role in the naming of natural objects. The languages under consideration are characterized by the use of appellatives when naming geographical objects-geomorphological terms characterizing the terrain, as well as hydronymic appellatives and soil science terms. Therefore, in the toponyms of Southern Siberia, names consisting of combinations of adjectives *kyzyl*, *ulan/ulaan* and appellatives are typical. It is noted that with greater frequency, there are names with the *kyzyl* component, less with the *ulaan* component. In rare cases, there is a multilayered toponymy: the upper layer is represented by a toponym with the Turkic component *Kyzyl*, the lower one with the Mongolian component *Ulan/ulaan*. Toponymic dictionaries, research on geology and local history, as well as field material were used in the work. The methods of comparison, description and etymological analysis were used in the work.

Keywords: *toponymy, etymology, nomination, motivation, adjectives kyzyl and ulan/ulan 'red', appellatives.*

Введение

Прилагательные кызыл и улан/улаан являются одними из самых регулярных, распространённых топонимов – оронимов, гидронимов, названий населенных пунктов в Южной Сибири. Целью статьи является описание топонимов с компонентами со значением «красный» в топонимической системе в Южной Сибири. Анализируется материал по алтайскому, тувинскому, хакасскому, шорскому языкам. В работе используются топонимические словари, исследования по геологии и краеведению, а также полевой материал. Используются методы сравнения, описания, этимологического анализа.

Анализ записей В. В. Радлова о географических названиях в Сибири был проведен Г. Ф. Благовой. Отмечается, что им были зафиксированы варианты изменений фонетического облика имени собственного в русском употреблении, также отмечены случаи калькирования тюркских топонимов в речи русского населения: *Кызыл Чар – Красноярск, Арасан – Тепло-Ключинская* и т.д. [Благова, 1970]. Из его же труда «Aus Siberian» О. Т. Молчанова приводит свыше 150 единиц топонимов Горного Алтая [Молчанова, 1982, с. 18-20].

Характерными чертами номинации географических объектов Сибири и Монголии являются учет физико-географических признаков местности, в частности, состав почвы и горных пород. Устанавливается, что чем древнее и прозрачнее название, тем более вероятна его связь с выраженным объектом. Например, алт. *Ак-суу* 'белая река', монг. *Цаган гол* 'белая река', *Чаган Узун* 'белая река', *Кызыл Кая* 'красная скала', букв. 'красный обрыв, яр' *Кызыл яр* 'Красноярск'.

Топонимы с компонентами *кызыл* и *улан/улаан* в Южной Сибири

Обширная зона территории России от Предкавказья до Алтая – в Восточной, Средней, Западной Сибири, Забайкалье, Монголии, Китае, Казахстане, Киргизии земля покрыта каштановой почвой, сформированной в условиях сухих степей. Такая почва образуется в пределах зоны с Севера на Юг. Выделяют три подтипа каштановых почв: темно-каштановые, каштановые и светло-каштановые. Наряду с каштановыми почвами в этой зоне широко распространены солонцы, солончаки, лугово-каштановые почвы (Ганжара, 2001). Естественно, особенности ландшафта отражены в топонимах на разных языках, иногда облачены в разные фонетические оболочки, но имеют одинаковое содержание, как например названия городов *Улан-Удэ* (Бурятия) и *Улангом* (Монголия), *Красноярск* ‘красный яр, кызыл яр’, районный центр *Улаган*, село *Кызыл-Озек* ‘красный лог, долина’, *Кызыл-Таш* ‘красный камень’, в Горном Алтае, урочище *Кызыл-Чын/Шын*, ‘красные ворота’, скала *Кызыл-Кайа* ‘красный скала’.

Красная гора в Горном Алтае, ставший местом обязательного посещения туристов, стал называться Марсом из-за терракотового цвета почвы. Существуют древние легенды и современные интерпретации о происхождении красных гор, в одной из них говорится, что когда-то верблюд и змея не поделили место обитания и подрались, в итоге вся гора была в крови, но никто не победил и заключили они договор: змея не заползает выше Акташа, на территорию современного Кош-Агачского района, верблюд не спускается ниже, и действительно, верблюд не может адаптироваться к нижним районам Горного Алтая, а в высокогорных районах не бывает змей. Основным мотивирующим признаком выступает красный цвет почвы с разными оттенками от темного до светлого: красный, румяный, рыжий, ярко-оранжевый, красновато-рыжий, кирпично-красный, бурый.

Известно, что в местах с наименованием *красный*, добывали глину для использования в быту. В XX в. в окрестностях сел *Кызыл-Озек*, Майма и города Горно-Алтайск Республики Алтай действовали кирпичные заводы. Местные жители издревле были знакомы с составом почвы, деловыми камнями, и в современных языках сохранена вся лексика. Археологи утверждают, что местными породами, служившими для производства украшений и предметов быта, были горный хрусталь, кальцит, мелкозернистый белый мрамор, микрокристаллический зеленый сланец и, возможно, другие пока не определённые породы. В керамическом производстве гунно-сарматского времени на Алтае выделяются две традиции – местная (лепная) и неместная (гончарная). В этот период населению были знакомы как местные традиции, характерные для Горного Алтая, которым присуще изготовление керамики с примесью дресвы, так и равнинные, для которых характерно использование шамота в керамическом производстве. Применение разнообразных рецептов формовочных масс было отмечено как характерная черта для поселений Северного Алтая и его предгорий, датированных первой половиной I тыс. н.э. – гунно-сарматским временем (Киреев, 2012, с. 91-108). В первой половине I тыс. н.э. лепная керамическая, возможно, как и деревянная, посуда являлась предметом широкого потребления в быту (Соенов, 2011, с. 57).

Прилагательное *кызыл* (алт., *кызыл* – красный; румяный; рыжий; ярко-оранжевый, красновато-рыжий, кирпично-красный, бурый) на территории Горного Алтая является типичным, регулярным, самым частотным в употреблении, отмечено в составе 85 географических объектов. С установлением советской власти в современной топонимии прибавилось еще одно значение красного, считает К. Б. Самтакова: *Кызыл Аркыт* букв. ‘Красный Аркыт’; *Кызыл Мааны* (колхоз) букв. ‘Красное Знамя’ (Национальная..., 2016]. К ним можно отнести и название городов *Кызыл* и *Улан-Батор*.

Топонимам с компонентом *кызыл* О.Т. Молчанова посвящает обширную статью в своем труде «Энциклопедия названий мест Горного Алтая». Она отмечает, что данные наименования географических объектов характеризуют цвет почвы и растений. В энциклопедию включаются такие названия как урочище *Кызыл*, *Кызыл-Ажу* (*Кызыл-Ашу*) букв. ‘красный перевал’, *Кызыл-Туу*, букв. ‘красная гора’, *Кызыл-Узук* букв. ‘красная седловина’, *Кызыл-Чай*, *Кызылчак*, урочище *Кызыл-Чын/Шын*, *Кызыл чоку* – букв. ‘красная вершина горы’, *Кызыл-Таш* (В. В. Радлов: *Kysyldasch/Kysyl-tasch*; *Кызылташ*, *Козыл-таш*, *Кызылдаш*), гора *Алтыгы-Кызыл-Таш*, гора *Устиги-Кызыл-Таш*, сочетание *кызыл таш*

букв. ‘красный или красно-коричневый камень’, *Кызыл-Тобрак* букв. ‘краснозем; краснозёмный’, кызыл тобё букв. ‘красная вершина, небольшое возвышение, холм’, *Кызыл-Туру* букв. ‘красное, красноваторыжее, бурое место’, гора *Кызыл-Кыр* (рус. Красногорка) букв. ‘красная гора, возвышенность, холмистая степь’, *Кызыл-Одуру* (В. В. Радлов: *Кусу/ Огуэй, Кызыл-Адуру. Кызылы Одру, Кызылы Олуру, Кызылы Орду*) букв. ‘красная тропинка, опоясывающая гору’, *Кызыл-Озёк* букв. ‘красный, красновато-рыжий, кирпично-красный, бурый лог’, река *Кызыл-Суу* букв. ‘красная вода’, *Кызыл-Тайга* букв. ‘красноваторыжая тайга, могильники *Кызыл-Джар* или *Кызыл-Жар* ‘красная надмогильная насыпь или красная надмогильная плита, гора *Кызыл-Жар-Чокусу* (*Кызыл-Джар-Чокусу*) букв. ‘с вершиной, макушкой *Кызыл-Жар*’, река *Кызыл-Жарык* (*Кызыл-Джарык, Кызыл-ярык*) букв. ‘красная теснина, ущелье’, *Кызыл-Кайа* (*Кызылгай, Кызылгай, Кызыл-Гая*) – красное, рыжее, бурое сухое безлесное место, где нет кустарников, солончак *Кызыл-Кезен* (*Кызыл-Кизень* или *Улан-Даба*) пер.-букв. ‘красная впадина, расщелина (горная)’, гора *Кызыл-Кизель-Конгур* или *Кызыл-Кожогор* букв. ‘красная гора с песком красного цвета’, река *Кызыл-Коол* (*Кызылгол*) букв. ‘красное русло реки, речная долина’, река *Кызыл-Кочко* букв. ‘красный оползень, обвал’, озеро *Кызыл-Кол* букв. ‘красное озеро *Кызыл-Кууур* букв. ‘красный, рыжий, красновато-рыжий, кирпично-красный, бурый солончак’ (Молчанова, 2018).

При наименовании мест, очевидно, что в основном использовались гидронимические апеллятивы, геоморфологические и почвоведческие термины в сочетании с прилагательным *кызыл*: перевал, гора, седловина, вершина горы, камень, краснозем, место, возвышенность, холмистая степь; тропинка, опоясывающая гору; лог, вода, река; тайга; надмогильная насыпь или надмогильная плита, теснина, ущелье, сухое безлесное место, впадина, расщелина (горная), гора с песком красного цвета, красное русло реки, оползень, озеро; солончак, – что является типичным для мест Горного Алтая и может быть отмечена как древнее, первичное наименование. Монгольское прилагательное *улан* ‘красный’ встречается в алтайских топонимах реже (*Улаган, Улан-Даба, Уландрык, Улан-Пыратты, Улан-Бура, Улан-Таш*).

А. А. Монгуш указывает на геологические характеристики мест Тувы, в наименовании которых содержится прилагательное *кызыл*. Происхождение топонимов с компонентов *кызыл* ученый связывает с красным цветом гранитов палеозойского возраста, в минеральном составе которых преобладает розовый микроклин, окисленные залежи серно-колчеданных руд, наличием *хаши* – твердого камня, халцедона, нефрита. Считается, что красный камень, красноцветных толщ и отложений камней и глины, придают воде характерный цвет. Автор пишет, что название реки на востоке Тувы *Кызыл-Хем* (букв. ‘красная река’) обусловлено тем, что в долине реки залегают четвертичные долинные базальты, покрытые коркой красноватого цвета, базальтовая лава растекалась вниз по долине реки на большое расстояние, более чем на 150 км, до поселков Эржей и Сизим. В верховьях реки *Кызыл-Даштыг* (букв. ‘с красными камнями’) вскрывают окисленные залежи серно-колчеданных руд, имеющих ярко выраженный красноватобурый цвет (Монгуш, 2016 с. 68). Река *Кызыл-Туруг-Хем* (*туруг* ‘отвесная скала, утес’) – река с крутым падением, протекающая мимо красной скалы Другие названия водных объектов: *Кызыл-Белдир* (*белдир* ‘слияние’), *Кызыл-Булун* (*булуң* ‘излучина реки), *Кызыл-Кудук* (*кудук* ‘колодец’), *Кызыл-Чул* (*чул* ‘горная речка с заболоченными берегами с красно-бурыми каменными россыпями по берегам и на дне’). Река *Кызыл-Торга* (*торга* ‘дятел’) названа, возможно, потому что по ее берегам водятся красные дятлы (Ондар, 2007 с. 281-284).

Компонент *кызыл* входит и в названия гор и горного ландшафта, окрашенных в красноватый, красно-бурый, бурый или рыжий цвет по цвету горных пород или растительности, которая их покрывает: *Кызыл-Даг* (*даг* ‘гора’), *Кызыл-Хая* (*хая* ‘скала’), *Кызыл-Тайга*, *Кызыл-Кожагар* (*кожагар* ‘холм, сопка’), *Кызыл-Тей* (*тей* ‘холм, сопка’), *Кызыл-Туруг* (*туруг* ‘отвесная скала, утес), *Кызыл-Туруг-Даг*, *Кызыл-Хавак* (*хавак* ‘пригорок, холмик, взлобок; вытянутая возвышенность’). *Кызыл-Куй* (*куй* ‘пещера’), *Кызыл-Одуруг* (*одуруг* ‘тропа, тропинка в горах’), *Кызыл-Эл* (*эл* ‘обрыв, яр; высокий крутой берег, подмываемый рекой’), *Кызыл-Чарык* (*чарык* ‘щель, трещина), *Кызыл-Шын* (*шың* ‘глубокий лог’, алт. *шың* ‘глухой, глухомань; глубокий ров между двух крутых гор’). С горы *Кызыл-Мажалык* (*мажалык* ‘холм, сопка’) раньше здесь брали глину для окрашивания верхней одежды. По цвету растительности названы: *Кызыл-Арыг* (*арыг* ‘пойменный лес’), *Кызыл-Хадын* (*хадың* ‘береза’) (Ондар, 2007, с. 281-283).

Всего в «Топонимическом словаре Тувы» Б. К. Ондара приводится 42 топонима с компонентом *кызыл* (Ондар, 2007, с. 281-287). Н. Д. Сувандии пишет, что с компонентом *кызыл* в Туве зафиксировано около 70 географических названий (2018, с. 200).

В тувинских топонимах типичными являются также наименования апеллятивов с прилагательным *кызыл*. Они, в большинстве своем, совпадают с принципом номинации в алтайском языке, также можно предположить, что наименование произошло в древности и изменений не претерпело: красная река, красные камни, красный (бурый, рыжий) лес, красное слияние рек, красный угол, излучина реки, красная гора, красная пыль, красная вершина, красный колодец, красная пещера, красная сопка, красная тропа, красная тайга, красный холм, красная скала, красная расщелина, красный лог, красный яр. Возможно, наименование *Кызыл-Торга* – это не красный дятел, а *торго/торко* ‘шёлк’, мотивирован визуальным спокойным течением или цветом (составом) воды. Наименования тувинских топонимов совпадают с названиями алтайских.

Как и в алтайских топонимах, в тувинских с меньшей частотой, встречаются наименования с компонентом *улаан* (монг.) ‘красный’. В «Топонимическом словаре Тувы» Б. К. Ондар приводит всего 4 топонима с компонентом *улаан* (Ондар, 2007, с. 403-404). Она отмечает, что в названии реки *Улаан-Быра* вторая часть гидронима этимологизируется монг. *бургас(ан)* ‘ива, верба; ивняк, тальник’, отсюда букв. ‘красный тальник, краснотал’. Автор приводит другие топонимы с компонентом *улаан*: гора *Улаан-Даг* (даг ‘гора’) букв., ‘красная гора’, *Улаан-Тайга* (тайга ‘таежное высокогорье, дремучий горный лес, чернь, тайга’), *Улаан-Эрги (Улаан-Эрег)*, где *эрег* ‘яр’; ‘букв. красный яр’ (Ондар, там же). Автор указывает на то, что местечко *Кызыл-Дыт* в долине реки Нарын старшее поколение называет *Улаан-Мот* (монг. *улаан* ‘красный’, *мот* ‘дерево’), а также что среди населения одновременно используются два топонима, называющие одно и то же место – *Кызыл-Даш* (даш ‘камень’) и *Улаан-Хаш* (монг. *хаш* ‘камень вообще, халцедон’) (Ондар, 2007, с. 282). Это свидетельствует о многослойности топонимов: на бытовавшее ранее монгольское название этой местности наложился тувинский топоним.

На территории Республики Хакасия и Кемеровской области частотность топонимов с компонентом *кызыл* не очень высока. В Хакасии отмечены реки *Кызылсуг*, *Кызылузунсу*, искусственное озеро Красное. В местах проживания шорцев, зафиксированы топонимы *Кызыл-Баш* ‘красная вершина’, *Кызыл-Гайа* ‘красная скала’, *Кызыл-суг* ‘красная река’, *Кызыл-чар гол* ‘река с красным яром’, гора *Кызыл-Топрак* ‘красная пыль’, *Кызыл-Каячак* ‘маленькая красная скала’.

Заключение

Таким образом, основным мотивирующим признаком топонимических наименований в Сибири и Монголии выступает цвет почвы и горных пород. Так, в топонимах с компонентом ‘красный’ проявились разные оттенки окраса пород: от насыщенного темного до светлого – красно-бурый, румяный, рыжий, ярко-оранжевый, красновато-рыжий, кирпично-красный. При наименовании мест, в основном, использовались апеллятивы – геоморфологические термины, гидронимические апеллятивы и почвоведческие термины, что является типичным для данной территории и может быть отмечено как древнее, первичное наименование объектов. Устанавливается, что *улан/улаан* ‘красный’ на территории Южной Сибири встречается реже, чем *кызыл*.

Литература

Благова, Г. Ф. В. В. Радлов и изучение топонимии Сибири // Развитие методов топонимических исследований. 1970. С.34–35.

Ганжара, Н. Ф. Почвоведение. Москва. 2001. URL: <http://topuch.ru/uchebniki-i-uchebnie-posobiya-dlya-studentov-visshih-uchebnih-v3/index46.html#pages> (дата обращения : 01.02.2022.)

Sarbasheva S.B. Adjectives *kyzyl* and *ulan/ulan* ‘red ... №1, 2022, 6-11 p.

Киреев, С. М. Украшения гунно-сарматского времени из Бугузуна (Горный Алтай) // Древности Сибири и Центральной Азии. Горно-Алтайск : ГАГУ. 2012. № 4(16). С. 85-93.

Молчанова, О. Т. Структурные типы тюркских топонимов Горного Алтая. Саратов, 1982.

Молчанова, О. Т. Энциклопедия названий мест Горного Алтая. Том I. – Щецин : Volumina.pl Daniel Krzanowski. 2018. 31 с.

Монгуш, А. А. О геологическом значении некоторых топонимов Тувы // Новые исследования Тувы. 2016. № 1. С. 64-74. URL: <https://nit.tuva.asia/nit/article/view/72> (дата обращения: 01.02.2022.)

Национальная идентичность в зеркале топонимии (Горный Алтай) : монография / П. В. Алексеев, К. Б. Самтакова, Т. П. Шастина и др. ; под редакцией П. В. Алексеева ; Министерство образования и науки РФ, Министерство образования и науки Республики Алтай, Горно-Алтайский государственный университет. Горно-Алтайск : РИО Горно-Алтайского университета. 2016. 149 с.

Ондар, Б. К. Топонимический словарь Тувы. 2-е изд. Кызыл : Тувинское книжное издательство. 2007. 552 с.

Соенов, В. И. Деревянная посуда из погребений могильника Курайка / В. И. Соёнов, Е. А. Штанаклова (Константинова) // Теория и практика археологических исследований. Барнаул : АлтГУ. 2011. Вып. 6. С.177–186.

Сувандии, Н. Д. Топонимы-цветообозначения в тувинском языке // Новые исследования Тувы. 2019. №4. URL: <https://nit.tuva.asia/nit/article/view/896> (дата обращения: 01.02.2022.). DOI: 10.25178/nit.2019.4.16

References

Blagova, G. F. V.V. Radlov i izuchenie toponimii Sibiri [Radlov and the study of Siberian Toponymy] // Razvitie metodov toponimicheskikh issledovanij [Development of toponymic research methods]. Moscow, 1970. pp.34-35. (in Russ.).

Ganzhara, N. F. Pochvovedenie [Soil science]. Available at: <http://topuch.ru/uchebniki-i-uchebnie-posobiya-dlya-studentov-visshih-uchebnih-v3/index46.html#pages> (accessed 01.02.2022.).

Kireev, S.M. Ukrasheniya gunno-sarmatskogo vremeni iz Buguzuna (Gornyj Altaj) [Jewelry of the Hunno-Sarmatian period from Buguzun (Gorny Altai)] / S.M. Kireev, E.A. Shtanakova (Konstantinova), Monosov V.M. // Drevnosti Sibiri i Central'noj Azii [Antiquities of Siberia and Central Asia]. Gorno-Altajsk : GAGU, 2012. № 4(16). pp.85-93. (in Russ.).

Molchanova, O. T. Strukturnye tipy tjurkskih toponimov Gornogo Altaja [Structural types of Turkic toponyms of the Altai Mountains]. Saratov, 1982. (in Russ.).

Molchanova, O. T. Enciklopedija nazvanij mest Gornogo Altaja [Encyclopedia of the names of places in the Altai Mountains]. Vol. I. Shhecín, 2018. pp.31. (in Russ.).

Mongush, A. A. O geologicheskom znachenii nekotoryh toponimov Tuvy [About the geological significance of some toponyms of Tuva] [online] // Novye issledovanija Tuvy [New studies of Tuva]. 2016, № 1. pp.64-74. Available at: <https://nit.tuva.asia/nit/article/view/72> (accessed 01.02.2022.).

Nacional'naja identichnost' v zerkale toponimii (Gornyj Altaj): monografija [National identity in the mirror of Toponymy (Gorny Altai) : monograph] / [P. V. Alekseev, K. B. Samtakova, T. P. Shastina i dr. ; pod redakciej P. V. Alekseeva]; Ministerstvo obrazovanija i nauki RF, Ministerstvo obrazovanija i nauki Respubliki Altaj, Gorno-Altajskij gosudarstvennyj universitet. Gorno-Altajsk : RIO Gorno-Altajskogo un-ta, 2016. 149 p. (in Russ.).

Ondar, B. K. Toponimicheskij slovar' Tuvy [Toponymic dictionary of Tuva]. 2nd ed. Kyzyl : Tuvinskoe knizhnoe izdatel'stvo, 2007. 552 p. (in Russ.).

Soenov, V.I. Derevjannaja posuda iz pogrebenij mogil'nika Kurajka [Wooden dishes from the burials of the Kurajka burial ground] / V.I. Sojonov, E.A. Shtanakova (Konstantinova) // Teorija i praktika arheologicheskikh issledovanij [Theory and practice of archaeological research]. Barnaul: AltGU, 2011. Vyp. 6. pp.177-186. (avt. vkl. 0,3 p.l.) (in Russ.).

Suvandii, N. D. Toponimy-cvetooboznachenija v tuvinskom jazyke [Toponyms-color designations in the Tuvan language] [online] // Novye issledovanija Tuvy. [New studies of Tuva] 2019, N 4. Available at: <https://nit.tuva.asia/nit/article/view/896> (accessed 01.02.2022.).

Для цитирования:

Сарбашева С. Б. Прилагательные кызыл, улан/улаан в топонимике Южной Сибири [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Sarbasheva S. B. Adjectives kyzyl, ulan/ulaan in the toponymy of Southern Siberia. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date:10.02.2022.

ЛЕКСЕМА ЧУВЕ 'ВЕЩЬ, ПРЕДМЕТ, НЕЧТО' КАК МАРКЕР ЭКСПРЕССИВНОСТИ В ТУВИНСКОМ ЯЗЫКЕ

THE WORD ЧУВЕ 'THING, OBJECT, SOMETHING' AS MARKER OF EXPRESSIVENES IN TUVAN

Анной Эллада Кан-ооловна, кандидат филологических наук, научный сотрудник сектора языка Тувинского института гуманитарных и прикладных социально-экономических исследований при Правительстве Республики Тыва.

Annai Ellada Kan-oolovna, Candidate of Philological Sciences, Researcher at the Language Sector of the Tuva Institute for Humanities and Applied Socio-Economic Research under the Government of the Republic of Tyva.

eannaj@mail.ru

ORCID 0000-0001-7892-0454



В статье рассматривается лексема *чүве* 'вещь, предмет, нечто', функционирующая в тувинском языке в качестве маркера экспрессивности. Определяется лексико-семантический подход в изучении вопроса экспрессивности в лексикологии тувинского языка. Отмечается многофункциональность и многозначность данного слова, также говорится о её амбивалентных семантических свойствах в разных речевых дискурсах. Наблюдается особенность данной лексемы при переводе на русский язык, так для выражения положительной коннотации автором предлагается перевод лексемы *чүве* 'вещь, предмет, нечто' междометиями положительных эмоций, а при реализации отрицательной коннотации использование семантического эквивалента в русском языке, демонстрируемое материалом статьи.

Ключевые слова: *экспрессия, экспрессивность, экспрессивное значение, семантика, словосочетания, словообразование, дискурс, тувинский язык.*

In this article, we analyze the word *чүве* 'thing, object, something' which functions as a marker of expressiveness in the Tuvan languages. We define the lexical-semantic approach towards expressiveness in Tuvan lexicology. The word 'chuve' is shown to be polysemantic and polyfunctional, showing ambivalent semantic traits in various types of discourse. Various nuances of translation into Russian are also observed: for example, we suggest that 'chuve' may be translated as a positive interjection in positive contexts, while in negative contexts, semantic equivalents demonstrated in the article may be used.

Keywords: *expression, expressiveness, expressive meaning, semantics, word combinations, word formation, discourse, Tuvan language.*

Введение

Экспрессия – выразительность, т.е. способность языка и его отдельных компонентов усиливать проявления чувств, настроений говорящих специальными средствами. Набор эмоций разнообразен и различен по интенсивности своего проявления. Чем интенсивнее эмоции, тем экспрессивнее их выражение в языке, средствами выражения эмоций выступают различные экспрессивные слова и словосочетания.

Существуют различные аспекты рассмотрения экспрессивности: семантический, лингвостилистический, коммуникативный, лингвокультурологический, лексикографический, синтаксический, социолингвистический, психолингвистический, и др.

Экспрессивность тувинского языка рассматривается в коммуникативном аспекте в работе В. Войкова (Voikov, 2014). В рамках семантического подхода отдельное место имеет лексико-семантический аспект, который определяет экспрессивность как «свойство лексической единицы, связанное с ее способностью актуализировать качественно-количественную характеристику реального предмета, выделить особенность данного предмета на фоне одноименных с ним предметов и выразить эмоциональную оценку предмета речи, данную ему от лица, говорящего» (Лукьянова 1986, с. 43).

Экспрессивное лексическое значение – это системное значение лексической единицы. В лексическом значении экспрессивного слова, помимо денотативно-сигнификативного компонента, присутствуют коннотативные компоненты интенсивности, эмотивности и эмотивной оценки, а образность рассматривается как самостоятельный, целостный, неделимый компонент, равно связанный как с коннотативным, так и с денотативно-сигнификативным содержанием экспрессивной лексической единицы (Лукьянова 2015, с. 189–198).

«Образно-характеризующее значение, которое формируется на базе основного значения, но отличается от него образностью выражаемого таким способом представления об объекте, действия или состояния» (Кокосникова, 2004, с. 25). Такая интерпретация вопроса представляет собой семантический подход в изучении экспрессивного фонда языка.

В статье рассматриваются семантика и функции особого маркера экспрессивности *чүве* (от сущ. *чүве* ‘вещь, предмет, нечто’) в тувинском языке. Избирается лексико-семантический подход в анализе значения данной лексемы. Материалом анализа послужили примеры, полученные автором непосредственно от информантов, и их текстов художественной литературы на тувинском языке.

Экспрессивная функция лексемы *чүве* ‘вещь, предмет, нечто’

Тувинский язык, являясь агглютинативным языком, имеет развитый аналитический строй, поэтому экспрессивность часто выражается посредством качественно-оценочных определительных словосочетаний. Лексема *чүве* ‘вещь, предмет, нечто’, употребляемая в составе словосочетаний для номинации людей, выступает в качестве маркера экспрессивности в тувинском языке. Лексема *чүве* ‘вещь, предмет, нечто’ в составе словосочетания актуализирует коннотативные значения оценочности, интенсивности и образности.

Рассматриваемая лексема имеет богатую внутрисловную и межсловную парадигмы в лексической системе тувинского языка. Неоценимый вклад в изучение лексической системы тувинского языка был сделан Б. И. Татаринцевым, особую теоретическую ценность в лексикологии тувинского языка имеет монография «Смысловые связи и отношения слов в тувинском языке» (1987), в которой рассмотрены смысловые характеристики слов в плане синхронии и диахронии, изучены связи и отношения в многозначном слове, выявлены их производные значения, подробно исследована метафоризация и другие способы образования многозначного слова. Изучая смысловые связи и отношения слов в тувинском языке, он исходит от внутрисловной парадигмы, «поскольку отношения и особенно связи значений в

многозначных словах представляются более выраженными, наблюдаемыми и системными, нежели связи и отношения значений на межсловном парадигматическом уровне» (Татаринцев, 1987, с. 8).

Тувинскому языку свойственна свободная сочетаемость слов, выявлено 144 словосочетания, в особенности продуктивным выступает маркер экспрессивности: *чүве* ‘вещь, предмет, нечто’ (по отношению к человеку), передающее ярко выраженное оценочное значение.

Д. А. Монгуш в статье «О служебных функциях слов *кижи*, *улус* и *чүве* в тувинском языке» отмечал способность лексемы *чүве* ‘вещь, предмет, нечто’ выполнять разного рода функции в предложении. «Слово *чүве* ‘вещь, предмет, нечто’ может употребляться при именных сказуемых в текстах, относящихся к разговорному стилю, и выражать модальное значение категорической достоверности, основанное на очевидности сообщаемого для говорящего: *Бо туругнуң бажында оргулааш бар чүве...* (В.Көк-оол). ‘На вершине этого утеса (это точно) ровное место есть (букв. ‘ровное место есть вещь’)’» (Монгуш, 2009, с. 109).

Данная лексема в тувинском языке является крайне многофункциональной. На нашем материале мы обращаем внимание на способность *чүве* ‘вещь, предмет, нечто’ придавать как положительное, так и отрицательное экспрессивное значение в сочетании с именами существительными и прилагательными. Возможность выступать в речи с контрастной оценочной коннотацией в разных контекстах, определяющая данную лексему как амбивалентную.

Рассмотрим примеры реализации *чүве* ‘вещь, предмет, нечто’ (по отношению к человеку) с положительной коннотацией:

(1) [Ах, кужжур эрес-дидим кысты, мен база шак-ла ынчаар дыңнаан мен. Чана бергеш, ашаанга чугаалаарга]: Эр хей чүвени!...(СС, АД, 19)

эр хей чүве=ни
молодец нечто=АСС

‘[Ах, какая храбрая, смелая девушка, я тоже слышала такое про неё. Когда сообщили её мужу, он сказал:] Ох и молодчина!...’

(2) Орлан чүвемни! Билбези чок: улуг хуулгаан төлгечимни! (СС, АД, 82)

орлан чүве=м=ни бил=без=и
бойкая нечто=POSS/1Sg=ACC знать=NEG =POSS /3Sg
чок улуг хуулгаан төлгечи=м=ни
нет большой волшебный гадалка=POSS/1Sg=ACC

‘Ох и бойкая ты моя! Чего только не знает, большая волшебная гадалка моя!’

Лексема *чүве* ‘вещь, предмет, нечто’ в рассмотренных примерах показывает крайне положительное, умилительное отношение к субъекту речи.

Также положительное, располагающее отношение к субъекту речи данная лексема выражает в разговорной речи в игре или неформальном общении девочек в форме обращения друг к другу:

(3) Эйт, чүве, сайзанактап ойнаалам че! (И-1)

эйт чүве=∅ сайзанакта=п
Сор нечто=NOM играть в постройку из камней=CV1
ойна=алам че
играть=IMP1PL давай

‘Эй, давай поиграем в постройку из игрушечных камней, дорогая (подружка)’

(4) Чүве, платьеңниң чаражын, кым садып берди? (И-2)

чүве=∅ платье=ң=ниң чараж=ы=н
нечто=NOM платье=POSS/2Sg=GEN красивый=POSS=ACC
кым сад=ып бер=ди
кто купить=CV1 давать=PP

‘(Дорогая) подружка, какое красивое у тебя платье, кто купил тебе?’

коммуникативных процессах, актуализирует то или иное значение для усиления впечатления от сказанного, лучшего донесения информации о внеязыковом факте, ситуации, воздействия на адресата.

Заключение

В тувинском языке лексема *чүве* ‘вещь, предмет, нечто’ помимо других служебных функций в предложении, также имеет способность функционировать в тувинском языке в качестве маркера экспрессивности, выражая модальное, оценочное значение категорической достоверности. Возможность выступать в речи с контрастной оценочной коннотацией в разных контекстах и дискурсах определяет данную лексему также как амбивалентную. При переводе на русский язык данную лексему при выражении положительной коннотации уместно переводить междометием, выражающим положительные эмоции, восхищение. А при реализации отрицательной коннотации можно использовать семантический эквивалент в русском языке, нами предлагаемый как рус. *туп*. Способность выражать и положительную, и отрицательную оценку в семантике, также передавать образ множества в языке определяет лексему *чүве* ‘вещь, предмет, нечто’ как экспрессивный маркер в тувинском языке.

Список условных обозначений источников примеров

И-1 – Информант 1: Ондар Хургулек Борисовна (1977 г.р.).

И-2 – Информант 2: Дамба Шораана Алексеевна (1993 г.р.)

СС, АД – Степан Сарыг-оол «Алдан дургун» Кызыл, 1987.

Список условных сокращений

Adj – имя прилагательное; ACC – винительный падеж; N – имя существительное; CAUS – понудительный залог; CV1 – форма соединительного деепричастия на =n; IMP – повелительное наклонение; NOM – основной падеж; Pl – множественное число; POSS – поссессивный аффикс; PP – форма причастия прошедшего времени на =ган; Sg – единственное число.

Литература

Кокошников, О. Ю. Семантическая структура многозначного глагола в хакасском языке в сопоставлении с тюркскими языками Южной Сибири. Новосибирск : И. Д. «Сова». 2004. 144 с.

Лукьянова, Н. А. Экспрессивная лексика разговорного употребления русского языка. Новосибирск. 1986. 21 с.

Лукьянова, Н. А. Экспрессивная лексика разговорного употребления в семантическом аспекте // Вестн. Новосиб. гос. ун-та. Серия : История, филология. 2015. Т. 14, вып. 9 : Филология. Новосибирск. 2015. С.183–201.

Татаринцев, Б. И. Смысловые связи и отношения слов в тувинском языке / Б. И. Татаринцев. Москва : Наука. 1987. 197 с.

Монгуш, Д. А. Тувинский язык и письменность // Избранные труды. Кызыл. 2009. 246 с.

Voinov V. Politeness Devices in the Tuvan Language. Wiesbaden : Harrassowitz Verlag, 2014. 199 p.

References

Kokoshnikova, O. Ju. Semanticheskaja struktura mnogoznachnogo glagola v hakasskom jazyke v sopostavlenii s tjurkskimijazykami Juzhnoj Sibiri. Novosibirsk : I.D. «Sova». 2004. 144 p. (in Russ.).

Luk'janova, N. A. Jekspressivnaja leksika razgovornogo upotreblenija russkogo jazyka / N. A. Luk'janova. Novosibirsk, 1986. 227 p. (in Russ.).

Luk'janova, N. A. Jekspressivnaja leksika razgovornogo upotreblenija v semanticheskom aspekte /

N. A. Luk'janova // Vestn. Novosib.gos. un-ta. – Serija: Istorija, filologija. 2015. Vol. 14, vyp. 9: Filologija. Novosibirsk, 2015. pp.183–201. (in Russ.).

Tatarincev, B. I. Smyslovyje svjazi i otnoshenija slov v tuvinskom jazyke / B. I. Tatarincev. Moscow : Nauka, 1987. 197 p. (in Russ.).

Mongush, D. A. Tuvinskij jazyk i pis'mennost' / D. A. Mongush // Izbr. tr. Kyzyl, 2009. 246 p. (in Russ.).

Voinov V. Politeness Devices in the Tuvan Language. Wiesbaden: Harrassowitz Verlag, 2014. 199 p. (in Russ.).

Для цитирования:

Анна́й Э. К. Лексема чу́ве ‘вещь, предмет, нечто’ как маркер экспрессивности в тувинском языке [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Annaj E. K. The Word чу́ве ‘thing, object, something’ as Marker of Expressiveness in Tuvan. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 15.12.2021.

КАЗАХСКАЯ ТЕРМИНОЛОГИЯ: СТАНОВЛЕНИЕ, ИЗУЧЕНИЕ, РАЗВИТИЕ

KAZAKH TERMINOLOGY: FORMATION, STUDY, DEVELOPMENT

Абдульманов Алтынбек Абуталиевич, кандидат филологических наук, ассоциированный профессор Западно-Казахстанского инновационно-технологического университета, кафедры филологии, культуры и методики преподавания.

Abdulmanov Altynbek Abutalievich, Candidate of Philological Sciences, Associate Professor of the West Kazakhstan University of Innovation and Technology, Department of Philology, Culture and Teaching Methods,

altynbek10@mail.ru

ORCID:0000-0003-3448-7251



Саматова Жанар Мендигереевна, магистр филологии, старший преподаватель Западно-Казахстанского инновационно-технологического университета, кафедры филологии, культуры и методики преподавания.

Samatova Zhanar Mendigereevna, Master of Philology, Senior Lecturer at the West Kazakhstan University of Innovation and Technology, Department of Philology, Culture and Teaching Methods.

zhanar.samatova80@mail.ru



В статье рассматриваются вопросы терминологии казахского языка. Авторы утверждают, что широко используются традиционные способы образования терминов: относительно структуры слова – морфологический, синтаксический; относительно передачи понятийного смысла – лексико-семантический, калькирование. В современных терминах, образованных морфологическим способом, главным отличием новых терминов является активизация ранее пассивных суффиксов. В терминообразовании является актуальным использование внутренних возможностей языка и в этой связи

важно изучение терминообразующего потенциала каждого ранее неизученного или малоизученного словообразующего элемента, выявление ареала и выстраивание генетической цепочки его значений. Это и дает расширение функционирования ранее пассивных словообразующих суффиксов. Активно развиваются также способ сложения основ и синтаксический (аналитический) способ; широко лексико-семантический способ (калькирование). Казахские ученые-терминоведы рассматривают положительно принцип возобновления и возвышения потенциала лексических единиц среди архаизмов, историзмов, профессионализмов и диалектизмов. Так, у слова в зависимости от его использования в том или ином контексте может расширяться и меняться смысловая окраска, реализуется её внутренний семантический потенциал. Аббревиатуры функционируют в большей степени в письменной речи и в непосредственной коммуникации предпочтение отдается полной номинации. Изучение вопросов казахской терминологии сегодня требует необходимости его рассмотрения во взаимосвязи с общими тенденциями развития терминологии в языках мира.

Ключевые слова казахский язык, принципы терминологии, заимствование, способы терминообразования.

The article discusses the issues of terminology of the Kazakh language. The author analyzes the modern terminology of Kazakhstan. The authors claim that traditional methods of term formation are widely used: morphological, syntactic with respect to the structure of the word; lexico-semantic, calculus with respect to the transfer of conceptual meaning. In modern terms formed morphologically, the main difference between the new terms is the activation of previously passive suffixes. The authors believe that the use of internal language capabilities is relevant in term formation, and in this regard, it is important to study the term-forming potential of each previously unexplored or little-studied word-forming element, identify the area and build a genetic chain of its meanings. This gives an extension of the functioning of previously passive word-forming suffixes. The authors believe that the use of internal language capabilities is relevant in term formation, and in this regard, it is important to study the term-forming potential of each previously unexplored or little-studied word-forming element, identify the area and build a genetic chain of its meanings. This gives an extension of the functioning of previously passive word-forming suffixes. The method of adding the basics and the syntactic (analytical) method are also actively developing; the lexico-semantic method (calculus) is widely used. Kazakh terminologists consider positively the principle of renewal and elevation of the potential of lexical units among archaisms, historicisms, professionalisms and dialectisms. Thus, a word, depending on its use in a particular context, can expand and change its semantic coloring, its internal semantic potential is realized. Abbreviations function more in writing and in direct communication, preference is given to the full nomination. Studying the issues of Kazakh terminology today requires the need to consider it in conjunction with the general trends in the development of terminology in the languages of the world.

Keywords: *Kazakh language, principles of terminology, borrowing, methods of term formation.*

Введение

Термины, как и любое и другое слово в языке служат обогащению национального языка. Будучи специальной лексической единицей, выражающей целое понятие, термин служит ключом к науке, искусству, политике, технике и т.д. Вследствие формирования глобального информационного пространства, динамики общественных и экономических процессов, направленных на мировую интеграцию, удельный вес заимствованных слов в лексической системе каждого языка постоянно растет. Считается, что около 90% новых слов, которые появляются в языке, – это термины.

При естественном сосуществовании разных народов происходит заимствование понятий и их идентификация на понятийном уровне; происходит фонетическая, грамматическая трансформация заимствованного слова. В. М. Аристова рассматривает три этапа внедрения заимствований: *проникно-*

вение (заимствованное слово только попадает в лексику и адаптируется к нормам языка), усвоение (действие народной этимологии: непонятное по содержанию слово наполняется близко звучащим или похожим по значению понятием), укоренение (термин полностью адаптируется и уже активно употребляется) (Аристова, 1978, с. 29). В. Г. Муромцева определяет заимствования как оправданные и неоправданные. По мнению ученого, оправданным является заимствование, которое означает в языке то понятие, которого ранее в этом языке не существовало (Муромцева, 1985, с. 61-62). Как указывает В. Белкин, в процессе перевода терминов с английского на арабский язык переводчики используют способы транслитерации и транскрибирования (FAX, «Facsimile»), описания и расшифровки (OS «Operating system») (Белкин, 1975).

Целью статьи является рассмотреть принятые способы терминообразования в казахском языке и перспективы этого направления. Материалом анализа послужили примеры из существующих терминологических словарей казахского языка.

Истоки, принципы и способы казахского терминообразования

Вопросами образования, науки и терминологии в Казахстане стали заниматься с конца XIX века. С открытием школ для всеобщего образования и возникла потребность в учебниках, а вместе с этим и составления терминов. Первые попытки создания научной терминологии принадлежат Ы. Алтынсарину. Более глубокое и масштабное изучение терминологии происходит в начале XX века. Разработка теоретической основы казахской терминологии принадлежит А. Байтурсынову. Его работу в этом направлении продолжили последователи, Е. Омаров, Ж. Аймауытов, Х. Досмухамедов, Н. Торекулов, К. Кеменгеров, Н. Каратышканов, Т. Шонанов и др. В 1926 году на Первом Всесоюзном тюркологическом съезде в Баку А. Байтурсынов представил основные принципы терминологии казахского языка:

1. Для терминов предпочтительно перед другими иностранными словами брать казахские слова с вполне соответствующим данному понятию значением.

2. В случае отсутствия в казахском языке подобного рода слов, заимствовать таковые из родственных казахскому языку.

3. Общеупотребительные международные термины должны приниматься с изменениями в соответствии с природой казахского языка. При наличии казахских слов, способных заменить их, должны помещаться оба, чтобы право выбора предоставить обществу.

4. Все не казахские слова, не согласующиеся с природой казахского языка, должны подвергаться изменениям соответственно «казахскому говору» (Первый..., 2001, с. 271-272).

Дальнейшее развитие казахской терминологии с 40-х годов XX века принимает иное экспансивное направление – искореняются национальные особенности, идет процесс русификации национальной терминологии.

После распада СССР казахская терминология в своем развитии ищет новые тенденции и возвращает во многом то, что было под запретом. Происходит естественный возврат к терминологическим принципам, установленный еще в начале XX столетия А. Байтурсыновым.

Способы казахского терминообразования

В современной казахской терминологии широко используются традиционные способы образования терминов: относительно структуры слова (морфологический, синтаксический); относительно передачи понятийного смысла (лексико-семантический, калькирование и т.д.) В отношении динамики развития того или иного способа образования терминов в различных сферах науки проводится множество

исследовательских работ. Сегодня исследования проводятся под руководством таких ученых, как А. Кайдаров, У. Айтбаев, Ш. Курманбайулы.

Согласно исследовательским работам по анализу утвержденных терминов казгостеркома, в современных терминах, образованных морфологическим способом, главным отличием новых терминов по сравнению с терминами советского периода является активизация ранее пассивных суффиксов. Например, активное использование суффикса *-м/-ым/ім* со значением результата какого-либо процесса (*алым* ‘пошлина’, *салым* ‘вклад’, *түсім* ‘поступление’, *бітім* ‘примерение’); а также его сочетания с суффиксом *-л/-ыл/-іл*: *-л/-ыл/-іл + -м/-ым/ім*, которое выражает значение процесса (*тыңдалым* ‘аудирование’, *басылым* ‘издание’) (Шерубай, 2007).

Так, сегодня в развитии собственной казахской терминологии является актуальным использование внутренних возможностей терминообразования. Изучение терминообразующего потенциала каждого ранее неизученного или малоизученного словообразующего элемента выстраиванием генетической цепочки значений, выявлением ареала значений, для передачи которых использовался когда-либо тот или иной данный аффикс. Это дает, как было продемонстрировано выше, расширение функционирования ранее пассивных словообразующих суффиксов (*сөгіс* ‘выговор’, *ұшақ* ‘самолет’, *қаржыгер* ‘финансист’); складываются сложные словообразовательные аффиксы с двумя компонентами (*жаз-у-шы* ‘писатель’, *аял-да-ма* ‘остановка’, *жаз-ыл-ым* ‘написание’).

Одним из динамичных и активно развивающихся способов образования новых терминов являются сложение основ и синтаксический (аналитический) способы. Сложением основ образованы, например, *еңбекақы* ‘оклад’, *жанкүйер* ‘болельщик’, *мұражай* ‘музей’, *ғаламтор* ‘интернет’, *сусабын* ‘шампунь’ (су ‘вода’ + сабын ‘мыло’). Аналитико-синтетическим способом – сложением корней и прибавлением суффикса, образованы термины *сутүзілім* ‘водообразование’ (су- ‘вода’ + түз- ‘проводить’ + -іл + -ім), терминжасам ‘терминообразование’ (термин + жас- ‘делать’ + -ам), құжатайналым ‘документооборот’ (*құжат-* ‘документ’ + *айнал-* ‘оборот’ + *-ым*) и др. С развитием аналитического способа терминообразования в казахском языке за последние двадцать лет появилось множество собственно казахских терминоэлементов: *-ақы* (*жөлақы* ‘проезд’, *еңбекақы* ‘оклад’, *жалақы* ‘зарплата’, *төлемақы* ‘оплата’, *қаламақы* ‘гонарар’, *шәкіртақы* ‘стипендия’, *пәтерақы* ‘оплата за квартиру’) от *ақы-* ‘плата, оплата’; *-жол* (*әуежолы* ‘воздушный путь’, *тасжолы* ‘асфальтная дорога’, *даңғылжол* ‘широкая дорога’, *күрежол* ‘магистральная дорога’, *құсжолы* ‘млечный путь’) от *жол-* ‘дорога’; *-түзілім* (*сутүзілім* ‘водообразование’, *газтүзілім* ‘газообразование’, *тұзтүзілім* ‘солеобразование’) от *түз-* ‘проводить’; *-жай* (*әуежай* ‘аэропорт’, *саяжай* ‘дача’, *жағажай* ‘пляж’, *мекенжай* ‘адрес’, *шинажай* ‘санаторий’, *жылыжай* ‘теплица’) от *жай* ‘помещение’; *-тас* (*сутас* ‘хрусталь’, *қырышқытас* ‘щебень’, *құмтас* ‘песчаник’, *шақпақтас* ‘кремень’, *көктас* ‘лазурит’) от *тас-* ‘камень’ и др.

Более широко в казахском терминотворчестве используется лексико-семантический способ, калькирование понятий на иностранном языке. Возвышается потенциал слова в языке до термина определенной отрасли, расширяя его лексический смысл до обозначения понятия. Таким образом, происходит процесс терминализации общеупотребительного слова.

Казахские ученые-терминоведы, по сравнению с другими учеными мира, рассматривают положительно принцип возобновления и возвышения потенциала лексических единиц среди архаизмов, историзмов, профессионализмов и диалектизмов. Так, у слова в зависимости от его использования в том или ином контексте может расширяться и меняться смысловая окраска, реализуется её внутренний семантический потенциал. Примерами могут служить термины *құжат* ‘документ’, *айдауыл* ‘конвой’, *қандауыр* ‘скальпель’ и др.

Как показывает практика развития казахской терминологии, в поиске семантических соответствий заимствуемых слов возможно обращение и к древним текстам. Изучая специальную лексику прошлых лет, оказывается возможным восстановить и возобновить употребление былых терминов, например, относящихся к военной терминологии *қарауыл* ‘караул’, *жасауыл* ‘отряд’, *айдауыл* ‘конвой’. Поэтому, мы считаем, что во внимании терминоведов должны быть в первую очередь внутренние возможности языка, благодаря которым обогащается лексика современной терминологией.

Так, у слова в зависимости от его использования в том или ином контексте может расширяться и меняться смысловая окраска, реализуется её внутренний семантический потенциал. Примерами могут служить термины *құжат* ‘документ’, *айдауыл* ‘конвой’. В большей части сокращение происходит по принципу русского языка и вариант на казахском языке калькируется. В результате сокращенное слово применяется на письме, но затруднительна в коммуникации. Так, практика показывает, что в непосредственной коммуникации появляется необходимость комментирования сокращения или приводить полный вариант термина. Поэтому вопрос с собственно казахским сокращенным термином остается очень сложным.

Хотя аббревиатуры функционируют в большей степени в письменной речи и в непосредственной коммуникации предпочтение отдается полной номинации, в последние двадцать лет идет тенденция увеличения роста сокращенных терминов из английского языка. Это – термины-паразиты, вошедшие в язык не только без фонетической трансформации, но и с сохранением графики. Например, USB, CD, Mr, PP и т.д. Данные сокращения все более широко используются как в письменной, так и в устной речи, что сильно засоряет собственно национальный язык.

Любое заимствование должно быть трансформировано в соответствии с грамматическими законами языка-реципиента, семантически осмысленно поиском альтернативного понятия из народного языка. Например, см. фонетическую трансформацию слов ‘кровать’ – *кереует*, ‘почта’ – *пошта*, ‘машина’ – *машин*; семантические соответствия слов ‘мороженое’ – *балмұздақ*, ‘панель инструментов’ – *саймандар тақтасы*, ‘насос’ – *сорғы*.

Заключение

Сегодня в казахском языке наблюдается как о в другом языке мира огромный приток новых иноязычных терминов и слов. Среди ученых-филологов идет горячий спор по вопросам заимствования. Наша позиция такова: 1) в процессе обогащения национальной терминологии, в первую очередь, необходимо обращение к собственным ресурсам национального языка; 2) любое заимствование должно быть трансформировано в соответствии с грамматическими законами языка-реципиента; 3) заимствование должно быть семантически осмыслено.

Изучение вопросов казахской терминологии сегодня требует необходимости его рассмотрения во взаимосвязи с общими тенденциями развития терминологии в языках мира с учетом особенностей собственно казахского терминообразования. Поэтому сегодня очень важно осмыслить позиции казахских ученых прошлого столетия разработать новые подходы к принципам заимствования в казахский язык.

Литература

- Аристова, В. М. Англо-русские языковые контакты. Ленинград : Издательство ЛГУ, 1978. 151 с.
 Первый Всесоюзный тюркологический съезд. 26 февраля – 5 марта (Стенографический отчет). Баку : Начыл Еви. 2011. 552 с.
 Шерубай Курманбайулы. Терминтанушы куралы /учебник-справочник. Астана : Ер-Даулет. 2007. 208 с.
 Белкин, В. М. Арабская лексикология. Москва : Наука, 1975. 200 с.
 Муромцева, В. Г. Развитие лексики украинского литературного языка во второй половине XIX - начале XX в. Харьков, 1985. 152 с.

References

- Aristova V. M. Anglo-russkie jazykovye kontakty [English-Russian language contacts]. Leningrad, 1978. 151 p. (in Russ.).
 Pervyj Vsesojuznyj tjurkologicheskij sjezd. 26 fevralja–6 marta (Stenograficheskij otchet) [All-Union

Turkological Congress February 26 – March 6 (verbatim report)]. Baku: Nacil Evi, 2013. pp.583. (in Russ.).

Sherubaj Kurmanbajuly Termintanushy kuraly / Textbook-reference. Astana : Er-Daulet, 2007. 208 p. (in Kaz.).

Belkin V. M. Arabskaja leksikologija [Arabic lexicology]. Moscow : Nauka, 1975. 200 p. (in Russ.).

Muromceva V. G. Razvitie leksiki ukrainskogo literaturnogo jazyka vo vtoroj polovine XIX - nachale XX v. [The development of the vocabulary of the Ukrainian literary language in the second half of the XIX - early XX century]. Har'kov, 1985. 152 p. (in Russ.).

Для цитирования:

Абдильманов А. А., Саматова Ж. М. Особенности развития казахской терминологии [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Abdulmanov A. A., Samatova Zh. M. Features of the development of Kazakh terminology Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...).

Дата поступления/Submission date: 15.12.2021.

«ОТЛОЖЕННЫЙ» ИМПЕРАТИВ В ТУВИНСКОМ ЯЗЫКЕ

THE “DEFERRED” IMPERATIVE IN THE TUVAN LANGUAGE

Ооржак Байлак Чаш-ооловна, доктор филологических наук, руководитель научно-образовательного центра «Тюркология», Тувинский государственный университет.

Oorzhak Baylak Chash-oolovna, Doctor of Philology, Head of the Scientific and Educational Center of Turkology, Tuvan State University.

oorzhak.baylak@mail.ru
ORCID: 0000-0001-8373-9107



Императив является универсальным грамматическим значением, которая имеется практически во всех известных языках. Выражая побуждение к действию, императив относит каузируемое действие в план будущего, поэтому он имеет категориальные связи с грамматическим временем. В некоторых языках дифференцировано представлены показатели «близкого будущего времени» и «отдаленного будущего времени» («отсроченного», «отложенного», «дистантного») императива. В императивной системе тувинского языка традиционно не выделяется грамматически выраженная оппозиция во времени. учеными установлено, что дистантный императив в языках мира может приобретать модальные значения и часто «воспринимается как имеющий более вежливое значение, чем императив немедленный, поскольку последний предполагает более высокую степень контроля говорящего над адресатом и, соответственно, более высокую степень принуждения» в выполнении действия субъектом. В оппозиции с традиционным показателем императива, которая выражается «чистой» основой, как и в других тюркских языках, в тувинском языке обнаруживается аналитическая форма =ар + показатель 2-го лица ед. и мн. ч. + *көрөм/көрүңерем*, выражающая значение «отложенного или дистантного императива», осложненное оттенком вежливости. Данная аналитическая форма в настоящее время не имеет определенного статуса в грамматической системе глагола тувинского языка, отличается регулярностью использования в повседневной речи и употреблением в художественной литературе (преимущественно в диалогах). Автор считает, что эту форму можно считать аналитическим показателем «отложенного» императива в тувинском языке.

Ключевые слова: тувинский язык, императив, «отложенный» императив, настоящее-будущее времени на =ар, в сочетании с местоимениями 2-го лица сен/силер, элемент *көрөм*.

Imperative is a universal grammatical meaning which is available in almost all languages. Expressing the urge to action, the imperative relates the causable action to the plan of the future, therefore it has categorical

connections with grammatical tense. In some languages, indicators of the ‘near future tense’ and the ‘distant future tense’ (‘deferred’, ‘distant’) imperative are differentiated. In the imperative system of the Tuvan language, grammatically expressed opposition in time is traditionally not distinguished. Scientists have found that the distant imperative in the languages of the world can acquire modal meanings and is often ‘perceived as having a more polite meaning than the immediate imperative, since the latter implies a higher degree of control of the speaker over the addressee and, accordingly, a higher degree of coercion’ in the performance of the action by the subject. In opposition to the traditional indicator of the imperative, which is expressed by a ‘pure’ basis, as in other Turkic languages, the Tuvan language reveals the analytical form =ar + indicator of the 2nd person SG and PL + *көрөм/көрүңерем*, expressing the meaning of ‘deferred or distant imperative’, complicated by a shade of politeness. This analytical form currently does not have a definite status in the grammatical system of the verb of the Tuvan language, differs in the regularity of use in speech and use in fiction (mainly in dialogues). The author believes that this form can be considered an analytical indicator of the ‘deferred’ imperative in the Tuvan language.

Keywords: *Tuvan language, imperative, ‘deferred’ imperative, present-future tense on =ar, in combination with pronouns of the 2nd person сен/силер, element көрөм.*

Введение

Императив является универсальным грамматическим значением и традиционно считается вторым по важности после индикатива в грамматическом устройстве любого языка. Он является самым сложным семантически и формально неоднородным наклонением. Первая проблема состоит в обосновании отнесения императива к категории наклонения и, в связи с этим, определения его категориального значения. Следующей стоит вопрос о его парадигме, какую парадигму он имеет в том или ином языке. Третья проблема о семантических связях императива как в общей семантике модальности, так и его связях с другими грамматическими категориями глагола.

Императив, выражая побуждение к действию, относит каузируемое действие в план будущего. Поэтому в тех языках, в которых функционирует не одна парадигма императива, говорят об оппозиции «немедленного» и «отсроченного» («отложенного», «дистантного») императива или императива «близкого будущего времени» и «отдаленного будущего времени» (см. об этом, например, в (Мальчуков, Храковский 2015, с. 17; Плунгян 2011, с. 365) и др.). Отмечается, что дистантный императив может модализироваться и часто «воспринимается как имеющий более вежливое значение, чем императив немедленный, поскольку последний предполагает более высокую степень контроля говорящего над адресатом и, соответственно, более высокую степень принуждения» (Мальчуков, Храковский 2015, с. 17). И наличие оппозиции временной дистанции в императиве является не таким уж редким явлением в языках мира. Например, дистантный императив отмечается в тунгусо-манчжурских языках (см. в (Malchukov 2001)).

Целью статьи является описание одного из значений полисемантической формы настоящего-будущего времени на =ар в тувинском языке – «отложенного» императива. Материалом описания послужили примеры из художественной литературы на тувинском языке.

Выражение значения «отложенного» императива в тувинском языке

Специализированные показатели темпоральной оппозиции в императиве среди тюркских языков имеется, по-видимому, только в якутском языке. Так, в системе якутского повелительного наклонения принято дифференцировать две парадигмы императива: «императив настоящего времени» и «императив будущего времени» (Харитонов 1947, с. 186–188) или другие термины – «повелительное категорическое» и «повелительное наклонение будущего времени» (Убрятова 1962, с. 121); повелительное наклонение

«близкого будущего времени» и «отдаленного будущего времени» (Коркина 1970, с. 147–161). Так, «императив настоящего времени», «повелительное категорическое» или повелительное наклонение «близкого будущего времени» в якутском языке представлено полной парадигмой, тогда как второе – «императив будущего времени», «повелительное наклонение будущего времени» или «отдаленного будущего времени» выделяется только в отношении 2-го лица.

Императив отдаленного будущего времени в якутском языке передает побуждение к совершению действий «в отличие от близкого будущего в сравнительно-отдаленном от момента речи времени» и проявляется в двух вариантах: 1) формой на =аар (2 лицо ед. число), =ымаар (2 лицо мн. число): *Дьонгун аҕалаар* ‘Приведи (потом, после) своих родных’; 2) формой на =аарый (2 лицо ед. число), =ымаарый (2 лицо мн. число): *Кэһиигитин ытаарынытыый, сеп дуо оҕолоор?! ‘Пошлите (потом) гостинцев, ладно, ребята?!’* (Коркина 1970, с. 161–162).

В других тюркских языках, в том числе и тувинском, для выражения семантики отложенного императива нет специализированных показателей. В тувинском языке это значение регулярно передается формой настоящего-будущего /будущего времени индикатива на =ар в сочетании с местоимениями 2-го лица сен/силер. Тогда как традиционный императив 2-го лица ед. числа во всех тюркских языках исключает участия местоимения и выражается “чистой” основой глагола: *Кел=Ø!* ‘Приходи!’; *Уду=Ø!* ‘Спи!’. Во множественном числе к основе присоединяется аффикс множественного числа: *Кел=иңер!* ‘Приходите!’; *Уду=ңар!* ‘Спите!’

Проявлению императивной семантики в форме на =ар 2-го лица способствуют также окружающий контекст, интонация и речевая ситуация. Например:

(1) – *Катап бөлдүрер эвеспе, Белек! Хааштай бээр кижии бар болза, меңээ чугаалаар сен* (МЭ, КХ)

| | | | |
|-----------------|-------------------------|--------------|-------|
| катап | бөл=дүр=ер | эвеспе | Белек |
| снова | собирать.в.кучу=CAUS=PF | PTCL | Белек |
| {хаашта=й | бэ=эр} | кижи | {бар |
| {возмущаться=CV | AUX=PF} | человек | {есть |
| бол=за} | меңээ | чугаала=ар | сен |
| AUX=COND} | мне | говорить=FUT | 2SG |

‘Смотри, нужно заставить собрата, Белек! Если кто будет возмущаться, скажи’.

(2) *Чаңгыс-даа аът артырбайн киир сүрер силер. Боттарыңар паром-биле кежип кээр силер* (ЧС)

| | | | | |
|------------|-------------|--------------------|----------------|-------|
| чаңгыс=даа | аът | артыр=байн | {киир.сүр=ер} | силер |
| один=PTCL | конь | оставлять=CV/NEG | {загонять=FUT} | 2PL |
| боттарыңар | паром-биле | {кеж=ип | | |
| сами | паром=INSRT | {переправляться=CV | | |
| кээр} | силер | | | |
| AUX=FUT} | 2PL | | | |

‘Не оставляя ни одного коня загонять будете. Сами переправитесь на пароме’.

(3) *Келир чылын бо-ла черге, бо-ла хүн келир силер* (КЧ, А, 128).

| | | | | |
|-----------|---------------|----------|-----------|----------|
| келир | чылын=Ø | бо=ла | чер=ге | бо=ла |
| следующий | год=NOM | это=PTCL | место=DAT | это=PTCL |
| хүн=Ø | кел=ир | силер | | |
| день=NOM | приходить=FUT | 2PL | | |

‘В следующем году придете на это же место, в этот же день’.

К форме настоящего-будущего времени может примыкать элемент *көр(ем)* (<көр= ‘смотреть, видеть’ + частица =ем), который выступает здесь как «модальная частица вежливости», которая вносит значение настойчивой просьбы совершения (пример 4) или не совершения действия (пример 5) в данном случае в неопределенном будущем времени:

(4) Демги кижги аалга кажан кээрил, айтырып көөр сен көрем (КЛ, УХШ, 59).

демги кижги=Ø аал=га кажан
 тот человек=NOM аал=DAT когда
 кэ=эр=ил {айтыр=ып кө=өр} сен
 приходит=FUT=Q {спросить=CV AUX=FUT} 2SG
 көрем
 PTCL

‘Пожалуйста, спросишь, когда придет в наш аал тот человек, ладно’.

(5) — Шын-дыр, оол, эмчиң бижшин алгаи, чорааш кел, бажың-биле барбас сен көрем, далаш чоруур сен (ЛЧ, ЧЧ, 246).

шын=дыр оол=Ø эмчи=ниң
 правильно=PTCL парень=NOM/VOC доктор=GEN
 бижги=и=н ал=гаш {чора=аш
 запись=POSS/3=ACC брать=CV {направляться=CV
 кел=Ø} баж=ың=биле
 приходит=IMP/2SG} голова =POSS/2SG=INSTR
 бар=бас сен көр=ем далаш чору=ур
 уходит=NEG.FUT 2SG PTCL=PTCL торопливо ехать=FUT
 сен
 2SG

‘Правильно, парень, взяв запись врача, съезди, пожалуйста, не пропадай, (а) поторапливайся’.

Грамматикализация глагола зрительного восприятия широко распространенное явление в языках мира. Служебные элементы, восходящие к глаголу *көр* = ‘смотреть, видеть’ в тюркских языках участвуют в формировании модальных и немодальных коммуникативно-прагматических значений, в число которых входит и с значение вежливости часто сопровождающая формы императива. Так, в тувинском языке элемент *көрөм/көрүңерем* входит в структуру складывающейся аналитической формы =*ар* + показатель 2-го лица ед. и мн. ч. *көрөм/көрүңерем*, причем элемент *көр* = всегда принимает синтетическую частицу смягчения =*ем*.

Для сравнения: служебный элемент, восходящий к *көр* = ‘смотреть, видеть’ может выступать и без частицы смягчения в специальной аналитической «вежливой» форме, образованной при помощи деепричастия на =*н* (*бер=ип көр* = ‘дай, пожалуйста’ ~ *бер=ип көр=ем* ‘дай, пожалуйста’). Последняя – еще более «вежливая» форма (см. об этом в (Voinov, 2014, с. 123-133)). Самой высокой степенью «вежливости» обладает аналитическая форма =*н көөр сен/силер бе*, где =*н* – форма деепричастия, *көөр* < *көр* = + форма будущего времени на =*өр*, показатели лица ед. и мн. числа *сен/силер*, вопросительная частица *бе*. Последняя форма относится к императиву. Установлено, что в отложенном императиве «вежливая» форма очень высокой степени не образуется. Представим для наглядности рассмотренные формы императива и отложенного императива по степени проявления значения «вежливости» в Таблице 1.

Таблица 1. Степени проявления значения «вежливости» в императиве и «отложенном» императиве
Table 1. Degrees of manifestation of the meaning of «politeness» in the imperative and the «deferred» imperative

| Степени «вежливости» | Императив | «Отложенный» императив |
|----------------------|--|--|
| Нейтральная | бер= ‘дай’ | бээр сен / силер ‘дашь/дадите’ |
| Высокая | бер=ип көр(ем) / көрүңер(ем) ‘дай / дайте, пожалуйста’ | бээр сен / силер көрөм / көрүңерем ‘пожалуйста, потом дашь/дадите’ |
| Очень высокая | берип көөр сен/силер бе ‘не дашь/дадите ли’ | _____ |

Что касается других тюркских языков, то в шорском языке и татарском языках отмечается, что форма будущего времени на =ар во 2-м лице так же, как в тувинском языке, передает побуждение, направленное в будущее. В шорском языке «не предполагается незамедлительное выполнение действия прямым исполнителем» [Nasilov, Ishakova 2001, с. 213]:

(5) *Sen a-aŋ son-da tašta-baan süriüş odur-ar-sıŋ!*

| | | | | |
|-----|--------|----------|--------------|---------|
| Sen | a-aŋ | son-da | tašta-baan | süriüş |
| You | he-GEN | back-LOC | leave-CV/NEG | {follow |

odur-ar-sıŋ!
sit:aux.-FUT-2SG}

‘You will follow him all the time’. (‘Вы будете следовать за ним все время’ – перевод наш) (пример взят из [Nasilov, Ishakova 2001, с. 213]).

В татарском языке формы будущего времени 2-го лица ед. и мн. числа регулярно используются для выражения вежливого побуждения. Например,

(6) *Andij-mondij hel bul-sa, tiz gene min-ga, Hesен-ne džiber-er-seng!*

| | | | | | |
|--------------|------|---------|-------|------|---------|
| andij-mondij | hel | bul-sa | tiz | gene | min-ga, |
| that-this | case | be-COND | quick | PTCL | I-DAT |

Hesен-ne džiber-er-seng!

Hesен-ACC send-FUT-2SG

‘If something happens, you will immediately send Xasan to me.’ (‘Если что-то произойдет, Вы немедленно пошлете Хасана ко мне!’ – перевод наш, пример из [Nasilov, Ishakova 2001, с. 213]).

Заключение

В тувинском языке временная дистанция в императивной парадигме формально не маркируется специальными показателями. Поэтому в императиве отсутствует грамматическая оппозиция по времени. При этом значение «отложенного императива» может передаваться формой будущего времени индикатива на =ар как одно из ее значений. Регулярное сочетание формы на =ар в значении «отложенного императива» с частицей көрем может свидетельствовать о том, что в современном тувинском языке складывается аналитическая форма «отложенного императива», осложненная значением вежливости.

Список обозначений грамматических значений в глоссах

2 – лицо, ACC – винительный падеж, AUX – вспомогательный глагол, CAUS – понудительный залог, COND – условное наклонение, CV – деепричастие, DAT – дательный падеж, FUT – форма будущего времени на =ар, GEN – родительный падеж, IMP – повелительное наклонение, INSTR – инструментальный (творительный) падеж, LOC – местный падеж, NEG – отрицательная форма, CV/NEG – отрицательное деепричастие, NOM – основной падеж, PL – множественное число, POSS – посессивность, PF – причастие будущего времени на =ар, PTCL – частица, Q – вопросительная частица, SG – единственное число, VOC – звательный падеж.

Источники

КЛ, УХШ – Кенин-Лопсан М. Улуг-Хемниц шапкыны. Кызыл : Тув. книж. изд-во, 1965. 198 с.

КЧ, А – Черлиг-оол К. Аялга. Кызыл : Тув. книж. изд-во, 1990. 192 ар.

ЛЧ, ЧЧ – Чадамба Л. Чогаалдар чыындызы. Кызыл : Тув. книж. изд-во, 1978. 318 с.

Литература

Коркина, Е. И. Наклонения глагола в якутском языке. Москва : Наука. 1970. 308 с.

Мальчуков, А. А., Храковский, В. С. Наклонение во взаимодействии с другими категориями: опыт типологического обзора. // ВЯ. № 6. 2015. С. 9–32.

Плунгян, В. А. Введение в грамматическую семантику: грамматические значения и грамматические системы языков мира. Москва : Изд-во РГГУ, 2011. 669 с.

Убрятова, Е. И. Согласование в якутском языке // Исследования по синтаксису тюркских языков [Research on the syntax of Turkic languages]. Москва : Изд-во вост. лит-ры, 1962. С.101–189.

Харитонов Л. Н. Современный якутский язык. Фонетика и морфология. Якутск, 1947. Ч.1. 313 с.

Malchukov A. Imperative constructions in Even // V. S. Xrakovskij (ed.). Typology of Imperative Constructions [Lincom Studies in Theoretical Linguistics 9]. München : Lincom Europa, 2001. P. 159–180.

Nasilov D., Ishakova X., Safarov Š, Nevskaya I. Imperative sentences in Turkic language. München : Lincom Press, 2001. 550 P.

Voinov V. Politeness Devices in the Tuvan Language. Wiesbaden: Harrassowitz Verlag, 2014. 199 P.

References

Korkina E. I. Naklonenija glagola v jakutskom jazyke [Verb moods in the Yakut language]. – Moscow: Nauka, 1970. 308 p. (in Russ.).

Mal'chukov A. A., Hrakovskij V. S. Naklonenie vo vzaimodejstvii s drugimi kategorijami: opyt tipologicheskogo obzora [Mood in interaction with other categories: the experience of typological review]. // Voprosy jazykoznanija, № 6. 2015. pp. 9–32. (in Russ.).

Plungjan V. A. Vvedenie v grammaticheskiju semantiku: grammaticheskie znachenija i grammaticheskie sistemy jazykov mira [Introduction to Grammatical Semantics: grammatical meanings and grammatical systems of the languages of the world]. Moscow : Publishing house RGGU, 2011. 669 p. (in Russ.).

Ubrjatova E. I. Soglasovanie v jakutskom jazyke [Coordination in the Yakut language] // Issledovanija po sintaksisu tjurkskih jazykov [Research on the syntax of Turkic languages]. Moscow : Izdatel'stvo vostochnoy literatury, 1962. pp.101–189. (in Russ.).

Haritonov L. N. Sovremennyj jakutskij jazyk. Fonetika i morfologija [Modern Yakut language. Phonetics and morphology]. Jakutsk, 1947. Part 1. 313 p. (in Russ.).

Malchukov A. Imperative constructions in Even // V. S. Xrakovskij (ed.). Typology of Imperative Constructions [Lincom Studies in Theoretical Linguistics 9]. München: Lincom Europa, 2001. pp.159–180. (in Russ.).

Nasilov D., Ishakova X., Safarov Š, Nevskaya I. Imperative sentences in Turkic language. München : Lincom Press, 2001. 550 p.

Voinov V. Politeness Devices in the Tuvan Language. Wiesbaden: Harrassowitz Verlag, 2014. 199 p.

Для цитирования:

Ооржак Б. Ч. «Отложенный» императив в тувинском языке [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Oorzhak B. Ch. The «deferred» Imperative in the Tuvan language. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 21.02.2022.

Раздел содержит статьи участников IX Международной конференции по компьютерной обработке тюркских языков "TurkLang 2021"

УДК 371.315.7

НАУЧНАЯ СТАТЬЯ

ПРОБЛЕМЫ МЕЖДИАЛЕКТНОГО МОРФОНОЛОГИЧЕСКОГО ВАРЬИРОВАНИЯ ПРИ РАЗМЕТКЕ В КОРПУСЕ МИНОРИТАРНОГО ЯЗЫКА (НА ПРИМЕРЕ ЧУЛЫМСКО-ТЮРКСКОГО)¹

PROBLEMS OF INTERDIALECT MORPHONOLOGICAL VARIATION WHEN PROCESSING MARKUP IN THE CORPUS OF A MINORITY LANGUAGE (AN EXAMPLE OF CHULYM-TURKIC)

Лемская Валерия Михайловна, заместитель директора Института иностранных языков и международного сотрудничества, доцент кафедры перевода и переводоведения; старший научный сотрудник лаборатории лингвистической антропологии Национального исследовательского Томского государственного университета; старший научный сотрудник отдела урало-алтайских языков Института языкознания РАН.



Lemskaya Valeriya Mihaylovna, Tomsk State Pedagogical University, Deputy Director of the Institute of Foreign Languages & International Cooperation, Associate Professor of the Chair of Translation & Translation Studies, National Research Tomsk State University, Senior Researcher, Laboratory of Linguistic Anthropology; Institute of Linguistics, the Russian Academy of Sciences, Senior Researcher, Department of the Ural-Altai languages.

lemskaya@tspu.edu.ru

ORCID: 0000-0003-3644-3463

В статье рассматриваются случаи корпусной репрезентации материала бесписьменного чулымско-тюркского языка, имеющего высокую степень языкового междиалектного варьирования. Описывается состав корпуса текстов, даются краткие характеристики их состава и содержания; указывается программное обеспечение, использованное в исследованиях ранее и в настоящее время. Приводится детальная характеристика фонетического и морфонологического варьирования языковых единиц среднечулымских говоров. Перечисляются случаи как архиморфемной репрезентации в корпусе, так и случаи отображения материала одного диалекта или говора спецификой другого. Приводятся некоторые возможные способы устранения нерегулярности отображения вариативных языковых единиц в корпусе

¹ Работа проведена в рамках проекта «Языковое и этнокультурное разнообразие Южной Сибири в синхронии и диахронии: взаимодействие языков и культур» (грант Правительства РФ № 14./26.31.0014).

(архиморфемная репрезентация не только деривационных и флективных морфем, но и корней, условная орфография и пр.).

Ключевые слова: чулымско-тюркский язык, диалекты, говоры, корпус текстов, бесписьменный язык, языковое варьирование.

The article deals with cases of corpus representation of the material of the unwritten Chulym-Turkic language, which has a high degree of linguistic inter-dialect variation. The composition of the corpus of texts is described, brief characteristics of their composition and content are given; the software used in research earlier and currently is indicated. A detailed description of the phonetic and morphological variation of the linguistic units of the Srednechulym dialects is given. The cases of both archimorphemic representation in the corpus and cases of displaying the material of one dialect or dialect by the specifics of another are listed. Some possible ways to eliminate the irregularity of the display of variable language units in the corpus (archimorphemic representation of not only derivational and inflectional morphemes, but also roots, conditional spelling, etc.) are given.

Key words: Chulym-Turkic language, dialects, dialects, text corpus, unwritten language, linguistic variation.

Введение

Чулымско-тюркский язык (в настоящее время представлен только среднечулымским диалектом) – язык автохтонного тюркоязычного населения среднего течения реки Чулым, распространенный на смежной территории Томской области (тутальский говор) и Красноярского края (мелетский говор). По данным Всероссийской переписи 2010 г. чулымцев всего насчитывалось 355 человек, владеющих этническим языком – 44. По нашим полевым данным лиц, свободно владеющих родным (этническим) языком среди чулымцев насчитывается не более пяти для каждого существующего ныне говора. До нескольких десятков человек еще могут понимать язык и составлять фразы и предложения (Лемская, 2016, с. 570–571).

Составляемый нами корпус чулымско-тюркских языковых данных представляет собой крайне разрозненное и несбалансированное собрание, в которое входят следующие материалы. Словарные статьи говоров кюэрик и кецик (нижнечулымский диалект) опубликованы в «Опыте словаря тюркских наречий» В. В. Радлова (Радлов, 1888–1911), текст на кюэрик – во втором томе «Образцов народной литературы тюркских племен, живущих в Южной Сибири и Дзунгарской степи» (Радлов, 1868, с. 689–705). Два текста – Молитва Господня и Символ Веры – на нижнечулымском (ежинский говор) и среднечулымском (тутальский говор) диалектах опубликованы в Томских епархиальных ведомостях (Н. Л-в, 1886, с. 15–21).

В публикациях П. Г. Иванова (1927, 1929) содержатся лишь несколько слов на кюэрик. Четыре текста на ежинском говоре нижнечулымского диалекта опубликованы А. П. Дульзоном в одной работе (Дульзон, 1952, с. 175–177) и один текст – в другой (Дульзон, 1966, с. 466). Кроме того, в двух публикациях (Дульзон 1973; 1977) доступен сравнительный материал: из 680 предложений на нижнечулымском (ежинский говор), среднечулымском (тутальский говор), эштинско-татарском и кызыльско-хакасском диалектах. Материал отражает собранные диалектные данные по специально разработанной анкете в 917 предложений, однако около трети собранного материала (предложения 681–917), насколько нам известно, не была опубликована. Материал по мелетскому говору (все 917 предложений) собраны Р. М. Бирюкович в 1971 г. (Бирюкович, 1971, с. 1–95) в рукописной тетради полевых материалов, однако пока не был обработан и не включен в корпус.

Основной пласт среднечулымского текстового материала, собранного в XX в., собран Р. М. Бирюкович и Р. А. Бони (Печерской) в 1971 и 1973 г. (Бирюкович, 1971, Т. III, VI, XI; Бирюкович, Бони, 1971, Т. V; Бони, 1971, Т. IV, 1973, Т. VII). Некоторые из собранных текстов были опубликованы (Бирюкович, 1980, с. 72–84; 1984, с. 77–86). В 1970 г. М. А. Абдрахмановым были также опубликованы два текста на среднечулымском (тутальском говоре) (Абдрахманов, 1970, с. 58–69). Четыре текста опубликованы

С. Е. Маловым (Малов, 1909, с. 35–46). Практически полностью представленный выше языковой материал был обработан – произведена морфологическая разметка с переводом – и опубликован автором (Лемская, 2010–2020, Т. 1–7).

С 2006 г. автором статьи ведется активнейшая полевая работа по документации нижнечулымского (ежинский говор, 2007–10 гг.) и среднечулымского диалектов (тутальский, мелетский говор). Было записано ок. 80 часов видео- и 200 часов аудиоданных². Зафиксированные материалы включают списки слов, перевод предложений, отдельные фразы, тексты монологической и полилогической (общение до трех носителей одновременно) речи. Значительная часть мелетских данных обработана и опубликована в двух томах (Кондияков, Лемская, 2021, Т. I, Т. II). Кроме того, корпус составляют два больших текста переводов Евангелий, выполненных носителем тутальского говора В. М. Габовым (Евангелие от Марка – отредактированный текст опубликован в 2019 г. (Евангелие от Марка, 2019); Евангелие от Иоанна – доступен только авторский перевод).

Часть корпуса нижнечулымского и среднечулымского (мелетского) материала выложена на ресурсе «Лингводок. Лингвистические платформы»³: есть примеры размеченных текстов, синхронизированных с аудиозаписями с полксемным выравниванием, которые далее были обработаны системой и помещены в формат словаря с аудиосопровождением⁴.

Другая часть чулымско-тюркского корпуса, собранная в ходе проектов по грантам Программы документации исчезающих языков (Endangered Languages Documentation Programme, ELDP) – проект MDP 0330, депозит №0496 за 2016–19 гг. – выложена на онлайн-ресурсе – в Архиве исчезающих языков (Endangered Languages Archive, ELAR⁵).

Всего в авторском корпусе нижнечулымского текстового материала (Радлов, 1868; Томские епархиальные ведомости, 1886; Дульзон, 1952; 1966; Лемская, 2007–2010) насчитывается 5673 словоупотреблений.

В корпусе среднечулымского текстового материала (Бирюкович, 1971; 1980; 1984; Бирюкович, Бони, 1971; Бони, 1973; Абдрахманов, 1970; Малов, 1909; Кондияков, Лемская, 2021; Лемская 2006–2021) – 125 текстов общим объемом 43,5 тыс. словоупотреблений.

Целью статьи является рассмотрение проблем корпусной репрезентации материала бесписьменного чулымско-тюркского языка. Статья написана на материале создаваемого корпуса.

Проблемы междиалектного морфонологического варьирования при разметке в корпусе чулымско-тюркского языка

В начале исследований чулымско-тюркского языка в 2005 г. для создания корпуса автором была выбрана программа «Field Linguist's Toolbox» (Рис. 1). Первоначально материал для наполнения корпуса извлекался из полевых тетрадей Р. М. Бирюкович и Р. А. Бони, частично публикаций А. П. Дульзона и его карточного рукописного словаря. В программе также обрабатывались новые данные, собираемые в ходе текущих проектов по документации.

Ввиду определенных ограничений представленной программы и для решения разноплановых исследовательских задач чулымско-тюркский исследовательский корпус с 2018 г. был перенесен и расширен в программе «SIL FieldWorks. Language Explorer» (Рис. 2).

По мере наполнения корпуса стали обнаруживаться следующие проблемы, связанные с материалом и его обработкой в указанной программе. Как было отмечено выше, чулымско-тюркские говоры отличаются высокой степенью вариантивности как на фонетическом, так и морфонологическом уровнях.

² Чулымско-тюркский язык // Малые языки России. Проект Института языкознания РАН. <https://minlang.site/lang/chulymsko-tyurkskiy>. Дата обращения: 01.11.2021.

³ ЛингвоДок. Лингвистическая платформа. http://lingvodoc.ispras.ru/corpora_all. Дата обращения: 01.11.2021.

⁴ Чулымские словари // ЛингвоДок. Лингвистическая платформа. http://lingvodoc.ispras.ru/dashboard/dictionaries_all (Chulyum). Дата обращения: 01.11.2021.

⁵ Архив языков, находящихся под угрозой исчезновения (Endangered Languages Archive). <https://www.elararchive.org/>. Дата обращения: 01.11.2021.

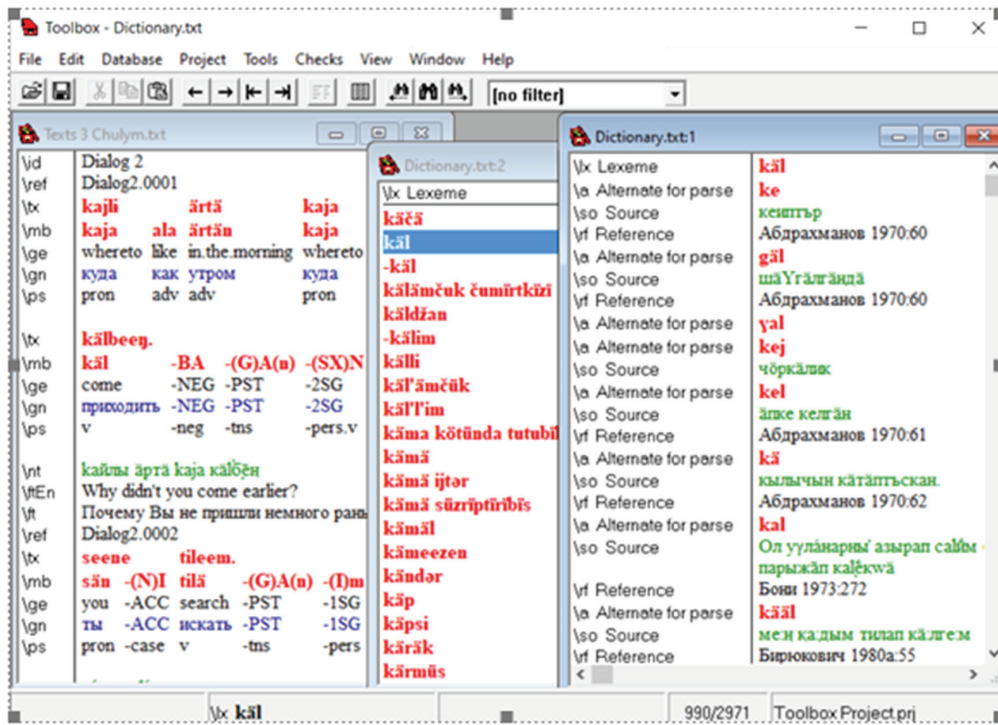


Рис. 1. Пример чулымско-тюркского корпуса в программе «Field Linguist's Toolbox».
 Fig. 1. An example of the Chulym-Turkic corpus in the “Field Linguist's Toolbox” program.

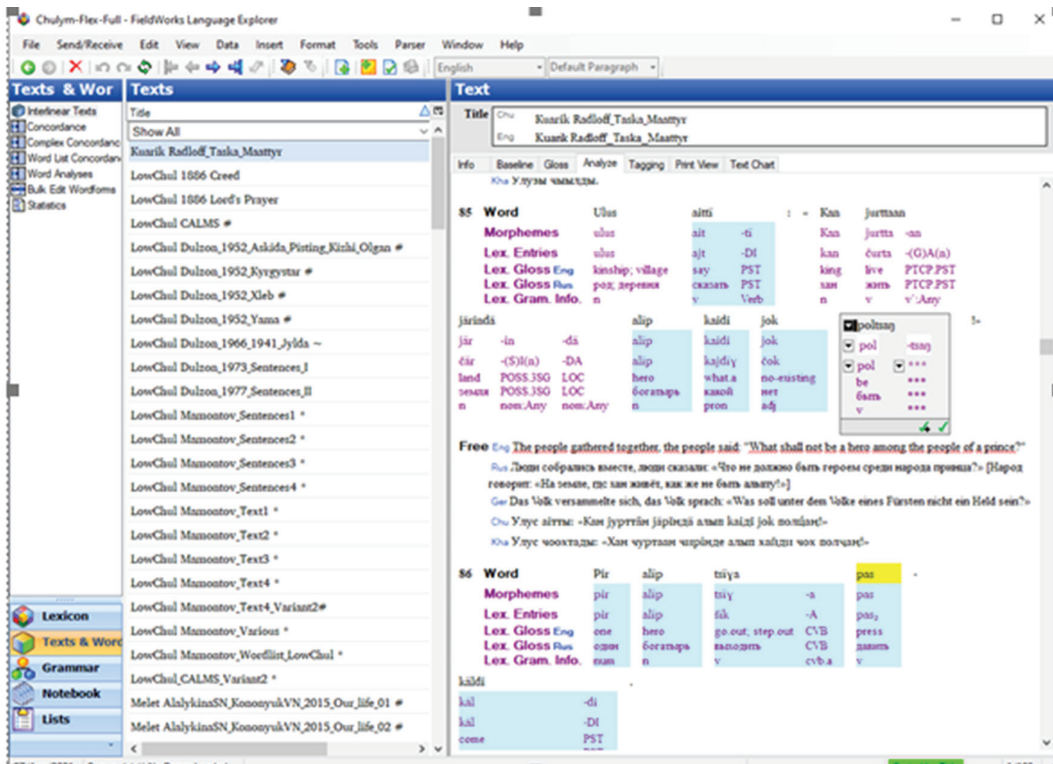


Рис. 2. Пример чулымско-тюркского корпуса в программе «SIL FieldWorks. Language Explorer».
 Fig. 2. An example of the Chulym-Turkic corpus in the “SIL FieldWorks program. Language Explorer”.

Особенно следует отметить вариативность чулымско-тюркских морфонологических форм (изменение фонетического облика морфем без изменения самих морфем на стыке частей слова), в первую очередь – аффиксов, далее – корней и составных слов.

По своим общетюркологическим типам нижнечулымский диалект относится к j-языкам, а среднечулымский – к z-языкам (Дыбо, 2020, с. 700). Для удобства далее проиллюстрируем варьирование на материале только говоров среднечулымского диалекта: тутальского и мелетского.

Фонетическое варьирование среднечулымских говоров (под которым мы понимаем алломорфизм в репрезентации форм без изменения их инвариантного состава) необходимо прежде всего отметить в следующих случаях:

1) в корнях слов:

а) варьирование согласных: т. *teate* ~ м. *εae* (<**sate*⁶) ‘волосы’; т. *af* ~ м. *as* (<**af*) ‘еда’; варьирование согласных на уровне среднечулымских говоров отражает общетюркские ряды чередования согласных фонем (Бирюкович, 1975);

б) варьирование гласных: в изолированных корнях система гласных обоих говоров, по всей видимости, не отличается. Однако, при наращивании аффиксов (у существительных это прежде всего аффиксы принадлежности 3 л., а у глаголов – одна из форм настоящего времени) корневая основа гласных меняется: в двусложных словах при условии, что последний гласный основы (а) (причем он может быть как конечным, так и прикрытым другим согласным кроме увулярного взрывного, либо велярного фрикативного), происходит его сужение в обоих говорах: в тутальском он заменяется на (з), а в мелетском – на (о). Следующие за видоизмененным корнем аффиксы строят свою гармонию согласно ассимилированному гласному: т. *partea* ‘весь, все’ > *partezzinge* (*partez-zin-ge* ‘весь-Poss.3Sg-Dat’) ‘всему’ ~ м. *partea* ‘весь, все’ > *parteozun* (*parteo-zun* ‘весь-Poss.Acc’) ‘всего’.

в) составные основы: т. *aate araazinda* (< *avate* (<**avate* ~ *ivate*) ~ *ivate* ara-(S)I(n)-DA ‘дерево середина-Poss.3Sg-Loc’) ‘в лесу’ ~ м. *azarozundun* (< *az-aro-zun-dun* < *ae* (<**avate* ~ *ivate*) ara-(S)I(n)-DIn ‘дерево середина-Poss.3Sg-AbI’) ‘из леса’; т. *aqpus* (< *aq pus* (<**a:q bu:z* ~ *tu:z*) ‘белый лед’) ‘белый’ ~ м. *aqvus* (< *aq pus* (<**a:q bu:z* ~ *tu:z*) ‘белый лед’) ‘белый’;

2) в аффиксах:

а) словообразовательные:

i) именные: т. *kyteygæte* (< **kiteig* ‘маленький’+(A)č) ‘щенок’ ~ м. *kyzyyæe* (< **kiteig* ‘маленький’+(A)č) ‘щенок’; т. *æptei* (< **æb* ~ **ib* + ČI) ‘жена’ ~ м. *æpɛy* (< **æb* ~ **ib* + ČI);

ii) глагольные: говоря о перечне вариативных элементов глагольного словообразования, следует отметить, что это скорее относится к синхронному состоянию языка, когда такими элементами могут выступать грамматикализованные лексемы и сочетания. Общетюркские глаголообразующие форманты в фонетическом варьировании скорее следуют тенденции, характерной для тюркских языков Южной Сибири в целом (отыменной глаголообразующий аффикс +LA- имеет шесть вариантов, различающихся по набору гласных – широкий (а) и узкий (е) – и согласных – сонанты (ɫ / l), (n) и (t), вариативность которых завистит от предшествующим аффиксу конечным фонемам основы: т. *taf* ‘камень-голыш’ +LA- > *tafta-* ‘бросать, кидать’ ~ м. *tas* ‘камень-голыш’ +LA- > *tasta-* ‘бросать, кидать’; т. ~ м. *raw* ‘веревка’ +LA- > *rawla-* ‘вязать’). Таким образом, синхронный словообразовательный глагольный суффикс представлен элементом Pa(r)-, восходящим к глаголу *par-* ‘уходить’ в полипредикативной конструкции, который в вариантах чулымско-тюркских говоров эксплицируется следующим образом: т. *togaſpagan* (< *togaſ-pa-gan* ‘встречать-Аух:уходить-Пtcp.Pst’) ‘встретившийся’ ~ м. *tovasvavannar* (*tovas-va-van-nar* ‘встречать-Аух:уходить-Pst-Pl’) ‘(они) встретили’.

б) словоизменительные:

i) именного словоизменения: в падежной парадигме существительных фонетическое варьирование проявляется в родительном падеже: суффикс -(N)IN имеет в мелетском варианты с чередованием узких

⁶Здесь и далее этимология лексических корней дана по материалу серии Этимологического словаря тюркских языков (1974-2000).

гласных без начального согласного (-*ij*, -*ij*, -*uj*, -*uj*); с начальным носовым сонантом (-*niŋ*, -*niŋ*, -*nuŋ*, -*nuŋ*); и с начальным глухим взрывным: -*tiŋ*, -*tiŋ*, -*tuŋ*, -*tuŋ*); в тутальском же говоре обнаруживается еще один вариант на звонкий взрывной согласный с узкими нелабиализованными гласными (-*dij*, -*dij*): т. ~ м. *øzunuj* (< *øz-y-nuj* ‘свой-Poss.3Sg-Gen’) ‘своих’; т. ~ м. *pistiŋ* (< *pis-tiŋ* ‘мы-Gen’) ‘наш’; т. *partea tæærdij* (< *partea tæærdij* ‘весь мир-Gen’) ‘всего мира’. Исходный падеж в обоих говорах представлен формами с чередованием узких гласных на глухой взрывной согласный (-*tin*, -*tin*, -*tun*, -*tün*) и звонкий взрывной согласный (-*din*, -*din*, -*dun*, -*dün*), однако периодически встречаются нерегулярные явления форм с широкими гласными (в тутальском говоре: -*dan*, в мелетском: -*dan* и -*tan*): т. *kizidin* (< *kizi-din* ‘человек-Abl’) ‘от человека’ ~ м. *kizidin* (< *kizi-din* ‘человек-Abl’) ‘от человека’; т. *zldin* (< *al-din* ‘деревня-Abl’) ‘от деревни’ ~ м. *aaldin* (< *al-din* ‘деревня-Abl’) ‘от деревни’; т. *tæilvidan* (< *tæilvi-dan* ‘лошадь-Abl’) ‘от лошади’; м. *turadan* (< *tura-dan* ‘город-Abl’) ‘из города’; м. *internettan* (< *internet-DXn* ‘Интернет-Abl’) ‘по Интернету’.

ii) глагольного словоизменения: глагольные аффиксы в среднечулымском диалекте значительно варьируют между говорами. Так, в суффиксах настоящего времени на -XdI (< *-*Atur*-) происходит сужение первого гласного аффикса: в тутальском говоре при выражении предикации на его месте возникает [з], а в мелетском – [о], а гласные последующих аффиксов следуют гармонии этого нового гласного: т. *parzdi* (< *par-zdi* < *par-XdI* ‘уходить-Prs’) ‘он(а) уходит’ ~ м. *parodu* (< *par-odu* < *par-XdI* ‘уходить-Prs’) ‘он(а) уходит’. Аффикс будущего времени на -(L)X(K) в тутальском говоре теряет начальный согласный, а узкий гласный чередуется (-*ik*, -*ik*, -*uk*, *ük*), в мелетском согласный присутствует, но имеет варианты: латеральный сонант [l] после основ, оканчивающихся на [r] или [q] (-*lik*, -*lik*, -*luk*, -*lyk*) и глухой дентальный взрывной [t] после основ, оканчивающихся на глухой согласный (-*tik*, -*tik*, -*tuk*, -*tyk*); кроме того, имеется вариант с отсутствующим начальным согласным и только с узкими нелабиализованными гласными после основ, оканчивающихся на гласный и носовой сонант [n] (-*ik*, -*ik*): т. *aparik* (< *apar-ik* ‘приносить-Fut’) ‘принесет’ ~ м. *aparlik* (< *apar-lik* ‘приносить-Fut’) ‘принесет’; т. *iteiksignar* (< *ite-ik-sij-nar* ‘пить-Fut-2Sg-Pl’) ‘(вы) будете пить’ ~ м. *ietikivs* (< *ie-tik-ivs* ‘пить-Fut-1Pl’) ‘(мы) будем пить’.

Морфонологическое варьирование (под которым мы понимаем существенное видоизменение морфем, которые на фонологическом уровне не являются очевидно относящимися к одной морфеме) среднечулымских говоров наиболее ярко представлено в словоизменительных показателях (аффиксах). Рассмотрим их типы:

а) именные: пожалуй, наиболее варьирующим аффиксом именного словоизменения является показатель орудно-совместного падежа на -(B)(I)LA(N). В тутальском говоре этот аффикс после основ, оканчивающихся на гласный, имеет варианты -*bla*, -*blæ*, а после основ, заканчивающихся на согласный – варианты -*la*, -*læ*: т. *kiziblae* (< *kizi-blae* ‘человек-Com’) ‘с человеком’; т. *kizilærlæ* (< *kizi-lærlæ* ‘человек-Pl-Com’) ‘с людьми’. В мелетском говоре имеется два типа с чередующимися широкими гласными (-*laŋ*, -*læŋ* и -*naŋ*, -*næŋ*): м. *kizinæŋ* (< *kizi-* ‘человек-Com’) ‘с человеком’; м. *otlaŋ* (< *ot-laŋ* ‘огонь-Com’) ‘(с) огнем’.

б) глагольные: в суффиксе настоящего времени на -(I)B(I)l(X) (< *-*(I)bolur-* < *-*(I)p olur-* ‘-Cvb Аух:сидеть’) имеется значительная вариация между среднечулымскими говорами. В тутальском говоре суффикс не имеет конечного гласного, общим для всех лиц является его форма -(I)bIl (-*ibil*, -*ibil*, -*ubul*, -*ubyl*), а в форме 2 л. у суффикса остается только та часть, которая восходит к деепричастию на *-*(I) p*: т. *ajdibil* (< *ajd-ibil* ‘сказать-Prs’) ‘(он(а)) рассказывает’; т. *ajdibziŋ* (< *ajd-ib-zij* ‘сказать-Prs-2Sg’) ‘(ты) рассказываешь’. В мелетском говоре этот суффикс не имеет среднего гласного, но происходит спонантизация интервокального звонкого губно-губного согласного (-*ivli*, -*ivli*, -*uvlu*, -*yvly*): м. *ajdivli* (< *ajt-ivli* ‘сказать-Prs’) ‘рассказывает’; в форме 2 л. наблюдается тот же принцип экспликации только грамматикализованного деепричастия на *-*(I)p*: м. *æurtavziŋnar* (< *æurta-v-zij-nar* < *teurta-(I)B(I)l(X)-(SX)N-LAr* ‘жить-Prs-2Sg-Pl’) ‘(вы) живете’. Аффиксы сказуемости также варьируют, восходя в разных парадигмах к разным общетюркским типам: личный аффикс глагола 1 л. ед. ч. в настоящем времени на -(X)dI в тутальском говоре имеет полную форму на -*mIn* (-*min*, -*min*, -*mun*, -*mun*), а в мелетском – его

аналог на $-m < *-(I)m$ ($-im, -im, -um, -um$): т. *ajdibilmin* ($< ajd-ibil-min$ ‘сказать-Prs-1Sg’) ‘(я) рассказываю’ ~ м. *ajdivlim* ($< ajd-ivli-m$ ‘сказать-Prs-1Sg’) ‘(я) рассказываю’.

Рассмотренные феномены варьирования среднечулымских говоров крайне затрудняют работу с языковым материалом при осуществлении разметки в корпусе. Однако благодаря технологии памяти переводов, применяемой в используемых нами программах («SIL FieldWorks. Language Explorer» и «Field Linguist’s Toolbox»), при глоссировании по введению варьируемого показателя в имеющиеся словарные карточки, либо после создания новой карточки, возникает возможность выбора необходимого значения и морфемную соотнесенность алломорфного речевого элемента в выпадающем списке (при условии, что программа сразу не распознала необходимое значение; Рис. 3).

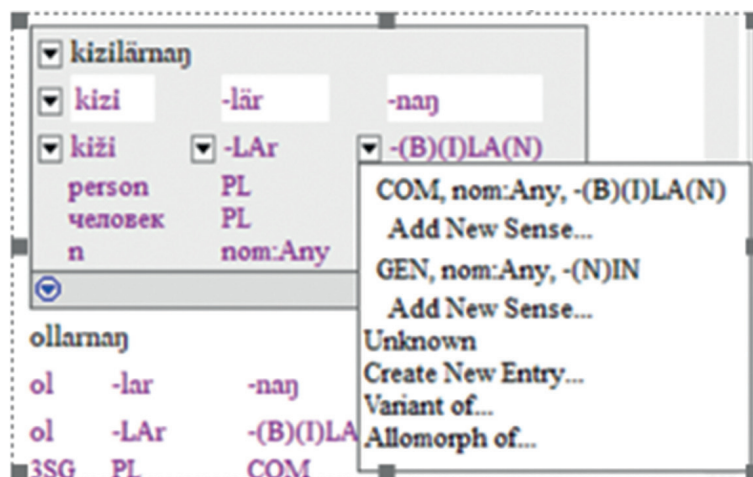


Рис. 3. Пример выпадающего списка алломорфов при глоссировании мелетского текста в программе «SIL FieldWorks. Language Explorer».

Fig. 3. An example of a drop-down list of allomorphs when glossing a Melets text in the “SIL FieldWorks. Language Explorer”.

При всем наборе возможностей, предоставляемых представленной программой, остается проблема записи заголовочного слова – исследователю необходимо выбрать тот вариант, который станет инвариантом для записи. В случае с некорневыми аффиксами мы выбрали распространенный в тюркологических исследованиях принцип архиморфемной репрезентации показателей, согласно которому в скобки берутся те фонемы, которые могут быть опущены в разных вариантах, строчными буквами записываются неварьируемые фонемы, а прописными – фонемы, варьирующие по законам чередования, характерного для каждого идиома. Так, архиморфемная запись орудно-совместного падежа $-(B)(I)LA(N)$ означает разного рода комбинации, некоторые из которых представлены на иллюстрациях



Рис. 4



Рис. 5

Рис. 4-5. Примеры глоссирования тутальского текста с архиморфемной репрезентацией аффиксов.
Fig. 4-5. Examples of glossing a Tatal text with an archimorphemic representation of affixes.

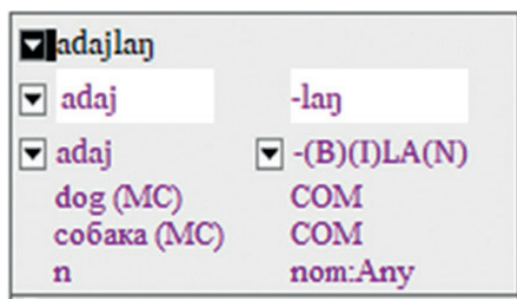


Рис. 6

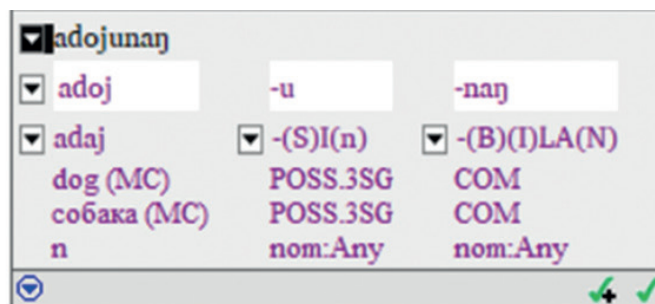


Рис. 7

Рис. 6-7. Примеры глоссирования мелетского текста с архиморфемной репрезентацией аффиксов.
 Fig. 6-7. Examples of glossing a Melets text with an archimorphemic representation of affixes.



Рис. 8

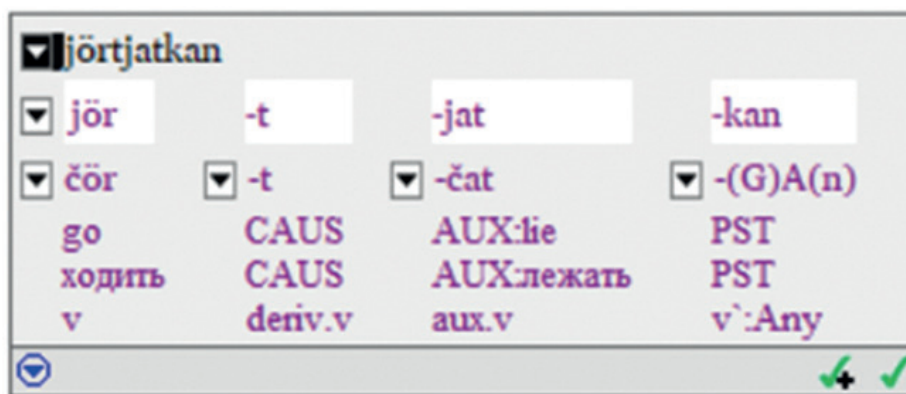


Рис. 9

Рис. 8-9. Примеры глоссирования нижнечулымского текста с репрезентацией лексических единиц тутальского говора среднечулымского диалекта в качестве заголовочных слов.

Fig. 8-9. Examples of glossing a Lower Chulym text with the representation of the lexical units of Middle Chulym dialect's Total subdialect as heading words.

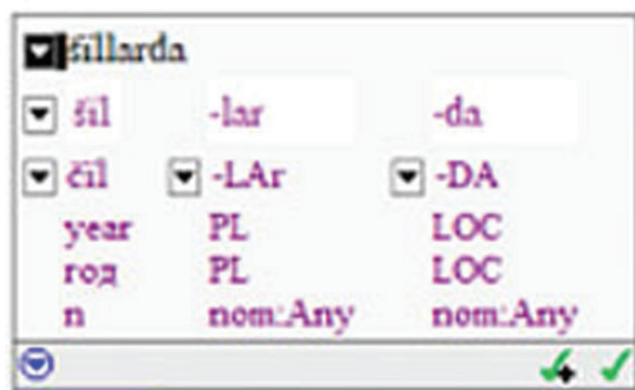


Рис. 10

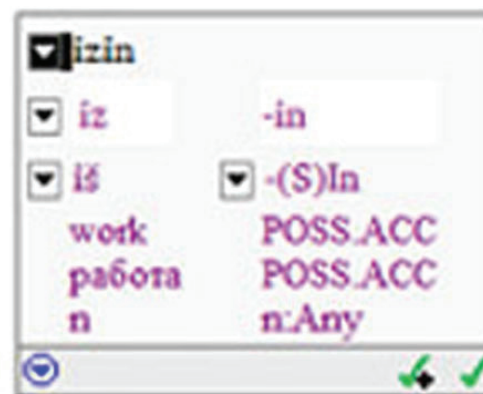


Рис. 11

Рис. 10-11. Примеры возникновения заголовочных слов, записанных на тутальском говоре, при глоссировании мелетских текстов.

Fig. 10-11. Examples of heading words written in the Total subdialect when glossing Melets texts.

Другим возможным решением проблемы репрезентации вариативных показателей в корпусе могла бы стать некая общая условная орфография – по этому принципу организована работа существующих парсеров, например, в Электронном корпусе хакасского языка⁷ и на Лингвистической платформе «ЛингвоДок»⁸.

Возможно, идеальной моделью было бы создание отдельной базы данных для каждого идиома – в случае чулымско-тюркского языка это были бы корпуса ежинского говора нижнечулымского диалекта (при этом единственный доступный текст на кюэрикском говоре, опубликованный В. В. Радловым в 1868 г., необходимо так же выносить в отдельную базу), тутальского говора среднечулымского диалекта и мелетского говора среднечулымского диалекта. Однако такая модель представляется наиболее трудоемкой и наименее оправданной – чулымско-тюркский текстовый материал представлен сравнительно небольшим объемом; кроме того, зачастую материал одного диалекта или говора позволяет установить те или иные особенности другого.

Заключение

Проблема разработки методики организации материала в корпусе бесписьменного языка, обладающего высокой степенью вариативности на уровне говоров, является актуальной задачей современной компьютерной лингвистики. Для чулымско-тюркского языка такая задача приобретает особую значимость в свете крайне малого количества активных носителей среднечулымского диалекта, утраты нижнечулымского диалекта и слабой документированности и изученности языка в целом.

Список использованных сокращений

m. – мелетский говор; т. – тутальский говор; 1Pl – форма 1 лица множественного числа; 1Sg – форма 1 лица единственного числа; 2Sg – форма 2 лица единственного числа; Abl – аблатив; Aux – вспомогательный глагол; Com – орудно-совместный падеж; Cvb – форма деепричастия; Dat – дательный падеж; Fut – форма будущего времени; Gen – родительный падеж; Loc – локатив; Pl – форма множественного числа; Poss.3Sg – притяжательная форма 3 лица единственного числа; Poss.Acc – аккузатив притяжательной формы; Prs – форма настоящего времени; Pst – форма прошедшего времени; Ptcp.Pst – форма причастия прошедшего времени.

Литература

Абдрахманов, М. А. Тексты чулымско-тюркского языка (средне-чулымский диалект) // Языки и топонимия Сибири. Т. II. Томск : Издательство Томского университета, 1970. С. 58-69.

Бирюкович, Р. М. Лексика чулымско-тюркского языка. Пособие к спецкурсу. Саратов : Издательство Саратовского университета, 1984. 88 с.

Бирюкович, Р. М. Материалы по чулымско-тюркскому языку [Полевые записи]. Т. III. Тегульдэт, 1971. 658 с.; Том VI. Пасечное, 1971. 361 с.; Том XI: нганасанский, чулымско-тюркский. Пасечное, 1971а. С.573-671.

Бирюкович, Р. М. О некоторых исторических чередованиях в чулымско-тюркском языке // Проблемы алтаистики и монголоведения. Выпуск 2. Серия лингвистики: материалы Всесоюзной конференции, Элиста, 17-19 мая, 1972. Москва : Издательство «Наука», Главная редакция восточной литературы, 1975. С. 274-279.

Бирюкович, Р. М. Чулымско-тюркские тексты // Сказки народов Сибирского Севера. Вып.3. Томск : Изд-во Том. ун-та, 1980. 191с. С. 72–84.

Бирюкович, Р. М., Бони, Р. А. Материалы по чулымско-тюркскому языку [Полевые записи]. Том V. Пасечное, 1971. 631 с.

⁷ Электронный корпус хакасского языка. <https://khakas.altai.ca.ru/>. Дата обращения: 01.11.2021.

⁸ ЛингвоДок. Лингвистическая платформа. http://lingvodoc.ispras.ru/corpora_all. Дата обращения: 01.11.2021.

- Бони, Р. А. Материалы по чулымско-тюркскому языку [Полевые записи]. Том IV. Пасечное, 1971. 629 с.; Том VII. Тегульдэт, 1973. 584 с.
- Дульзон, А. П. Материалы по диалектологии тюркских наречий Западной Сибири // Вопросы хакасской филологии: сборник научных трудов. Абакан : Хакасский научно-исследовательский институт языка, литературы и истории, 1977. С. 150-178.
- Дульзон, А. П. Материалы по диалектологии тюркских наречий Западной Сибири // Ученые записки Хакасского научно-исследовательского института языка, литературы и истории. Серия филологическая №2. Вып. XVIII. Абакан, 1973б. С. 187-207.
- Дульзон, А. П. Чулымские татары и их язык // Ученые записки Томского государственного педагогического института. Том IX. 1952. С. 76-211.
- Дульзон, А. П. Чулымско-тюркский язык // Языки народов СССР. Том II. Тюркские языки. М.: Издательство «Наука», 1966. С. 446-466.
- Дыбо, А. В., Абубакирова, Л. Ф., Айбазова, З. К., Зимин, М. М., Коровина Е. В., Мальцева В. С., Мудрак О. А., Ренковская Е. А., Савельев А. В., Хисамов О. Р., Шарова В., Шеймович А. В. Новые результаты в генеалогической классификации тюркских диалектов («случаи с аффрикатами») // *Oriental Studies*. 2020. Vol. 13. Is. 3. pp. 696-713.
- Евангелие от Марка на чулымском языке с параллельным русским Синодальным переводом. Томск : Изд-во «Иван Федоров». 132 с.
- Иванов, П. Г. Сибирские турки и их наречия. Из курса лекций, читанных в туземном краеведческом кружке при Томском государственном университете туземцам-студентам Томских ВУЗ в 1925-26 акад. году. Томск, 1927. 31 с.
- Иванов, П. Г. Татары среднего течения реки Чулыма // Труды Томского областного музея. Томск, 1929. Т. 2. С. 87-91.
- Кондияков, А. Ф., Лемская В. М. Чулымский язык д. Пасечное Тюхтетского района Красноярского края. 2007–2021 гг.: монография. Т. I: словарь, формы слов и грамматические примеры. Томск: изд-во Том. ун-та, 2021, 204 с.; Т. II: Тексты с переводом и анализом. Томск: Изд-во Том. ун-та, 2021. 246 с.
- Лемская, В. М. Материалы этнолингвистических экспедиций кафедры языков народов Сибири Томского государственного педагогического университета к чулымским тюркам д. Кафтанчиково Томского района Томской области Российской Федерации: нижнечулымский диалект [Полевые записи: электронный ресурс], 2007-2010.
- Лемская, В. М. Материалы этнолингвистических экспедиций кафедры языков народов Сибири Томского государственного педагогического университета к чулымским тюркам Тегульдетского района Томской области и Тюхтетского района Красноярского края Российской Федерации: среднечулымский диалект [Полевые записи: электронный ресурс], 2006-2021.
- Лемская, В. М. Чулымско-тюркские тексты // Сборник аннотированных фольклорных и бытовых текстов обско-енисейского языкового ареала. Коллективная монография. Т. I, Томск: «Ветер», 2010, с. 263-314; Т. II, Томск: Аграф-Пресс, 2012, с. 184-237; Т. III, Томск: «Вайар», 2013, с. 295-345; Т. IV, Томск: ТМЛ-Пресс; Вайар, 2015, с. 217-291; Т. V, Томск: ТМЛ-Пресс; Вайар, 2017, с. 103-177; Т. VI, Томск: Аграф-Пресс; Вайар, 2020, с. 233-276; Т. VII, Томск: Аграф-Пресс; Вайар, 2020. С. 102-127.
- Лемская, В. М. Чулымско-тюркский (чулымский) язык // Колл. монография «Язык и общество. Энциклопедия». Отв. ред. В. Ю. Михальченко. Москва : «Азбуковник», 2016. С. 569-571.
- Малов, С. Е. Отчет о командировке студента Восточного факультета Сергея Ефимовича Малова [Текст] / С. Е. Малов // Известия Русского комитета для изучения Средней и Восточной Азии в историческом, археологическом и антропологическом отношениях. Санкт-Петербург, 1909. №9. С.35-46.
- Н. Л-въ. Поѣздка Преосвященнаго Макарія, Епископа Бійскаго, къ Чулымскимъ инородцамъ Маріинскаго округа // Томскія епархіальныя вѣдомости. Томск, 1886. С. 15-21.
- Радлов, В. В. Образцы народной литературы тюркских племен, живущих в Южной Сибири и Дзунгарской степи. Часть II. Поднаречия Абаканские (сагайское, койбальское, качинское), кызыльское и чулымское (кюэрик). Санкт-Петербург, 1868. 420 с.
- Радлов В. В. Опыт словаря тюркских наречий. Т. I. Гласные. СПб, 1893. 1023 с.; Т. II. СПб, 1899.

971 с.; Т. III. Санкт-Петербург, 1905. 1200 с.; Т. IV. Санкт-Петербург, 1911. 1115 с.

Этимологический словарь тюркских языков: общетюркские и межтюркские лексические основы. Москва : Наука, 1974-2000.

References

Abdrakhmanov M. A. Teksty chulymsko-tyurkskogo yazyka (sredne-chulymskiy dialekt) [Texts of the Chulym-Turkic language] // Yazyki i toponimiya Sibiri. T. II. Tomsk: Izdatel'stvo Tomskogo universiteta, 1970. pp.58-69. (in Russ.).

Biryukovich R. M. Leksika chulymsko-tyurkskogo yazyka. Posobie k spetskursu [Vocabulary of the Chulym-Turkic language. The manual for the special course]. Saratov : Izdatel'stvo Saratovskogo universiteta, 1984. 88 p. (in Russ.).

Biryukovich R. M. Materialy po chulymsko-tyurkskomu yazyku [Polevye zapisi] [Materials on the Chulym-Turkic language [Field records]. T. III. Tegul'det, 1971. 658 s.; Tom VI. Pasechnoe, 1971. 361 s.; Tom XI: nganasan, chulym-Turkic. Pasechnoe, 1971a. pp.573-671. (in Russ.).

Biryukovich R. M. O nekotorykh istoricheskikh cheredovaniyakh v chulymsko-tyurkskom yazyke [About some historical alternations in the Chulym-Turkic language] // Problemy altaistiki i mongolovedeniya. Vypusk 2. Seriya lingvistiki: materialy Vsesoyuznoy konferentsii, Elista, 17-19 maya, 1972. Moscow : Izdatel'stvo «Nauka», Glavnaya redaktsiya vostochnoy literatury, 1975. pp.274-279. (in Russ.).

Biryukovich R. M. Chulymsko-tyurkskie teksty. Chulym-Turkic texts [Materials on the Chulym-Turkic language. Field records] // Skazkinarodov Sibirskogo Severa. Vyp.3. Tomsk: Izd-vo Tom.un-ta, 1980. 191 p. pp.72-84.

Biryukovich R. M., Boni R. A. Materialy po chulymsko-tyurkskomu yazyku. Polevye zapisi [Materials on the Chulym-Turkic language. Field recordings]. Tom V. Pasechnoe, 1971. 631 p. (in Russ.).

Boni R. A. Materialy po chulymsko-tyurkskomu yazyku. Polevye zapisi [Materials on the Chulym-Turkic language. Field recordings]. Tom IV. Pasechnoe, 1971. 629 с.; Tom VII. Tegul'det, 1973. 584 p. (in Russ.).

Dul'zon A. P. Materialy po dialektologii tyurkskikh narechiy Zapadnoy Sibiri [Materials on the dialectology of the Turkic dialects of Western Siberia] // Voprosy khakasskoy filologii: sbornik nauchnykh trudov. Abakan: Khakasskiy nauchno-issledovatel'skiy institut yazyka, literatury i istorii, 1977. pp.150-178. (in Russ.).

Dul'zon A. P. Materialy po dialektologii tyurkskikh narechiy Zapadnoy Sibiri [Materials on the dialectology of the Turkic dialects of Western Siberia] // Uchenye zapiski Khakasskogo nauchno-issledovatel'skogo instituta yazyka, literatury i istorii. Seriya filologicheskaya №2. Vyp. XVIII. Abakan, 1973b. pp.187-207. (in Russ.).

Dul'zon A. P. Chulymskie tatory i ikh yazyk [Chulym Tatars and their language] // Uchenye zapiski Tomskogo gosudarstvennogo pedagogicheskogo instituta. Tom IX. 1952. pp.76-211. (in Russ.).

Dul'zon A. P. Chulymsko-tyurkskiy yazyk [Chulym-Turkic language] // Yazyki narodov SSSR. Tom II. Tyurkskie yazyki. Moscow: Nauka, 1966. pp.446-466. (in Russ.).

Dybo A. V., Abubakirova L. F., Aybazova Z. K., Zimin M. M., Korovina E. V., Mal'tseva V. S., Mudrak O. A., Renkovskaya E. A., Savel'ev A. V., Khisamov O. R., Sharov A. V., Sheymovich A. V. Novye rezul'taty v genealogicheskoy klassifikatsii tyurkskikh dialektov («sluchai s affrikatami») [New results in genealogical classification of Turkic dialects ("cases with affricates")] // Oriental Studies. 2020. Vol. 13. Is. 3. pp.696-713. (in Russ.).

Evangelie ot Marka na chulymskom yazyke s parallel'nym russkim Sinodal'nym perevodom [The Gospel of Mark in Chulym with a parallel Russian Synodal translation]. Tomsk: Izd-vo «Ivan Fedorov». 132 p. (in Russ.).

Ivanov P. G. Sibirskie turki i ikh narechiya. Iz kursa lektsiy, chitannykh v tuzemnom kraevedcheskom kruzhek pri Tomskom gosudarstvennom universitete tuzemtsam-studentam Tomskikh VUZ v 1925-26 akad. godu. [Siberian Turks and their dialects. From a course of lectures given in the native local history circle at Tomsk State University to native students of Tomsk universities in 1925-26 acad. year]. Tomsk, 1927. 31 p. (in Russ.).

Ivanov P. G. Tatory srednego techeniya reki Chulyma [Tatars of the middle course of the Chulym River] // Trudy Tomskogo oblastnogo muzeya. Tomsk, 1929. T. 2. pp.87-91. (in Russ.).

Kondiyakov A. F., Lemskaya V. M. Chulymskiy yazyk d. Pasechnoe Tyukhtetskogo rayona Krasnoyarskogo kraya. 2007–2021 gg.: monografiya. [Chulymsky language, Pasechnoye village, Tyukhtetskiy district, Kras-

noyarsk Krai. 2007-2021: monograph]. T. I: slovar', formy slov i grammaticheskie primery. Tomsk: izd-vo Tom. un-ta, 2021, 204 s.; T. II: Teksty s perevodom i analizom. Tomsk: Izd-vo Tom. un-ta, 2021. 246 p. (in Russ.).

Lemskaya V. M. Materialy etnolingvisticheskikh ekspeditsiy kafedry yazykov narodov Sibiri Tomskogo gosudarstvennogo pedagogicheskogo universiteta k chulymskim tyurkam d. Kaftanchikovo Tomskogo rayona Tomskoy oblasti Rossiyskoy Federatsii: nizhnechulymskiy dialekt [Polevye zapisi: elektronnyy resurs] [Materials of ethnolinguistic expeditions of the Department of Languages of the Peoples of Siberia of Tomsk State Pedagogical University to the Chulym Turks of the village of Kaftanchikovo of the Tomsk region of the Tomsk region of the Russian Federation: Nizhnechulym dialect [Field records: electronic resource]. 2007-2010.

Lemskaya V. M. Materialy etnolingvisticheskikh ekspeditsiy kafedry yazykov narodov Sibiri Tomskogo gosudarstvennogo pedagogicheskogo universiteta k chulymskim tyurkam Tegul'detskogo rayona Tomskoy oblasti i Tyukhtetskogo rayona Krasnoyarskogo kraya Rossiyskoy Federatsii: srednechulymskiy dialekt [Polevye zapisi: elektronnyy resurs] [Material of ethnolinguistic expeditions of the Department of Languages of the Peoples of Siberia of Tomsk State Pedagogical University to the Chulym Turks of the Teguldet district of the Tomsk region and the Tyukhtet district of the Krasnoyarsk Territory of the Russian Federation: Srednechulym dialect [Field records: electronic resource], 2006-2021.

Lemskaya V. M. Chulymsko-tyurkskie teksty [Chulym-Turkic texts] // Sbornik annotirovannykh fol'klornykh i bytovykh tekstov obsko-eniseyskogo yazykovogo areala. Kollektivnaya monografiya. T. I, Tomsk: «Veter», 2010, s. 263-314; T. II, Tomsk: Agraf-Press, 2012, s. 184-237; T. III, Tomsk: «Vayar», 2013, c. 295-345; T. IV, Tomsk: TML-Press; Vayar, 2015, s. 217-291; T. V, Tomsk: TML-Press; Vayar, 2017, s. 103-177; T. VI, Tomsk: Agraf-Press; Vayar, 2020, pp. 233-276; T. VII, Tomsk: Agraf-Press; Vayar, 2020. pp. 102-127. (in Russ.).

Lemskaya V. M. Chulymsko-tyurkskiy (chulymskiy) yazyk [Chulym-Turkic (Chulym) language] // Koll. monografiya «Yazyk i obshchestvo. Entsiklopediya». Otv. red. V. Yu. Mikhal'chenko. Moscow: «Azbukovnik», 2016. pp.569-571. (in Russ.).

Malov S. E. Otchet o komandirovke studenta Vostochnogo fakul'teta Sergeya Efimovicha Malova [Tekst] [Report on the business trip of a student of the Oriental Faculty Sergey Efimovich Malov [Text]] / S. E. Malov // Izvestiya Russkogo komiteta dlya izucheniya Sredney i Vostochnoy Azii v istoricheskom, arkhologicheskom i antropologicheskom otnosheniyakh. St. Peterburg, 1909. №9. pp.35-46. (in Russ.).

N. L-v". Poъzdka Preosvyashchennago Makariya, Episkopa Biyskago, k" Chulymskim" inorodtsam" Mariinskago okruda // Tomskiya eparkhial'nyya vъdlomosti. Tomsk, 1886. pp. 15-21. (in Russ.).

Radlov V. V. Obraztsy narodnoy literatury tyurkskikh plemen, zhivushchikh v Yuzhnoy Sibiri i Dzungarskoy stepi. Chast' II. Podnarechiya Abakanskie (sagayskoe, koybal'skoe, kachinskoe), kyzyl'skoe i chulymskoe (kyuerik) [Samples of folk literature of Turkic tribes living in Southern Siberia and Dzungarian steppe. Part II. The sub-dialects of Abakan (Sagai, Koibal, Kachin), Kyzyl and Chulym (kuerik)]. SPb, 1868. 420 p. (in Russ.).

Radlov V. V. Opyt slovary tyurkskikh narechiy. T. I. Glasnye [The experience of the dictionary of Turkic adverbs. Vol. I. Vowels.]. SPb, 1893. 1023 s.; T. II. St. Peteburg, 1899. 971 s.; T. III. St. Peterburg, 1905. 1200 p.; T. IV. St. Peterburh, 1911. 1115 p. (in Russ.).

Etimologicheskii slovar' tyurkskikh yazykov: obshchetyurkskie i mezhtyurkskie leksicheskie osnovy [Etymological dictionary of Turkic languages : common Turkic and inter - Turkic lexical bases]. Moscow: Nauka, 1974-2000. (in Russ.).

Для цитирования:

Лемская В. М. Проблемы междиалектного морфонологического варьирования при разметке в корпусе миноритарного языка (на примере чулымско-тюркского) [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Lemskaya V. M. Problems of Interdialect Morphonological Variation when Processing Markup in the Corpus of a Minority Language (an Example of Chulym-Turkic). Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 29.11.2021.

УДК 811.512.1

ПЕРВЫЕ ТЮРКСКИЕ КИРИЛЛИЧЕСКИЕ КНИГИ НА ПЛАТФОРМЕ ЛИНГВОДОК

FIRST TURKIC CYRILLIC BOOKS ON THE LINGVODOC PLATFORM

Норманская Юлия Викторовна, доктор филологических наук, профессор. Институт системного программирования им. В. П. Иванникова РАН, главный научный сотрудник, зав. лабораторией «Лингвистические платформы», Институт языкознания РАН, ведущий научный сотрудник отдела «Урало-алтайских языков».

Normanskaja Julia Viktorovna, Doctor of Philology, Professor. Ivannikov's Institute of System Programming RAS, Head of the Laboratorim "Linguistic platforms", Institute of Linguistic RAS, leading researcher of the Department "Ural-Altai languages".

julianor@mail.ru.

ORCID: 0000-0002-2769-9187



В статье представлены результаты анализа первых кириллических книг на башкирском, татарском, крымско-татарском, казахском и узбекском языках, созданных в XIX-начале XX вв., и размещенных на платформе ЛингвоДок (lingvodoc.ispras.ru). Удалось выявить, 1) для башкирского: что создателем первой книги на башкирском языке (Букварь 1892) не являлся В. В. Катаринский, как традиционно считалось ранее, а северо-западный диалект в XIX в. был близок к прабашкирскому состоянию и с точки зрения морфонологии не отличался от восточно-башкирского диалекта. 2) Для татарского: в XIX в. кряшенский мишарский и казанский татарский отличались, в первую очередь, наличием vs. отсутствием перелома на пракрыпчакских широких гласных (*o, *ö, *e); а казанско-татарские кряшенские говоры различались между собой частотой огубления гласных ы, е во втором слоге, после огубленного гласного первого слога. 3) Для крымско-татарского языка было показано, что один и тот же диалект лег в основу первых словарей и литературного языка, поэтому можно считать, что литературный язык создавался не в 1928 году после лингвистической конференции, а в XIX веке крымско-татарскими просветителями. 4) Для казахского литературного языка было установлено, что в его основу положены говоры северо-западных казахских говоров Павлодара-Актюбинска, и у его истоков стоял В. В. Катаринский. 5) Для узбекского языка удалось показать, как в 1930 году была изменена основа литературного с самаркандско-бухарской на ташкентскую, а кириллические книги на узбекском языке активно издавались еще до Революции.

Ключевые слова: лингвистические платформы, цифровизация, архивы, башкирский, татарский, крымско-татарский, казахский, узбекский язык.

The article presents the results of the analysis of the first Cyrillic books in Bashkir, Tatar, Crimean Tatar, Kazakh and Uzbek languages, created in the XIX-early XX centuries, and posted on the LinguoDoc platform

(lingvodoc.ispras.ru). It was possible to identify, 1) for Bashkir: that the creator of the first book in the Bashkir language (Primer 1892) was not V. V. Katarinsky, as it was traditionally considered earlier, and the northwestern dialect in the XIX century was close to the Prabhaskir state and from the point of view of morphology did not differ from the Eastern Bashkir dialect. 2) For Tatar: in the XIX century, Kryashen Mishar and Kazan Tatar were distinguished, first of all, by the presence of vs. the absence of a fracture on the Prakypchak broad vowels (*o, *ö, *e); and the Kazan-Tatar Kryashen dialects differed among themselves in the frequency of the vowel stub y, e in the second syllable, after the stub vowel of the first syllable. 3) For the Crimean Tatar language, it was shown that the same dialect formed the basis of the first dictionaries and the literary language, so we can assume that the literary language was created not in 1928 after a linguistic conference, but in the XIX century by Crimean Tatar educators. 4) For the Kazakh literary language, it was found that it was based on the dialects of the northwestern Kazakh dialects of Pavlodar-Aktobe, and V. V. Katarinsky stood at its origins. 5) For the Uzbek language, it was possible to show how in 1930 the basis of literature was changed from Samarkand-Bukhara to Tashkent, and Cyrillic books in Uzbek were actively published even before the Revolution.

Keywords: *linguistic platforms, digitalization, archives, Bashkir, Tatar, Crimean Tatar, Kazakh, Uzbek.*

Введение

В настоящее время многие диалекты языков России не имеют стандартизованных описаний фонетики и грамматики, существующие очерки сделаны по разным стандартам, неполны и труднодоступны. Сами языки и диалекты и архивы, в которых содержатся материалы по ним, находятся под угрозой исчезновения⁹. Энтузиасты из регионов часто пытаются самостоятельно выработать графическую систему для записи текстов на своем диалекте для создания словарей, букварей и учебников. В большинстве случаев, она различается у разных носителей, и способы фиксации диалектов в XXI веке, фактически, оказываются на менее стандартизованном уровне, чем у создателей первых славянских книг в начале XI-XIII вв. Даже национальные языки недостаточно описаны, в большинстве грамматик нет четких правил, учащиеся в школах делают многочисленные ошибки и не имеют возможности выучить свой язык на высоком уровне. Это в ряде случаев провоцирует конфликты в регионах, например, в Башкирии в северо-западных районах, где диалект значительно отличается от литературного языка, нет возможности изучать его в школе, поскольку нет его грамматических описаний, учебников.

Целью статьи является представление результатов анализа первых кириллических книг на башкирском, татарском, крымско-татарском, казахском и узбекском языках XIX–нач. XX вв. Материалом послужили тексты первых кириллических книг, размещенные на платформе ЛингвоДок.

О платформе ЛингвоДок

В ЕС в настоящее время все языковые корпуса и словари, создаваемые в 25 странах собраны на единой европейской платформе «Clarin»¹⁰, в которую сдают свои материалы ученые из 1700 институтов, помимо основного постоянного финансирования Правительством ЕС Clarin имеет и грантовую поддержку из различных фондов, общий бюджет этой платформы сопоставим с финансированием европейского университета. Это позволяет сохранять диалектное богатство стран Европы.

В РФ в рамках сотрудничества Института Языкознания РАН и Института системного программирования РАН с 2012 года ведется разработка виртуальной лаборатории «ЛингвоДок» (lingvodoc.ispras.ru, главный редактор – Норманская Ю. В.), на которой есть возможность хранения и анализа материалов по языкам и диалектам Российской Федерации. В настоящее время на ЛингвоДоке

⁹ По данным ЮНЕСКО 2009, 136 языков России находятся под угрозой исчезновения.

¹⁰ По ссылке <https://www.clarin.eu/>.

уже собрано более 1000 словарей и 300 корпусов текстов на диалектах языков России, зарегистрировано более 250 активных пользователей-лингвистов, которые могут в он-лайн режиме обрабатывать свои материалы и сравнивать их с другими с помощью специальных авторских программ, воспроизводящих ту работу, которую ранее ученые делали вручную.

Возможности платформы ЛингвоДок:

- одновременная и независимая работа многих исследователей;
- уникальные программы, воспроизводящие экспериментально-фонетическую, этимологическую и морфологическую работу исследователя в 100 раз быстрее.
- возможность автоматической проверки ошибок в обработанных данных;
- возможность создания РИДов для отчетности и работы с материалом.

– работа со словарями:

- создание любых столбцов; добавление любого текста, аудиофайлов, разметки спектрограмм в фонетической программе Праат; создание этимологических связей между словами в разных словарях.
- возможность автоматической сегментации опроса носителей, подгруженного в Телеграмм-канал «Поддержка ЛингвоДок» на отдельные слова;
- программы обработки и анализа данных: фонетический анализ; поиск этимологий; анализ когнатов диалектов и нескольких языков; акустический анализ когнатов; фонологическое статистическое расстояние; фонематический анализ; реконструкция когнатов диалектов и нескольких языков;

– работа с корпусами текстов:

- загрузка аудиофайлов любого объема; (аудио) корпусов в формате Элан; текстов в формате Ворд. odt;
- автоматическое создание словарей из корпусов текстов;
- обработка с помощью парсеров: эрзянского, мокшанского, удмуртского, коми, казахского, татарского языков; возможность быстрого создания новых парсеров и их интеграции в ЛингвоДок;
- удобный интерфейс для он-лайн снятия вручную омонимии, возникшей после обработки текста парсером;

– картографирование лингвистических особенностей:

- поисковые запросы любого типа сложности и отражение их на карте;
- построение географических ареалов;
- вывод результатов в виде он-лайн фрагментов аудиословарей и корпусов с возможностью их редактирования и в формате файла Эксель;
- возможность сохранения построенной он-лайн карты в виде ссылки и ее автоматическое пополнение при добавлении новых материалов на ЛингвоДок.

Подробнее о применении этих опций в меню «Помощь»¹¹.

В настоящее время на ЛингвоДоке представлено более 1000 аудиословарей и 300 корпусов текстов по диалектам языков народов мира. На ней представлены уникальные аудиословари по исчезающим или уже исчезнувшим языкам и диалектам России, например, по восточно-мансийскому языку словарь¹² и тексты от последнего носителя языка были записаны М.К.Аmeliной в 2012-2017 гг.

Помимо современных аудиословарей на ЛингвоДоке представлено значительное количество корпусов кириллических книг, их конкордансы и кириллические словари, созданных на языках народов России в XVIII-XIX вв. Для тюркских языков большинство текстов не были известны специалистам ранее.

подавляющее большинство этих книг были переведены в результате деятельности Переводческой Комиссии Православного Миссионерского Общества. Она была создана во второй половине XIX века (1869) и находилась в ведении Братства Св. Гурия, образованного в 1867 году в г. Казани. У истоков Переводческой Комиссии стояли видные миссионеры Н. И. Ильминский, В.В. Миротворцев, И. Я. Яковлев. С начала 90-х годов XIX века ее возглавлял профессор Казанской Духовной Академии М.А.

¹¹ По ссылке <https://github.com/ispras/lingvodoc-react/wiki>.

¹² Доступен он-лайн по ссылке <http://lingvodoc.ispras.ru/dictionary/657/3/perspective/657/4/view>.

Машанов. Комиссия работала по официально признанной системе Н. И. Ильминского, в основе которой лежит первоначальное обучение детей с опорой на родной язык. Н. И. Ильминский был приверженцем простых, понятных переводов, предлагал учитывать при работе особенности родного языка. Каждый перевод, подготовленный тем или иным лицом, посылался для чтения и исправления лицам, хорошо владевшим родным языком. Только после того, как он становился безупречным со стороны языка, точности и назидательности, рукопись рекомендовалась к печати.

Важность этих первых полных текстовых и грамматических памятников для истории уральских и алтайских языков нельзя переоценить. Окончание эпохи создания этих текстов совпало с Октябрьской революцией 1917. В XX столетии, в основном, видимо, по идеологическим причинам изучение этих памятников находилось под запретом. В настоящее время нет даже полного каталога этих текстов, словарей и грамматик.

Первые тюркские кириллические книги на платформе ЛингвоДок

В настоящей статье кратко представим наиболее интересные результаты, полученные нами при анализе первых кириллических книг на башкирском, татарском, крымско-татарском, казахском и узбекском языках.

Башкирские первые кириллические книги

В соавторстве с З. Н. Экба и Р. Н. Каримовой нами были проанализированы книги на трех башкирских диалектах, см. подробнее (Норманская, Каримова, Экба 2017; Экба, Норманская, Каримова 2019; Норманская 2020):

* **южном диалекте:** (Букварь 1892)¹³;

* **восточном диалекте:** (Бессонов 1907)¹⁴;

* **северо-западном диалекте:** (Катанов 1901)¹⁵;

* **восточном или северо-западном диалекте** ¹⁶: (Евангелие 1899)¹⁷; 4) (Учение 1899)¹⁸.

* **восточном и южном диалектах:** (Словарь 1899)¹⁹ содержит формы как из южного, так и из восточного диалекта.

Наиболее ранняя из них – это (Букварь 1892). Ранее считалось, что его автором является В. В. Катаринский, который создал и (Словарь 1899), где он официально указан. Лингвистический анализ,

¹³Диалектная принадлежность этой книги рассмотрена в (Норманская, Каримова, Экба 2017).

¹⁴Диалектная принадлежность книги (Бессонов 1907) указана в названии и в тексте самой книги, возможности соотнесения материала с современными восточными башкирскими говорами разобраны в (Экба, Норманская, Каримова 2019), конкорданс книги доступен он-лайн по ссылке <http://lingvodoc.ispras.ru/dictionary/3974/1/perspective/3974/2/view>.

¹⁵Имеются в виду записи, сделанные на башкирском диалекте Белебеевского уезда, найдены в «Государственном архиве Республики Татарстан» <http://lingvodoc.ispras.ru/dictionary/2685/2450/perspective/2685/2451/view>.

¹⁶Насколько нам известно, ранее (Евангелие 1899; Учение 1899) детально проанализированы не был. Как показал анализ их морфонологических особенностей, книги были написаны на восточном, либо северо-западном диалекте.

¹⁷Глоссированный корпус (Евангелие 1899) доступен он-лайн на ЛингвоДок <http://lingvodoc.ispras.ru/dictionary/2479/8025/perspective/2479/8031/view>.

¹⁸Глоссированный корпус (Учение 1899) доступен он-лайн <http://lingvodoc.ispras.ru/dictionary/1700/5/perspective/1700/8/view>.

¹⁹(Словарь 1899) доступен он-лайн по ссылке <http://lingvodoc.ispras.ru/dictionary/714/5/perspective/714/6/view>.

проведенный нами, см. подробнее (Норманская, Каримова, Эжба 2017), позволил установить, что в Букваре, самой ранней башкирской кириллической книге, есть весьма много существенных отличий от Словаря. Во-первых, использованы другие графические символы для специфических башкирских звуков. Во-вторых, Букварь сделан на основе материала двух разных диалектов, и формы второго диалекта приводятся в скобках, они являются весьма специфическими, и отличаются от всех современных диалектов Башкирии. В качестве рабочей гипотезы, которая нуждается в дальнейшей проверке, можно предположить, что эти формы характерны для башкир-нагайбаков, проживающих в тесном языковом союзе с татарами в Челябинской области. Это хорошо согласуется и с историческими данными о том, что деятельность Уфимской миссии на территории Башкирии была начата только в 1896 году уже после создания Букваря (1892).

В результате анализа словника северо-западного «уфимского говора» башкирского языка²⁰, найденного в архиве Н.Ф.Катанова в «Государственном архиве Республики Татарстан» и записанного, вероятно, в 1898 г. в Белебеевском районе Республики Башкирия, с точки зрения графики, был выявлен ряд архаических черт, в частности, сохранение ПТ *i, *ç, *s, и инновации, свойственные северо-западным говорам, например, ПТ *ib > ū, и, особенности, не выявленные в других башкирских диалектах: ПТ *ig > u.

Было проведено также сравнение морфонологических особенностей северо-западного белебеевского говора с другими башкирскими диалектами вышперечисленных книг и дополнено сведениями о чередованиях в рассматриваемых суффиксах в материалах (Катанов 1901) по мещерскому диалекту с. Ыслак Белебеевского уезда, тептярскому диалекту д. Турунтаиш, кряшенскому татарскому д. Утарово. Сравнение с этими говорами было важно, чтобы исследовать морфонологические сходства и различия татарских и башкирских говоров в XIX в. с целью датировки возникновения близости северо-западных диалектов башкирского и современного татарского языка.

В результате сравнения были получены следующие результаты:

I. *-Lar множественное число

| | |
|---|---|
| Татарские диалекты, мещерский, тептярский и южный башкирский (Букварь 1892) | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -лV * p -нVp (в основах на -н,-м) | -тVp (в основах на глухие согласные) -дVp (в основах на -л, -н,-м) -з/дVp (в основах на -й,-р, -з/д, -у) -лVp (в основах на гласные, кроме -у)** |

* Здесь и далее V обозначает гласный, который изменяется по правилам сингармонизма в зависимости от гласного первого слога.

** Здесь и далее упоминаются не все возможные типы основ, а только те случаи, которые были зафиксированы в рассматриваемых книгах.

²⁰ Словник доступен он-лайн <http://lingvodoc.ispras.ru/dictionary/2685/2450/perspective/2685/2451/view>.

II. Падеж

II.1. Генитив (Родительный падеж)

| | |
|---|---|
| Татарские диалекты, мещерский, тептярский и южный башкирский (Букварь 1892) | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -нVң | -тVн (в основах на глухой согласный) -дVн (в основах на -м, -л, -н, -ң) -э/дVн (в основах на -р, -э/д, -й, -у) -нVн (в основах на гласный) |

II.2. Аккузатив (Винительный падеж)

| | |
|---|---|
| Татарские диалекты, мещерский, тептярский и южный башкирский (Букварь 1892) | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -нV | -тV (в основах на глухой согласный) -дV (в основах на -м, -н, -ң) -э/дV (в основах на -р, -л, -э/д, -й, -у) -нV (в основах на гласный) |

II.3. Локатив (местно-временной падеж)

| | |
|---|---|
| Татарские диалекты, мещерский, тептярский | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -дV (после основ на звонкий согласный) -тV (после основ на глухой согласный) | -тV (в основах на глухой согласный) -лV (в основах на гласный) -дV (в основах на -м, -н, -ң) -э/дV (в основах на -р, -у) |

II.4. Аблатив (Исходный падеж)

| | |
|---|--|
| Татарские диалекты, мещерский, тептярский и южный башкирский (Букварь 1892) | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -тVн -дVн (в основах -д, -л) -нVн (в основах на -н) | -тVн (в основах на глухой согласный) -дVн (в основах на -м, -н, -ң) -э/дVн (в основах на -р, -у) -лV (в основах на гласный) |

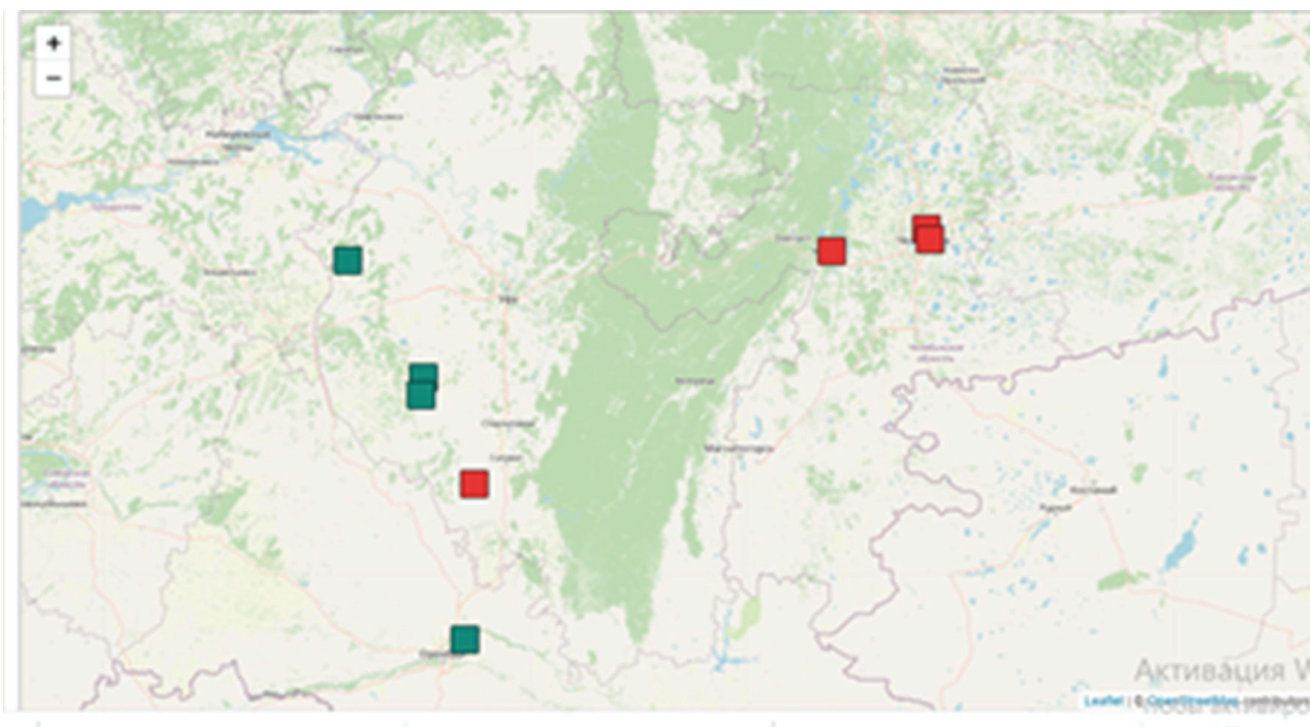
III. Абстрактный суффикс

| | |
|---|---|
| Татарские диалекты, мешцерский, тептярский | Восточный диалект (Бессонов 1907) Северо-западный диалект (Катанов 1901) (Евангелие 1899) (Учение 1899) |
| -ны (после основ на носовые) -лы (в остальных случаях) | -ты (после основ на глухой согласный) -ды (после основ на л-) -з/ды (после основ на -р) -лы (после основ на гласные) |

Проведенный анализ показывает, что в XIX в. система морфонологических чередований северо-западного диалекта полностью совпадала с языком книг, написанных на восточных говорах. Видно, что для всех разобранных типов суффиксов характерны два типа чередований. Первый (красный цвет на карте) – свойственен татарскому, мешцерскому, тептярскому и южному башкирскому, второй (бирюзовый цвет на карте) – восточным и северо-западным первым башкирским кириллическим книгам. Эти типы представлены на карте 1.

Карта 1. Типы морфонологических чередований в книгах XIX в. на татарских, мешцерских, тептярских и башкирских диалектах.

Map 1. Types of morphonological alternations in books of the 19th century in Tatar, Meshchera, Teptyar and Bashkir dialects.



На карте видно, что пункты северо-западного диалекта с севера и юга окружены были татарскими, мешцерскими, тептярскими и южными башкирскими диалектами, которые имели другой тип чередования. Вероятно, в связи с этим в XX в. произошли изменения, которые сблизили северо-западный диалект с татарским и южным башкирским.

Татарские первые кириллические книги

Нами были проанализированы девять кириллических книг, созданных в период с 1862 по 1908 г. а татарском языке: 1) Священная история от сотворения мира до кончины Иосифа по книге Бытия, изложенная на народном татарском языке (1862) (Священная история, 1862); 2) Святое Евангелие Господа нашего Иисуса Христа на народно-татарском языке (1866) (Евангелие, 1866); 3) Первый опыт словаря народно-татарского языка по выговору крещенных татар Казанской губернии. составитель Николай Остроумов (1876) (Словарь, 1876); 4) Словарь татарского языка (1880) (Словарь, 1880); 5) Требникъ на татарском языке (1881 г.) (Требник, 1881); 6) Опыты переложения христианских вероучительных книг на татарский и другие инородческие языки (1883 г.) (Опыты, 1883); 7) Чын ден княгясе (Учение о православной вере) (1897 г.) (Учение, 1897); 8) Земная жизнь Господа нашего Иисуса Христа» (1900 г.) (Земная жизнь, 1900); 9) Святое Евангелие господа нашего Иисуса Христа на татарском языке (1908 г.) (Евангелие, 1908).

Было установлено, что графико-фонетические особенности книг 1-8 весьма близки, был выявлен лишь один признак, который позволяет их различить: **в каком проценте случаев под влиянием (о), (ө) 1-го слога происходит огубление негубных кратких (ы), (е) последующих слогов.** По этому признаку выделяются три группы татарских первых кириллических книг:

1. Графическое отражение губного сингармонизма, огубление кратких (ы), (е) второго слога в позиции после (о), (ө) первого слога присутствует во всех словоформах в книгах (Словарь, 1880; Земная жизнь, 1900).

2. Сингармонизм присутствует в подавляющем большинстве словоформ в (Словарь, 1876; Требник 1881) (словоформы с огублением встречаются примерно в 20 раз чаще, чем без огубления).

3. В (Евангелие, 1866), наоборот, примеры с огублением достаточно редкие, примерно в 5 раз чаще встречаются словоформы без огубления (ы), (е) второго слога. В (Евангелие, 1908) примеров с огублением (ы), (е) второго слога несколько больше, чем в (Евангелие, 1866), но все равно случаи без огубления встречаются примерно в три раза чаще.

Эти три группы памятников можно соотнести с тремя типами татарских кряшенских говоров. Учитывая исторические данные о местах, в которых концентрировалась миссионерская православная деятельность, можно предположить, что это соответственно заказанский говор среднего диалекта; нагайбакский говор и высокогорский говор.

Наиболее ранняя из доступных нам 9 книга (Священная история, 1862) имеет значительные отличия от других книг, которое заключается в то, что в (Священная история, 1862) так называемого поволжский перелом гласных, видимо, прошел не до конца: регулярное употребляются графемы у, ү в соответствии с лит. о, ө, графемы и в соответствие с лит. е, э (при этом лит. у, ү, и присутствует в тех же лексемах, что и в памятнике). Это соответствует данным из монографии (Махмутова, 1978: 34), где указывается, что в мишарских диалектах кыпчакский процесс передвижения гласных Птю *i > e, *e > i, *u > o, *o > u, *ü > ö, *ö > ü не завершился. Вероятно, что (Священная история, 1862) является первой известной нам кириллической книгой на мишарском татарском языке.

Крымско-татарские первые кириллические книги

Для анализа первых крымско-татарских словарей был создан электронный сводный словарь²¹, в который вошли данные литературного языка и первых крымско-татарских словарей (Боданинский, 1873, Татарско-русский словарь 1902, Русско-татарский (крымско-татарский) 1906). Ранее считалось, что первые памятники написаны на южном (огузском) диалекте крымско-татарского языка, а литературный

²¹ Словарь доступен по ссылке он-лайн

<http://lingvodoc.ispras.ru/dictionary/719/12652/perspective/719/12653/view>.

язык создан на основе срединного (кыпчакского) диалекта. Для верификации этой гипотезы мы проанализировали первые кириллические крымско-татарские словари с точки зрения отражения в них диагностических кыпчакских и огузских признаков.

Ниже приводятся примеры диагностических признаков их отражения в первых словарях и литературном языке.

— **Ногайский б- vs. гагаузский б-, р-:** (Боданинский, 1873), б-/в-, (Татарско-русский словарь, 1902) п-/б-/в-, (Русско-татарский (крымско-татарский) 1906) п-/б-/в-, литературный п-/б-;

— **Ногайский т- vs. гагаузский д-, т-:** (Боданинский, 1873), д-, (Татарско-русский словарь, 1902) д-/т-, (Русско-татарский (крымско-татарский) 1906) д-/т-, литературный д-/т-;

— **Ногайский к- vs. гагаузский г-:** (Боданинский, 1873; Татарско-русский словарь, 1902; Русско-татарский (крымско-татарский) 1906), литературный г-/к-;

— **Ногайский ш - vs. гагаузский щ, ж:** (Боданинский, 1873; Татарско-русский словарь, 1902; Русско-татарский (крымско-татарский) 1906), литературный ч-;

— **Ногайский с - vs. гагаузский ш-:** (Боданинский, 1873; Татарско-русский словарь, 1902; Русско-татарский (крымско-татарский) 1906), литературный -ш-;

Итак, оказывается, что не удается выявить признаки, надежно различающие диалект первых словарей и современного крымско-татарского языка, между ними существует полная языковая преемственность. Поэтому можно считать, что литературный язык создавался не в 1928 году после лингвистической конференции, а в XIX веке крымско-татарскими просветителями. При этом все признаки, которые имеют однозначную рефлексацию, указывают на огузский характер как первых словарей, так и литературного. Можно предположить, что «ногайские» формы в первых словарях и литературном языке – результат контактов. Очевидным образом, проанализированный материал чрезвычайно важен для представления о распространении огузских и кыпчакских языков, как сегодня, так и в XIX веке.

Казахские первые кириллические книги

По сведениям библиографа А. Г. Каримуллина, в рамках деятельности Переводческой комиссии на протяжении 1860—1917 гг. на казахский язык было переведено 72 книги вероучительного содержания (Каримуллин 1971: 158). А. В. Дыбо, Ю. В. Норманской (Дыбо, Норманская, 2016), А. А. Гаджиевой (Гаджиева 2019) были проанализированы 6 первых кириллических книг на разных диалектах казахского языка, созданные в рамках деятельности Переводческой комиссии, ранее неизвестные ученым. Они были обнаружены в Национальной библиотеке Финляндии:

1) Училище благочестія на киргизскомъ языке. Езгиликке ўйрететин кнеге. Изданіе 3. Православнаго Миссіонерскаго общества. Казань: типографія М. А. Чирковой, 1892 (Училище, 1892);

2) Букварь для киргизовъ. Жазуга ўйрететин кнеге. Казань: типографія Императорскаго университета, 1892 (Букварь, 1892);

3) Крещение Руси на киргизскомъ языке. Ауліе княз Владимир тура динге бзи де крип; орыс калкында крғизгени. Казань: типографія М. А. Чирковой, 1892 (Крещение, 1892);

4) Киргизско-русскій словарь. Оренбургъ: типо-литографія Б. А. Бреслина, 1897 (Словарь, 1897)²²;

5) Букварь для киргизъ. Букуар. Казань: Центральная типографія, 1908 (Букварь, 1908)²³.

²² Словарь доступен он-лайн по ссылке

<http://lingvodoc.ispras.ru/dictionary/775/5/perspective/775/6/view>.

²³ Глоссированный корпус доступен по ссылке

<http://lingvodoc.ispras.ru/dictionary/681/5/perspective/681/9/view>; конкорданс книги доступен по ссылке <http://lingvodoc.ispras.ru/dictionary/1447/5/perspective/1447/6/view>.

б) «Акыл бере тугун кнеге» («Изъ премудрости Исуса сына Сирахова») — Казань: Типография и Литография В. М. Ключникова, 1891 (Премудрости, 1891)²⁴;

Из истории известно, что Киргизская миссия во главе с Филаретом Синьковским осуществляла свою деятельность, располагаясь до 1883 г. в Усть-Каменогорске, затем с января по декабрь 1883 г. в Семипалатинске, а с декабря 1883 г. в казачьем поселке Буконь, недалеко от города Кокпекты. В 1894 г. была открыта миссия Оренбургской епархии, действовавшая в Уральской и Тургайской областях. (Елисеев, 1900, с. 2—4). Учитывая эти факты и языковые особенности первых трех книг, см. подробнее (Дыбо, Норманская, 2016), мы пришли к выводу, что наиболее ранние первые три книги — (Училище, 1892, Букварь, 1892, Крещение, 1892) — были созданы в Семипалатинске. Более поздняя книга — (Словарь, 1897) была создана в Оренбургской области. Это подтверждается и архивными свидетельствами о переводчиках-киргизах Балгимбаеве, Бийсеневе, Сарыбатырове, Джумалиеве и Бахтигирееве, принимавших участие в создании книг см. подробнее (Гаджиева, 2019, с. 54).

А. А. Гаджиева, анализируя языковые особенности наиболее ранней книги (Премудрости, 1891), показывает, что она имеет отличия от всех ранее изученных книг: 1) упреждение гласных переднего ряда: лит. ұ — ү, лит. о — ö (тенденция упреждения есть и у носителя костанайского говора²⁵), 2) узкие гласные второго слога становятся более широкими под влиянием губной гармонии: омтолоб — лит. каз. ұмтылыпш ‘стремясь’; котолмас — лит. каз. құтылмас ‘не избежится’; көндөгүге — лит. каз. көндігуге ‘чтобы привыкнуть’. Эта особенность была также обнаружена нами и в современном костанайском говоре. На сайте lingvodoc.ispras.ru она не зафиксирована в словарях южного, восточного и центрального диалектов казахского языка. Таким образом, А. А. Гаджиева приходит к выводу, что эта книга была создана в Усть-Каменогорске (Гаджиева, 2019, с. 68-69).

Загадочной книгой оставался лишь Букварь (1908), графика которого полностью соответствует современным произносительным нормам литературного казахского языка. Лингвистический анализ этой книги позволил установить ее существенную близость с первым казахским словарем, созданным в Павлодаре. Экспедиция к казахам Павлодарской и Актюбинской областей подтвердила наши выводы о большом сходстве их говоров и современного литературного казахского языка, поскольку в них присутствует сингармонизм по огубленности, который не отмечен в диалектной речи казахов, проживающих в районе Алма-Аты. Таким образом, мы можем предположить, что литературный казахский язык формировался на основании северо-западных казахских говоров Павлодара-Актюбинска, и у его истоков стоял В. В. Катаринский.

Узбекские первые кириллические книги

В научной литературе, посвященной изучению диалектов узбекского языка, периоду становления литературного языка, фактически не упоминаются первые кириллические книги, созданные на узбекском языке до 1917 г. В (Решетов, 1958, с. 22) указывается, что носители узбекского языка вплоть до конца первой четверти XX века пользовались арабским алфавитом.

Анализ истории преподавания и изучения узбекского языка в Туркестанской учительской семинарии данный в (Германова, 2016), показывает, что, начиная с 1873 г., Н. П. Остроумов, В. П. Наливкин, М. С. Андреев, С. Лапин, И. А. Беляев создали на кириллице несколько десятков узбекских книг. Ранее эти книги, насколько нам известно, не были проанализированы тюркологами с точки зрения особенностей отраженных в них диалектов.

²⁴ Глоссированный корпус доступен по ссылке <http://lingvodoc.ispras.ru/dictionary/1752/7470/perspective/1752/7473/view>; конкорданс книги доступен по ссылке <http://lingvodoc.ispras.ru/dictionary/1752/7483/perspective/1752/7484/view>.

²⁵ См. подробнее аудиословарь современного говора <http://lingvodoc.ispras.ru/dictionary/2705/15/perspective/2705/16/view>.

В настоящее время на платформе ЛингвоДок представлены два доступных нам издания, подготовленных для обучения студентов Туркестанской учительской семинарии: В. П. Наливкин «Сартовско-персидская хрестоматия», 1887²⁶ и И. А. Беляев «Руководство к изучению сартовского языка. Т. 1: Общий курс / Сост. по методу Туссена и Лангеншейдта», 1896²⁷, их подробный анализ представлен в (Норманская, 2019). В результате было выявлено, что в их основе лежат самаркандско-бухарские говоры, при этом в обеих книгах присутствуют дублетные формы и определенная часть недублетных форм в (Наливкин, 1887), например, слова с рефлексацией пратюркского *a > o в структурах *CaCaC, *aCaC, *CaCa указывающие на ферганские диалекты. Факт цитирования этих форм в виде дублетов указывает, что авторы осознавали одновременное использование двух диалектных норм. Интересно, что в рассматриваемых книгах нет диагностических признаков ташкентских говоров, отличающих их от ферганских и самаркандских диалектов и подробно рассмотренных в (Шаабдурахманов, 1966): рефлексации o-o в структурах *CaCaC, *aCaC, *CaCa, отпадения ауслатных согласных -р, -к, -h, -й.

Таким образом, первые кириллические книги на узбекском языке дают нам интересные исторические данные о доминировании самаркандско-бухарских диалектов в Ташкенте в конце XIX века. Вероятно, именно поэтому в 1926 г. орфографическая конференция проголосовала большинством голосов за наличие сингармонизма в узбекском литературном языке, поскольку сингармонизм присутствовал и в первых кириллических книгах, и в самаркандско-бухарских говорах. Поэтому не вполне справедливыми кажутся высказывания В. В. Решетова, в статье, посвященной становлению узбекского литературного языка, где он пишет, что сингармонизм «отсутствует в городских говорах, которые, несмотря на открытое и активное сопротивление сторонников сингармонизма, оказались определяющими в наши дни» (Решетов, 1958, с. 26). Вероятно, за сопротивлением стояла борьба за влияния между представителями ташкентской и самаркандско-бухарской интеллигенции, в результате которой победа осталась за Ташкентом, куда была в 1930 году перенесена столица из Самарканда, и чей говор стал фонетической основой для литературного языка. Но становится ясно, что произошло это менее ста лет назад, а ранее на самаркандско-бухарском диалекте русскими учеными была развита кириллическая письменность, подготовлено несколько десятков книг, дальнейшее изучение которых позволит глубже понять историю узбекского языка и народа.

Заключение

Анализ на платформе ЛингвоДок первых кириллических книг на тюркских языках позволил выявить,

– что создателем первой книги на башкирском языке (Букварь, 1892) не являлся В. В. Катаринский, как традиционно считалось ранее, а северо-западный диалект в XIX в. был близок к прабашкирскому состоянию и с точки зрения морфонологии не отличался от восточно-башкирского диалекта.

– В XIX в. крышненский мишарский и казанский татарский отличались, в первую очередь, наличием vs. отсутствием перелома на пракыпчакских широких гласных (*o, *ö, *e); а казанско-татарские крышненские говоры различались между собой частотой огубления гласных ы, е во втором слоге, после огубленного гласного первого слога;

– для крымско-татарского языка было показано, что один и тот же диалект лег в основу первых словарей и литературного языка, поэтому можно считать, что литературный язык создавался не в 1928 году после лингвистической конференции, а в XIX веке крымско-татарскими просветителями;

– что в основу казахского литературного языка были положены северо-западные казахские говоры Павлодара-Актюбинска, и у его истоков стоял В. В. Катаринский.

²⁶ Конкорданс этой книги доступен по ссылке <http://lingvodoc.ispras.ru/dictionary/1716/6/perspective/1716/7/edit>.

²⁷ Конкорданс этой книги доступен по ссылке <http://lingvodoc.ispras.ru/dictionary/1716/5598/perspective/1716/5600/edit>.

– для узбекского языка удалось показать, как в 1930 году была изменена основа литературного с самаркандско-бухарской на ташкентскую, а кириллические книги на узбекском языке активно издавались еще до Революции.

Наличие корпусов и конкордансов первых кириллических книг и словарей на платформе ЛингвоДок позволяет любому читателю верифицировать полученные выводы на полном материале, поскольку эти книги, в основном, не доступны в библиотеках РФ.

Литература

Беляев, И. А. Руководство к изучению сартовского языка. Т. 1: Общий курс / Сост. по методу Туссена и Лангеншейдта, Ташкент, 1896. 104 с.

Бессонов, А. Г. Бессонов, А. Г. Букварь для башкир. Казань, 1907. 46 с.

Боданинский, А. Русско-татарский букварь для чтения в первоначальных народных школах Таврической губернии. Одесса, 1873. 64 с.

Букварь для башкир. Оренбург, 1892.

Букварь для киргизовъ. Жазуга ўйрететин кнеге. Казань, 1892.

Букварь для киргизъ. Букуар. Казань, 1908.

Гаджиева, А. А. Анализ системы вокализма и консонантизма в первых кириллических книгах на казахском языке // Урало-алтайские исследования. 2019, 3 (34). С. 53—71.

Германова, В. У истоков литературного узбекского языка // <https://vesti.uz/u-istokov-uzbekskogo-literaturnogo-yazyuka/> (дата обращения 15.11.2021).

Дыбо, А. В., Норманская, Ю. В. Первые кириллические книги на казахском языке как источники для изучения истории диалектов и создания литературной нормы // Урало-алтайские исследования. 2016, 4 (23). С. 138—152.

Святое Евангелие Господа нашего Иисуса Христа на народно-татарском языке, Казань, 1866.

Святое Евангелие от Матфея на башкирском языке (опыт). Казань, 1899.

Святое Евангелие господа нашего Иисуса Христа на татарском языке. Казань, 1908.

Елисеев, Е. Записки миссионера Буконского стана Киргизской миссии за 1892—1899 гг. Санкт-Петербург, 1900. 188 с.

Земная жизнь Господа нашего Иисуса Христа. Казань, 1900.

Каримуллин А. Г. У истоков татарской книги. От начала возникновения до 60-х годов XIX в. Казань, 1971. 224 с.

Катанов Н. Ф. Отчет о поездке, совершенной с 20 июля по 20 августа 1898 г. в Белебеевский уезд Уфимской губ. // УЗ Императорского Казанского университета. 1 – 3 книги. Казань, 1901. С. 1–90.

Аулие князь Владимир тура динге бзи де крип; орыс калкын-да кргизгени. Крещение Руси на киргизскомъ языке. Казань, 1892. 27 с.

Махмутова, Л. Т. Опыт исследования тюркских диалектов: мишарский диалект татарского языка. Москва : Наука, 1978. 271 с.

Наливкин, В. П. Сартовско-персидская хрестоматия. Самарканд, 1887.

Норманская Ю.В. Первые кириллические учебные издания на узбекском языке и их диалектная принадлежность // Слова мои останутся, а сам я ухожу... Scripta in memoriam D.M.Nasilov. Москва, 2019. С. 239–254.

Норманская, Ю. В. Рукописный словарь уфимского гвора башкирского языка, записанный Н.Ф.Катановым, как свидетельство архаичности северо-западного диалекта башкирского языка в конце XIX в. // Башкирская идентичность в тюркоязычной культурной парадигме России, Уфа, 2020, часть 2, том 1, с. 10 – 20.

Норманская, Ю. В., Каримова, Р. Н., Эмба, З. Н. В. В. Катаринский — автор первой кириллической книги на башкирском языке? // Урало-алтайские исследования. 2017, 2 (25). С. 46–52.

Опыты переложения христианских вероучительных книг на татарский и другие инородческие языки, Казань, 1883. 356 с.

- Акыл бере тугун кнеге. Изъ Премудрости Исуса сына Сирахова. Казань, 1891.
- Решетов В.В. Опорный диалект при образовании узбекского национального языка // Вопросы диалектологии тюркских языков. Баку, 1958. С. 22 – 36.
- Полный русско-татарский словарь: крымско-татарского наречия / Сост. Заатов О. Симферополь, 1906. 120 с.
- Священная история от сотворения мира до кончины Иосифа по книге Бытия, изложенная на народном татарском языке. Казань, 1862. 137 с.
- Первый опыт словаря народно-татарского языка по выговору крещенных татар Казанской губернии. составитель Николай Остроумов. Казань, 1876. 145 с.
- Словарь татарского языка. Казань, 1880.
- Киргизско-русский словарь. Оренбургъ, 1897.
- Краткий русско-татарский словарь: крымское наречие / сост. Танатар С. М. [и др.]; ред. Гаспринский И. Одесса, 1902. 25 с.
- Требникъ на татарском языке 1881.
- Чын ден княгясе [Учение о православной вере]. Казань, 1897.
- О загробной жизни по учению православной церкви. Оренбург. 1899.
- Езгиликке үйрететин кнеге. Училище благочестія на киргизскомъ языке. Издание 3. Православнаго Миссионерскаго общества. Казань, 1892.
- Шаабдурахманов, Ш. Некоторые фонетические особенности ташкентского диалекта // Вопросы диалектологии тюркских языков. Баку, 1966. С. 48 – 54.
- Экба, З. Н., Норманская Ю. В., Каримова Р. Н. Три диалекта в «Букваре для башкир» А.Г.Бессонова // Урало-алтайские исследования. 2019, 2 (33). С. 101—111.

References

- Belyaev, I. A. Rukovodstvo k izucheniyu sartovskogo yazyka. T. 1: Obshchiy kurs [Guide to the study of the Sartov language. Vol. [1: General course] / Sost. po metode Tussena i Langensheydta [Comp. according to the method of Toussaint and Langenscheidt], Tashkent, 1896. 104 p. (in Russ.).
- Bessonov, A. G. Bukvar' dlya bashkir [Primer for Bashkirs]. Kazan, 1907. 46 p. (in Russ.).
- Bodaninsky, A. A Russko-tatarskiy bukvar' dlya chteniya v pervonachal'nykh narodnykh shkolakh Tavricheskoy gubernii [Russian-Tatar primer for reading in the original public schools of the Tauride province]. Odessa, 1873. 64 p. (in Russ.).
- Bukvar' dlya bashkir [Primer for Bashkirs]. Orenburg, 1892. (in Russ.).
- Bukvar' dlya kirgizov". Zhazuga üyretetin knege [Primer for Kyrgyzstan. Boucoire]. Kazan, 1908. (in Russ.).
- Bukvar' dlya kirgizov". Zhazuga üyretetin knege [Primer for the Kyrgyz]. Kazan, 1892. (in Russ.).
- Chyn den knyagyase. Uchenie o pravoslavnoy vere [Doctrine of the Orthodox faith]. Kazan, 1897.
- Dybo A. V., Normanskaya Yu. V. Pervye kirillicheskie knigi na kazakhskom yazyke kak istochniki dlya izucheniya istorii dialektov i sozdaniya literaturnoy normy [The first Cyrillic books in the Kazakh language as sources for studying the history of dialects and creating a literary norm] // Uralo-altayskie issledovaniya. 2016, 4 (23). pp. 138-152. (in Russ.).
- Ekba, Z. N., Normanskaya Yu. V., Karimova R. N. Tri dialekta v «Bukvare dlya bashkir» A.G.Bessonova [Three dialects in the "Primer for Bashkirs" by A.G.Bessonov] // Uralo-altayskie issledovaniya. 2019, 2 (33). pp. 101-111. (in Russ.).
- Eliseev E. Zapiski missionera Bukonskogo stana Kirgizskoy missii za 1892—1899 gg. [Notes of the missionary of the Bukon camp of the Kirghiz Mission for 1892-1899]. St. Petersburg, 1900. (in Russ.).
- Ezgilikke üyretetin knege. Uchilishche blagochestiya na kirgizskom" yazyke. Izdanie 3. Pravoslavnago Missionerskago obshchestva [Ezgilikkeperiretetin knege. The School of Piety in the Kyrgyz language. Edition 3. Of the Orthodox Missionary Society]. Kazan, 1892.
- Gadzhieva, A. A. Analiz sistemy vokalizma i konsonantizma v pervykh kirillicheskiykh knigakh na

- kazakhskom yazyke [Analysis of the system of vocalism and consonantism in the first Cyrillic books in the Kazakh language] // Uralo-altayskie issledovaniya. 2019, 3 (34). pp. 53-71. (in Russ.).
- Germanova, V. U istokov literaturnogo uzbekskogo yazyka [At the origins of the literary Uzbek language]. Available at: <https://vesti.uz/u-istokov-uzbekskogo-literaturnogo-yazyka/> (accessed 15.11.2021).
- Gospoda nashego Iisusa Khrista.Zemnaya zhizn' Gospoda nashego Iisusa Khrista [The earthly life of 1900 is the earthly life of our Lord Jesus Christ]. Kazan, 1900. (in Russ.).
- Karimullin, A. G. U istokov tatarskoy knigi. Ot nachala vzniknoveniya do 60-kh godov XIX v. [At the origins of the Tatar book. From the beginning of the emergence to the 60s of the XIX century]. Kazan, 1971. (in Russ.).
- Katanov, N. F. Otchet o poezdke, sovershennoy s 20 iyulya po 20 avgusta 1898 g. v Belebeevskiy uезд Ufimskoy gub. [Report on a trip made from July 20 to August 20, 1898 to Belebeysky district of Ufa province] // UZ Imperatorskogo Kazanskogo universiteta. 1-3 books. - Kazan, 1901, pp. 1-90. (in Russ.).
- Kirgizsko-russkiy slovar' [Kyrgyz-Russian dictionary]. Orenburg, 1897. (in Russ.).
- Kratkiy russko-tatarskiy slovar': krymskoe narechie [A short Russian-Tatar dictionary: Crimean adverb] / comp. Thanatar S. M. et al.; ed. Gasprinsky I. Odessa, 1902. (in Russ.).
- Makhmutova L. T. Opyt issledovaniya tyurkskikh dialektov: misharskiy dialekt tatarskogo yazyka [The experience of studying Turkic dialects: Mishar dialect of the Tatar language]. Moscow : Nauka, 1978. (in Russ.).
- Nalivkin, V. P. Sartovsko-persidskaya khrestomatiya [Sartov-Persian anthology]. Samarkand, 1887. (in Russ.).
- Normanskaya, Y. V., Karimova R. N., Ekba Z. N. V. V. Katarinskiy — avtor pervoy kirillicheskoy knigi na bashkirskom yazyke? [Katarinsky is the author of the first Cyrillic book in the Bashkir language?] // Uralo-altayskie issledovaniya. 2017, 2 (25). pp. 46-52. (in Russ.).
- Normanskaya, Yu. V. Pervye kirillicheskie uchebnye izdaniya na uzbekskom yazyke i ikh dialektnaya prinadlezhnost' [The first Cyrillic educational publications in Uzbek and their dialect affiliation] // Slova moi ostanutsya, a sam ya ukhozhu... Scripta in memorium D.M.Nasilov. Moscow, 2019. pp. 239-254. (in Russ.).
- Normanskaya, Yu. V. Rukopisnyy slovar' ufimskogo gvora bashkirskogo yazyka, zapisanny N.F.Katanovym, kak svidetel'stvo arkhachnosti severo-zapadnogo dialekta bashkirskogo yazyka v kontse XIX v. [A handwritten dictionary of the Ufa folklore of the Bashkir language, recorded by N.F.Katanov, as evidence of the archaism of the Northwestern dialect of the Bashkir language at the end of the XIX century]. // Bashkirskaya identichnost' v tyurkoyazychnoy kul'turnoy paradigme Rossii. Ufa, 2020, part 2, volume 1, pp. 10-20. (in Russ.).
- O zagrobnoy zhizni po ucheniyu pravoslavnoy tserkvi [The doctrine of 1899 is about the afterlife according to the teachings of the Orthodox Church]. Orenburg. 1899. (in Russ.).
- Opyty pereozheniya khristianskikh verouchitel'nykh knig na tatarskiy i drugie inorodcheskie yazyki [Experiments of translating Christian doctrinal books into Tatar and other foreign languages], Kazan, 1883. (in Russ.).
- Pervyy opyt slovarya narodno-tatarskogo yazyka po vygovoru kreshchennykh tatar Kazanskoy gubernii. sostavitel' Nikolay Ostroumov [The first experience of the dictionary of the national Tatar language on the pronunciation of the baptized Tatars of the Kazan province. compiled by Nikolai Ostroumov]. Kazan, 1876. (in Russ.).
- Polnyy russko-tatarskiy slovar': krymsko-tatarskogo narechiya [Complete Russian Tatar Dictionary: Crimean Tatar Dialect] / Comp. Zaatov O. Simferopol, 1906. (in Russ.).
- Reshetov, V. V. Opornyy dialekt pri obrazovanii uzbekskogo natsional'nogo yazyka [Basic dialect in the formation of the Uzbek national language] // Voprosy dialektologii tyurkskikh yazykov. Baku, 1958. pp. 22-36. (in Russ.).
- Shaabdurakhmanov, Sh. Nekotorye foneticheskie osobennosti tashkentskogo dialekta [Some phonetic features of the Tashkent dialect] // Voprosy dialektologii tyurkskikh yazykov. Baku, 1966. pp. 48 - 54. (in Russ.).
- Slovar' tatarskogo yazyka. [Dictionary of the Tatar language]. Kazan, 1880. (in Russ.).
- Svyatoe Evangelie Gospoda nashego Iisusa Khrista na narodno-tatarskom yazyke, Kazan' [The Holy Gospel of our Lord Jesus Christ in the Tatar language]. Kazan, 1866. (in Russ.).

Svyatoe Evangelie gospoda nashego Iisusa Khrista na tatarskom yazyke [The Holy Gospel of our Lord Jesus Christ in the Tatar language]. Kazan, 1908. (in Russ.).

Svyatoe Evangelie ot Matfeya na bashkirskom yazyke (opyt) [The Holy Gospel of Matthew in the Bashkir language (experience)]. Kazan, 1899. (in Russ.).

The Sacred History of 1862 is a sacred history from the creation of the world to the death of Joseph according to the Book of Genesis, set forth in the Tatar folk language. Kazan, 1862. (in Russ.).

Trebnik" na tatarskom yazyke [Breviary in the Tatar language] 1881.

Akyl bere tugun knege. Iz" Premudrosti Iisusa syna Sirakhova [From the Wisdom of Jesus the son of Sirach]. Kazan, 1891.

Для цитирования:

Норманская Ю.В. Первые тюркские кириллические книги на платформе ЛингвоДок [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Normanskaja J. V. First Turkic cyrillic books on the LingvoDoc platform. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 23.11.2021.

ЭЛЕКТРОННЫЕ БАЗЫ ДАННЫХ ТУВИНСКИХ ФОЛЬКЛОРНЫХ ТЕКСТОВ

ELECTRONIC DATABASES OF TUVAN FOLKLORE TEXTS



Ондар Менги Васильевна, кандидат филологических наук, заместитель директора ГБУ «Центр развития тувинской традиционной культуры и ремесел».

Ondar Mengi Vasilyevna, Candidate of Philological Sciences, Deputy Director of the State Budgetary Institution "Center for the Development of Tuvan Traditional Culture and Crafts".

mengi89@yandex.ru
ORCID: 0000-0002-8270-7071

В статье представлена информация о созданных электронных базах данных по фольклорным произведениям тувинского народа, которые входят в Электронный корпус текстов тувинского языка. Описываются структура и принципы организации баз данных по текстам тувинских героических сказаний и народным песням. Указывается, что фольклорные тексты специфичны в силу того, что многие из них имеют свои варианты. Многие тувинские тексты являются заимствованными из монгольского фольклора. Так, многие варианты произведений, хранящиеся в архивах, еще ждут их включения в базы данных. В связи с этим, рассматриваются проблемы и вопросы, связанные с их цифровизацией, обновлением и пополнением с учетом особенностей языка и текстологии произведений. Планируется включить в разрабатываемые базы данных аудио- и видеоприложения. И самое главное – с созданием баз данных по текстам героических сказаний и народных песен тувинцев появилась реальная возможность работать более эффективно в их исследовании.

Ключевые слова: электронные базы данных, фольклорные тексты, музыкальный фольклор, героические сказания.

The article presents information about the created electronic databases on folklore works of the Tuvan people, which are included in the Electronic corpus of texts of the Tuvan language. The structure and principles of the organization of databases on the texts of Tuvan heroic tales and folk songs are described. It is indicated that folklore texts are specific due to the fact that many of them have their own variants. Many Tuvan texts are borrowed from Mongolian folklore. So, many versions of works stored in archives are still waiting for their inclusion in the database. In this regard, the problems and issues related to their digitalization, updating and replenishment are considered, taking into account the peculiarities of the language and textology of works. It is planned to include audio and video applications in the databases being developed. And most importantly, with the creation of databases on the texts of heroic tales and folk songs of Tuvinians, there was a real opportunity to work more effectively in their research.

Keywords: *electronic databases, folklore texts, musical folklore, heroic tales.*

Введение

О важности использования компьютерных технологий в гуманитарных исследованиях начали говорить еще в середине XX века (Дикарева, 2011, 19). В особенности те, кто связал свою судьбу с гуманитарной наукой, в последнее время испытывает потребность в инновационных технологиях, которые повысили бы скорость сбора материала и эффективность его обработки.

Начиная с 2009 г. силами сотрудников Тувинского государственного университета, научными сотрудниками НОЦ «Тюркология», преподавателями и студентами филологического и физико-математического факультетов, проводятся работы по созданию и совершенствованию Электронного корпуса текстов тувинского языка (далее ЭКТТЯ), призванного обеспечить высокие темпы научного исследования, изучения и актуализации тувинского языка (Салчак, 2012). В настоящее время данный корпус содержит несколько подкорпусов, в числе которых значится подкорпус текстов фольклорных произведений, одной из задач которого является разработка баз данных (далее – БД) фольклорных текстов.

Работа по созданию электронных баз данных по имеющимся фольклорным текстам вызвала определенную сложность тем, что до сих пор в тувиноведении отсутствуют работы, посвященные методам создания БД. Поэтому работа по созданию БД по фольклору является авторской работой. В настоящее время для ЭКТТЯ, подкорпуса фольклора созданы следующие БД: «Тувинские героические сказания», «Антропонимы в тувинских героических сказаниях», «Клише и стандарты в тувинских героических сказаниях» и «Тувинские народные песни». Все указанные БД прошли регистрацию в Федеральной службе по интеллектуальной собственности.

Целью статьи является представление электронных баз данных по фольклорным произведениям тувинского народа в Электронном корпусе текстов тувинского языка (ЭКТТЯ). Материалом статьи послужили базы в ЭКТТЯ.

Электронные базы данных тувинских фольклорных текстов

Целью настоящей статьи является описание структуры созданных нами баз данных текстов героических сказаний и народных песен. Фольклорные тексты специфичны в силу того, что многие из них имеют свои варианты – часто результаты заимствования из фольклора других народов. Так, среди тувинских фольклорных текстов имеются заимствованные из монгольского фольклора, которые проникли в тувинский фольклор двумя путями: письменным и устным (Донгак, 2010, с. 49). Письменным путем проникли тексты, переведённые из монгольского языка, и, функционирующие в печатном виде. Заимствованными устным путем являются тексты, распространившиеся по Туве, благодаря сказителям, сказочникам, певцам и другим нашим талантливым предкам. Поэтому созданные нами БД включают различные варианты фольклорных произведений.

БД по текстам тувинских героических сказаний. В подкорпусе фольклора ЭКТТЯ БД «Тувинские героические сказания» является самой обширной и включает 52 текста, 6 из которых являются повторяющимися в разных изданиях. Она была создана на основе опубликованных текстов тувинских героических сказаний.

Многовариантным из них является текст «Бокту-Кириш, Бора-Шээлей», который имеет 4 разных опубликованных варианта, исполненных разными сказителями. В своей научной работе С. М. Орус-оол фиксирует, что данное сказание имеет 18 вариантов, которые хранятся в Научном фонде Тувинского института гуманитарных и прикладных социально-экономических исследований при Правительстве Республики Тыва (Орус-оол, 2001). Поскольку неопубликованные варианты этого произведения в настоящую БД не входят, ее следует дополнять неопубликованными рукописными текстами из архивных фондов в дальнейшем, желательно снабдить их аудиофайлами, т.к. интерес представляет и исполнительское искусство сказителя.

На основе имеющейся БД «Тувинские героические сказания» создана БД «Антропонимы в тувинских героических сказаниях». Она была разработана с учётом структуры личных имен персонажей, которые характеризуются своей многокомпонентностью, включающей кличку коня, само личное имя и титул персонажа. Поэтому здесь используется сравнительно простая метаразметка, которая имеет следующую структуру: 1. Имя женское или мужское. 2. Количество компонентов. 3. Название титула в структуре личного имени. В текстах тувинских героических сказаний в среднем встречается около 14 личных имен. Наибольшее количество личных имен употребляется в тексте «Бокту-Кириш, Бора-Шээлей» в исполнении Иргита Ширинена. Там было обнаружено 30 антропонимов. Всего 5 личных имен употребляется в тексте «Анан-Даваа».

Третьей базой является БД «Клише и стандарты в тувинских героических сказаниях». Клише и стандарты, используемые в тексте героического сказания, относят к опорным, текстообразующим компонентам в структуре текста, обязательным элементам каждого эпического текста. Эти устойчивые фразы собраны в БД методом сплошной выборки и с учетом типических мест в сказаниях. По мнению С. М. Орус-оол, клише и стандарты «помогают запоминать большие тексты» (2001, 18).

Составляя БД тувинских героических сказаний, собственных имен, стандартов и клише по всем текстам, мы обнаружили, что в них присутствуют явные заимствования из монгольского языка. А сопоставляя уже целые тексты, выявили, что многие тексты в корпусе текстов героических сказаний являются заимствованными и соответствуют монгольским.

На основе БД по героическим сказаниям проводятся исследования. Защищена диссертация на соискание ученой степени кандидата филологических наук Ондар М.В. на тему «Особенности языка тувинских героических сказаний (в сопоставительном аспекте)» (2020).

БД по текстам музыкального фольклора. По музыкальному фольклору нами была разработана БД «Тувинские народные песни». Она представляет интерес тем, что в корпус текстов песен также были включены народные песни, сохранившиеся у китайских и монгольских этнических тувинцев. Так, в корпус текстов тувинских народных песен было включено более 1000 текстов, которые встречаются в опубликованных сборниках песен. В ходе работы по созданию БД песен было выявлено, что многие тексты являются повторяющимися, встречаются в разных сборниках. Тут также встает проблема вариативности – один из главных признаков фольклорного текста. Так, при создании БД была учтена вариативность текстов песен. В настоящее время работа над БД по народным песням еще продолжается. Разрабатывается структура и содержание данных баз. В будущем планируется включить в БД аудио- и видеоприложения.

Заключение

Указанные электронные БД в дальнейшем будут обновляться и пополняться с учетом особенностей языка и текстологии произведений. Самое главное – появилась реальная возможность работать более эффективно, не упуская детали, вовлекая в анализ одновременно практически все изданные тексты героических сказаний и народных песен тувинцев.

Базы данных

Бавуу-Сюрюн М. В., Далаа С. М., Монгуш Ч. М., Ондар М. В., Клише и стандарты в текстах тувинских героических сказаний. Свидетельство о государственной регистрации базы данных № 2017620024 от 10.01.2017.

Бавуу-Сюрюн М. В., Далаа С. М., Монгуш Ч. М., Ондар М. В. Тувинские героические сказания. Свидетельство о государственной регистрации базы данных № 2017620090 от 19.01.2017.

Далаа С. М., Бавуу-Сюрюн М. В., Ондар М. В. Антропонимы в тувинском героическом эпосе. Свидетельство о государственной регистрации базы данных № 2018621622 от 18.10.2018.

Далаа С. М., Монгуш А. К., Ондар М. В., Тувинские народные песни. Свидетельство о государственной регистрации базы данных № 2019620654 от 19.04.2019.

Литература

Дикарева, С. С. Технологии иноязычного образования в эпоху электронных корпусов текстов // Ученые записки Таврического национального университета им. В.И. Вернадского. Серия «Филология. Социальные коммуникации». Том 24(63) №1. Часть 1, 2011. С.17-23.

Донгак, А. С. Индо-тибето-монгольская «обрамленная повесть» в тувинской сказочной традиции». Абакан, 2010. 178 с.

Ондар, М. В. Особенности языка тувинских героических сказаний (в сопоставительном аспекте). Дисс. на соискание ученой степени канд. филол. наук. Москва, 2020. 302 с.

Орус-оол, С. М. Тувинские героические сказания (текстология, поэтика, стиль). Москва, 2001. 431 с.

Салчак, А. Я. Электронный корпус текстов тувинского языка // Новые исследования Тувы. № 3, 2012. URL : https://www.tuva.asia/journal/issue_15/5231-salchak.html (дата обращения: 12.11.2021.).

References

Dikareva, S.S. Tehnologii inozazychnogo obrazovaniya v jepohu jelektronnyh korpusov tekstov [Technologies of foreign language education in the era of electronic text corpora] // Uchenye zapiski Tavricheskogo nacional'nogo universiteta im. V.I. Vernadskogo. Serija «Filologija. Social'nye kommunikacii». [Scientific notes of V.I. Vernadsky Tauride National University. The series "Philology. Social communications"]. Vol. 24(63) No. 1. Part 1, 2011. pp. 17–23. (in Russ.).

Dongak, A.S. Indo-tibeto-mongol'skaja «obramlennaja povest'» v tuvinskoj skazochnoj tradicii» [Indo-Tibeto-Mongolian "framed story" in the Tuvan fairy tale tradition"]. Abakan, 2010. 178 p. (in Russ.).

Ondar, M. V. Osobennosti jazyka tuvinskih geroicheskikh skazanij (v sopostavitel'nom aspekte) [Features of the language of Tuvan heroic tales (in a comparative aspect)]. Diss. na soiskanie uchenoj stepeni kand. filol. nauk [Diss. for the degree of Candidate of Philology Sciences]. Moscow, 2020. 302 p. (in Russ.).

Orus-ool, S. M. Tuvinskie geroicheskie skazanija (tekstologija, pojetika, stil') [Tuvan heroic tales (textology, poetics, style)]. Moscow, 2001. 431 p. (in Russ.).

Salchak, A. Ja. Jelektronnyj korpus tekstov tuvinskogo jazyka [. Electronic corpus of Tuvan language texts] // Novye issledovaniya Tuvy. [New studies of Tuva] № 3, 2012. Available at : https://www.tuva.asia/journal/issue_15/5231-salchak.html (accessed: 12.11.2021.).

Для цитирования:

Ондар М.В. Электронные базы данных тувинских фольклорных текстов [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Ondar M. V. Electronic databases of Tuvan folklore texts. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 15.09.2021.

УДК 811.512.133+ 81'33

DEVELOPMENT OF SOFTWARE FOR A TEXT CORPUS IN UZBEK

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ТЕКСТОВОГО
КОРПУСА НА УЗБЕКСКОМ ЯЗЫКЕ

Турсунов Мухаммадсолих Саъдин угли, ассистент кафедры программной инженерии, Самаркандский филиал Ташкентского Университета информационных технологий имени Мухаммеда аль-Хорезми, Самарканд, Узбекистан.

Tursunov Mukhammadsolikh Sadin ugli, Department of Software Engineering Samarkand Branch of Tashkent University of Information Technologies named after Muhammad al-Khwarezmi, Samarkand, Uzbekistan.

muhammadsolih927@gmail.com

ORCID: 0000-0002-6485-3630

В статье рассматриваются теоретические и практические вопросы, возникшие в процессе создания корпуса узбекского языка. При создании корпуса узбекского языка исследовался зарубежный опыт и виды программного обеспечения. Современный корпус будет разработан на узбекском языке как сбалансированный, масштабный и универсальный корпус. Перед созданием корпуса были изучены теоретические и практические методы. На протяжении всего процесса для решения конкретных проблем использовалось различное программное обеспечение, и созданный корпус будет открытым исходным кодом для некоммерческого использования. В статье описаны начальные этапы построения корпуса и требования к его созданию. Разработку корпуса можно рассматривать как технологический процесс, осуществляемый в несколько этапов. Наиболее важными этапами являются токенизация, разметка и преобразование текстов в формат программного обеспечения для управления корпусом.

Ключевые слова: корпус, разметка, токен, морфологическая разметка, синтаксическая разметка, семантическая разметка, экстралингвистическая разметка, менеджер корпуса, согласование, AntConc, Sketch Engine.

This article is dedicated to theoretical and practical issues that arose during the creation of corpus for Uzbek language. Foreign experience and software have been investigated while creating the corpus of the Uzbek language. The modern corpus will be designed and developed in the Uzbek language as a balanced, large-scale and universal corpus. The theoretical and practical methods have been studied before the creation of the corpus. Throughout the process, different softwares have been used to solve specific problems. The created corpus will be an open source for non-commercial use. The article describes the initial stages of the structure of the corpus and the requirements for the creation of modern corpus. The corpus search system allows to search for words, phrases, grammatical forms and different meanings of words and present them in a user-friendly format. There are several types of corpus, such as national corpus, author corpus, book corpus, newspaper corpus, oral corpus, and the purpose of this article is the development of a corpus based on the texts of a fiction writing. The development of a corpus might be considered as a technological process that is carried out in

several stages. The most important steps are tokenization, markup, and conversion of the markup texts into the format of the corpus control software. Software tools and methodologies have also been developed for specific tasks, for instance, analysis of ideas, translation, collection and analysis of target information. The corpus of the Uzbek language will be created with the help of Uzbek philologists, researchers, information technology specialists and programmers.

Keywords: *corpus, markup, token, morphological markup, syntactic markup, semantic markup, extralinguistic markup, corpus manager, concordance, AntConc, Sketch Engine.*

Introduction

We consider a corpus as a set of electronic texts and a software package that serves to create the corpus and perform search operations in it. The corpus search system allows to search for words, phrases, grammatical forms and different meanings of words and present them in a user-friendly format. There are several types of corpus, such as national corpus, author corpus, book corpus, newspaper corpus, oral corpus, and the purpose of this article is the development of a corpus based on the texts of a fiction writing.

1. Stages of corpus creation

Methods elaborated for the function of the language corpus and the organization of the corpus workflow usually lead to improved quality of the results obtained in the corpus, as a result of the production of flawless and sustainable development (Grosman, 2020).

The development of a corpus might be considered as a technological process that is carried out in several stages. The most important steps are tokenization, markup, and conversion of the markup texts into the format of the corpus control software.

There are many years of research in natural language processing and computational linguistics, and tools based on this research have been developed, and many algorithms and software tools for text analysis, such as lemmatizers, tokenizers, triggers, tags, etc. Software tools and methodologies have also been developed for specific tasks, for instance, analysis of ideas, translation, collection and analysis of target information (Kuhn? 2019).

1.1. Tokenization

The first task of tokenization is to identify and distinguish lexical units in a text. Determining the boundaries of text and sentences is an important processing step in language processing programs since most of them work at the word (morphological analysis) or sentence level (disassembly, parsing, machine translation). The term token is used to denote lexical units. The key problems of the tokenization function are related to the lack of words in alphabetical order, some undefined characters and well-defined word boundaries in special writing systems (Jurish). In general, tokens have the same meaning as word forms. However, the term “token” is used to refer to lexical units, not word, because in some cases units smaller than a word (separate morpheme) or units larger than a word (word combinations) can also be used as tokens. The source text will be divided into components in tokenization (tokens: words, punctuation marks, dates, etc.). The software that performs this algorithm is called a tokenizer, A tokenizer called UzbekToken has been realized for texts in the Uzbek language. First the text is split into words based on white spaces, then punctuation marks will be removed from the words. Abbreviations “e.g. TUIT, UN, CIS, etc.” and date entries “e.g. 09.04.2018” will be also taken as tokens. The list of abbreviations that might be found in the text is created in advance and entered into a separate file. Likewise, the list of fixed phrases and idioms that are tokens will also be saved in a separate file, for example, finally, therefore, etc. These files will be considered in extraction tokens from texts. The results of the tokenizer are presented in Table 1:

Table 1. The results of the tokenizer. Результаты токенизатора.

| № | Source text | List of tokens |
|---|---|---|
| 1 | O'zbekiston Respublikasi 02.03.1992 kuni BMTga a'zo bo'lgan | O'zbekiston Respublikasi 02.03.1992 kuni BMTga a'zo bo'lgan |
| 2 | TATU Samarqand filiali 2005 yilda o'z faoliyatini boshladi | TATU Samarqand filiali 2005 yilda o'z faoliyatini boshladi |

1.2. Marking

Marking up the text is the most basic stage; as a result of its work, the corpus will be formed. Marking is the process of attaching special tags to text and its components. There are two types of special tags: linguistic tags and extralinguistic 'external' tags. Linguistic tags contain information that describes the lexical, grammatical, and other characteristics of text elements. The types of linguistic markup can be: morphological markup, syntactic markup, semantic markup, etc. (Nicolaev, 2016). Extralinguistic tags describe information about the author, attitude, year and place of publication, genre, subject in addition to the language itself (Kamaruddin).

Marking in proportion to the enormous dimensions of the case is a long and laborious process. Therefore, this should be performed automatically. If any marking is not yet automated, then you will have to do it manually.

For morphological labeling, a software tool called *tegger* was developed that allows most of the work to be done automatically. As a work of the work of the software of automatic morphological markup 'tagger', each lexical unit, or rather each word form, is given a grammatical characteristic (part of speech, the numeral, case, tense, "person, etc."). In foreign terminology, the term part-of-speech tagging 'POS-tagging' is used (Li). In fact, the morphological sign includes not only the sign of the part of speech, but also signs of the grammatical categories regarding to this part of speech, and the normal word form as lemma.

This can be considered as the main type of marking, because, first, morphological analysis is better automated than others, secondly, it serves as the basis for the subsequent stages of analysis, that is, the processes of syntactic and semantic analysis.

The markup information will be presented as a structure. The morphological markup of a fragment of the Uzbek text *Boysari farzandli bo'ldi. Shohimardon bilan tunab qoldi* is presented in Table 2.

Table 2. Marking words: XML-form.


```

<?xml version="1.0" encoding="windows-1251"?><text><p>
<s>
<w> boysari <ana lemma="BOYSARI" pos="[ot],[at.ot],[shaxs_n.]"/> </w>
<w> farzandli <ana lemma="FARZAND" pos="[ot],[tur.ot],[qar.LMG],[shaxs.ot],
[b.k.],[birl.son],[3-sh.]"/> </w>
<w> bo`ldi <ana lemma="BO`L" pos="[f],[must.f],[har.f],[14.2]"/> </w></s>
<s>
<w> shohimardon <ana lemma="SHOHIMARDON" pos="[ot],[at.
ot.],[shaxs_n.]"/> </w>
<w> bilan <ana lemma="BILAN" pos="[ko`m.],[sof ko`m.],[vos.m.]"/> </w>
<w> tunab <ana lemma="TUNAB" pos="[f],[must.f],[har.f],[o`-siz f.],[b-li
f.],[an.n.],[sod.f.],[t.f.],[rav-sh]"/> </w>
<w> qoldi <ana lemma="QOL" pos="[f],[must.f],[holat.f.],[o`-siz f.],[b-li f.],
[an.n.],[sod.f.],[t.f.],[x.m.],[o`.z.],[III sh.b.]"/> </w>
</s></p></text>

```

The following tags have been used in these records: <text> - text, <p> - paragraph, <s> - sentence, <w> - using a word, <ana> - attributes, <lemma> - lemma, <pos> - word markup.

Syntactic markup is performed based on data from morphological analyzes. This type of markup depends on the accepted formal syntax model, it describes the syntactic relationships between lexical units and various syntactic constructions "subordinate clause, predicate and etc". It syntactically combines discrete units, words, phrases, sentences in accordance with the rules (Aronoff, 2008). To mark up syntax, it is required to develop software tools called a parser. As a result of the work of the automatic syntactic markup program, syntactic links between words and phrases will be fixed, information about its characteristics "sentence type, phrase syntactic function, etc." will be attached to the syntax units.

Unlike morphology, syntax does not have separate methods of describing structures and syntactic relations. Typically, markup syntax is performed in grammar links or directly in terms of grammar compound components.

Semantic markup. Semantic tags define the semantic categories to which a word or phrase belongs and the narrow subdivision categories that represent its meaning (Penev, 2010). Semantic markup of the corpus requires preliminary specialization of word meanings, solution of ambiguity and ambiguity, categorization of words, separation of thematic classes, evaluation and characteristics of derivation.

1.3. Corpus control system

If special software – corpus manager exists, the text corpus becomes a powerful tool in the hands of a linguist.

A modern corpus manager should be able to following:

- concordance formation;
- search for contexts not only by words, but also by phrases;
- sort lists by several criteria selected by the user;
- providing the ability to describe the found word forms in an extended context;
- provide statistics on individual corpus elements;
- save and print results;
- the ability to work not only with individual files, but also with an unlimited number of corpuses;
- quickly respond to inquiries and get results quickly;
- be convenient and understandable for both beginners and experienced users.

Modern corpus systems should be able to solve more complex problems, such as compiling frequency dictionaries of words, compiling lists of colloquial expressions 'ideoms', and forming lexical-semantic groups.

The corpus of the Uzbek language has been developed uzbekcorpora.uz and preliminary results were obtained. according to the preliminary result, the epic “Alpomish” has been included in the database.

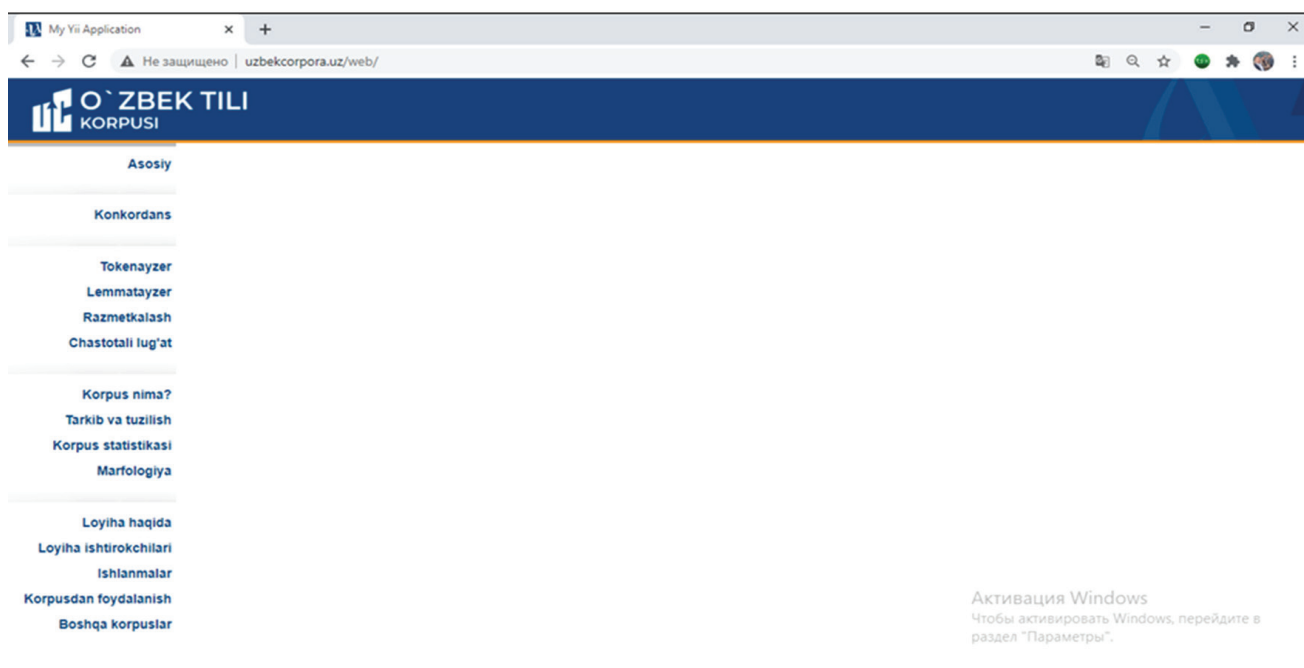


Fig. 1. Uzbek language corpus interface

There is a search function for the Concordans corps item in the Corpus Manager menu. The corpus search interface is illustrated in the Fig. 2.

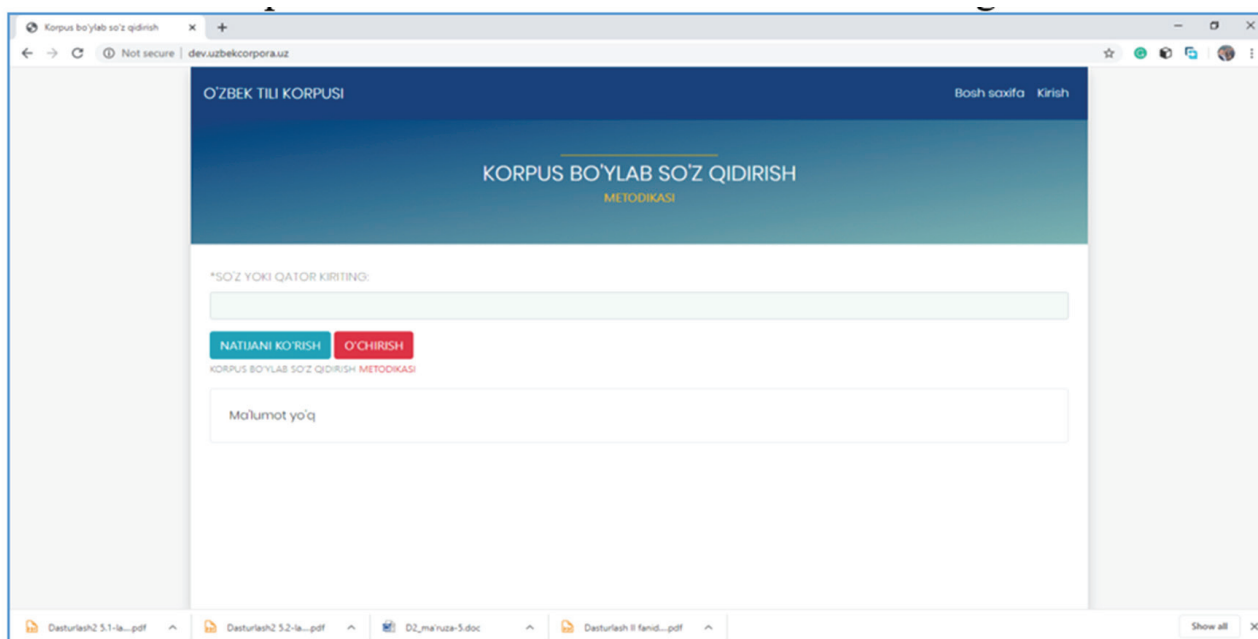


Fig. 2. Concordance

There is a function to search for a word or phrase throughout the entire corpus in the Concordance window. It provides a field for entering the desired word or phrase and a button to get the search result.



Fig. 3. Search word entry window

Information will be entered to a space designed to search for a word from the corpus and ‘View results’ button will be used to view results. As an experiment, the Alpomish epic was loaded into the corpus to obtain very first results. Boybo'ri will be entered in the blank field to get result and ‘View results’ button will be pressed. In the window below ‘figure 4.’ can be seen that Boyb'ori is included in the 3 paragraphs of the text file in the corpus.



Fig. 4. Results of search

Searching for the word Boybo'ri will return the number of paragraphs that contain the word and the paragraph itself.

By clicking on an arbitrary word in the resulting paragraphs, a lemma and a morphological word label are formed on the screen. ‘Fig. 5, 6, 7’.

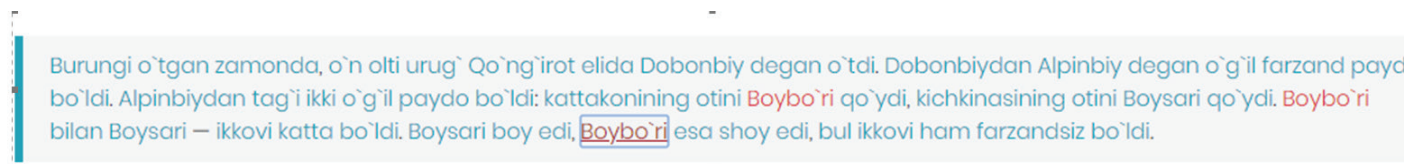


Fig. 5. Paragraph appearing in search results

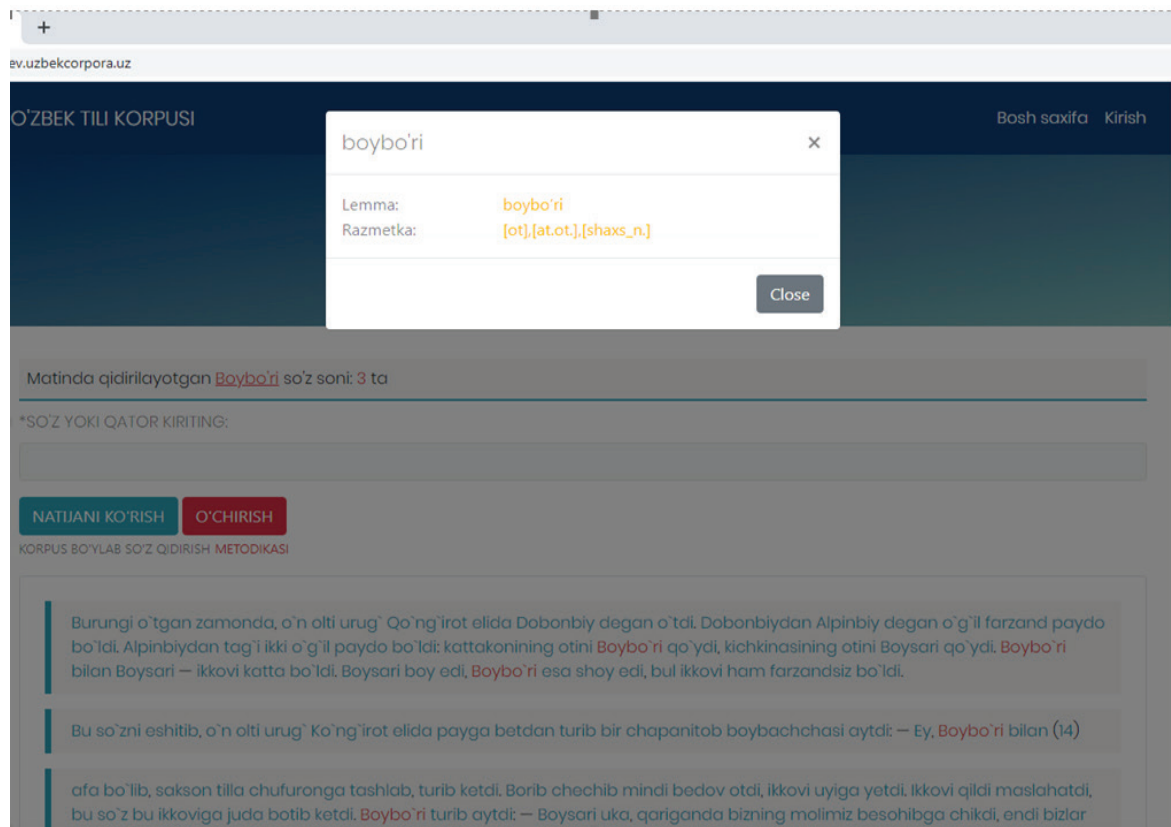


Fig. 6. Boybo'ri has been selected and lemma and markup of the word

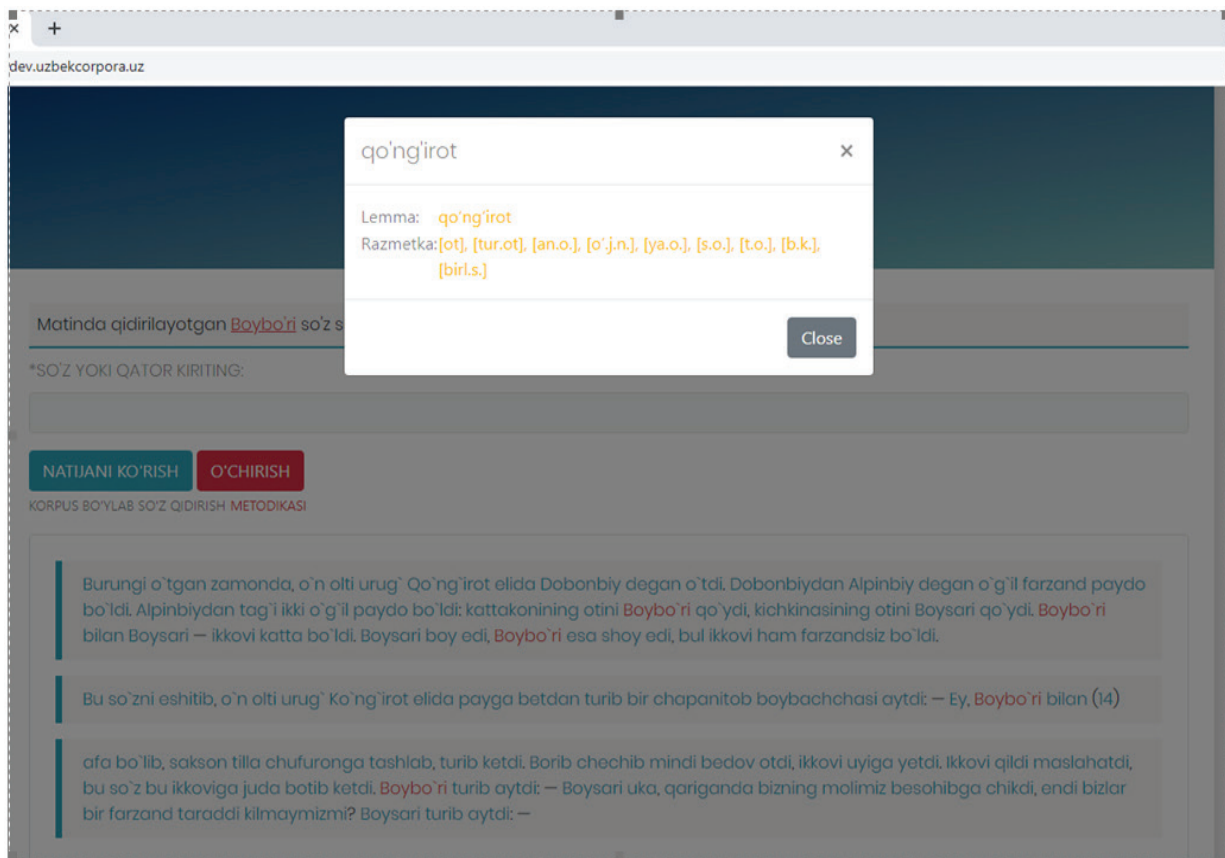


Fig. 7. Qo'ng'iro't has been selected and lemma and markup of the word

The term ‘corpus quality’ varies depending on the intended use of the corpus. In empirically oriented theoretical linguistics, carefully chosen procedures and cleaning up indestructible redundant messages are important, however for many computational linguistics and language technology tasks, aggressive corpus cleansing is necessary to achieve good results. Technically, these differences are not only related to software architecture, but also to different software configurations. It is also important to focus on increasing the processing speed and data collection across all corpus-oriented projects (Biemann, 2013).

To better understand the general principles of how corpus managers work, we'll look at a few of them. Tony McEnerly and Andrew Hardy have described four generations of corpus control softwares. They believed that most modern corpus linguistics software was third generation. These software systems offer a wide range of functions including statistical methods, multilingual accounting and a user-friendly interface. Examples of such software include WordSmith Tools, MonoConc Pro, ParaConc, AntConc (McEnerly, 2012).

The AntConc system is often used for small corpuses. It is free software with a user-friendly interface and many functions for handling various encodings and text formats. It is possible to perform followings by using AntConc:

- view a text file;
- forming a concordance in the context window for a word;
- draw graphs for concordance;
- splitting the n-gram of a word participating in the context window frame;
- distinguishing collocates of a word based on the association measure;
- forming lists of colors or absolute frequencies of words or lemmas;
- highlight keywords.

Its major drawback is that it does not work well with large corpuses. (Anthony, 2005).

4th generation corpus software includes Sketch Engine, CQPweb, Wmatrix, Mark Davis corpus system, etc.

The main innovative feature of CQPweb is its flexibility, its more generalized data model makes it suitable for any corpus. The following analysis options are available on CQPweb: concordation, collocation, distribution of tables and parts, frequency lists, keywords and tags (Hardie, 2012).

Wmatrix is essentially a comparative stylistic analysis process. An automatic method of comparing text corpuses to describe their «differences» is provided by the Wmatrix software developed by Paul Ryson (Leech).

By These software tools are capable of handling large amounts of data. They store corpuses in a web server database and the data is pre-indexed to speed up searches.

Many linguistic functions have been implemented in the Sketch Engine, created under the direction of Adam Kilgarrif at Masaryk University in the Czech Republic. This system is developed for the Sketch Engine corpus service use on a subscription basis. Currently, the service can provide users with more than 350 corpora in different languages of the world. The Sketch Engine system performs the separation of phrases in the corpus that indicate the strength of connections between phrases ‘components’, as well as the performance of itself, taking into account syntactic formulas “for example, a verb + correct addition”, construction of lexical-semantic fields for a given word, comparison of connectivity synonyms, separation of keywords. The core of this system is distributed free of charge by NoSketch Engine. This manager is used for many national corpuses “Czech, Slovak, Bulgarian, Croatian, etc.”. This program loads the previously marked texts and creates a corpus based on them. The work with the corpus data can be carried out in several modes: concordance mode for a given word or phrase or formation of frequency lists (Pavel, 2007).

Conclusion

The corpus of the Uzbek language will be created with the help of Uzbek philologists, researchers, information technology specialists and programmers. The construction of the corpus is carried out in strict accordance with the originally received standards. The Uzbek Language Corpus is an important resource for researchers interested in the linguistics of the Uzbek language.

References

- Aronoff M., Meir I., Padden C. and Sandler W. The Roots of Linguistics Organization in a New Language. Published in final edited form as: *Interact Stud.* 2008; 9(1): 133–153. doi: 10.1075/is.9.1.10aro.
- Anthony, L. “AntConc: Design and Development of a Freeware Corpus Analysis Toolkit for the Technical Writing Classroom.” 2005 IEEE International Professional Communication Conference Proceedings.
- Biemann Ch. and Bildhauer F. “Scalable Construction of High-Quality Web Corpora.” *The Journal for Language Technology and Computational Linguistics (JLCL 2013 – Band 28 (2))*.
- Grosman, J. S., Furtado P. H. T. and Rodrigues, A. M. B. Eras : Improving the quality control in the annotation process for Natural Language Processing tasks <https://www.journals.elsevier.com/information-systems> 93 (2020) 101553.
- Jurish, B. and Würzner K.M. Word and Sentence Tokenization with Hidden Markov Models. *The Journal for Language Technology and Computational Linguistics (JLCL 2013 – Band 28 (2) – 61-83)*.
- Kamaruddin R., Ahmad N.A., Saad J.M., Kamaruddin M., Seruji Z. Attitudes as extra linguistic factor in learning second language among foreign students in 5 universities in Malaysia. *International Journal of Asian Social Science*.
- Kamaruddin J. Performance Evaluation of English Part of Speech Tagging Based on Multi-feature Knowledge Algorithm. *Journal of Physics: Conference Series*.
- Kuhn, J. Computational text analysis within the Humanities: How to combine working practices from the contributing fields? *Lang Resources & Evaluation* (2019) 53:565–602.
- Nikolaev I., Mitrenina O., Lando T/ «Прикладная и компьютерная лингвистика». 146-бет, URSS научная книга.
- McEnery T., Hardie A. “Corpus linguistics: method, theory and practice.” Cambridge University Press, 2012.
- Leech G. “Virginia Woolf meets Wmatrix,” <https://journals.openedition.org/esa/1405>, 4/2013.
- Hardie A., “CQPweb — combining power, flexibility and usability in a corpus analysis tool.” *International Journal of Corpus Linguistics* 17:3 (2012), 380–409.
- Pavel, R. “Manatee/Bonito — A Modular Corpus Manager // 1st Workshop on Recent Advances in Slavonic Natural Language Processing.” Brno: Masaryk University, 2007. pp. 65-70.
- Penev L., Agosti D., Georgiev T., Catapano T. and Miller J. Semantic tagging of and semantic enhancements to systematics papers: ZooKeys working examples. Published online 2010 Jun 30. doi: 10.3897/zookeys.50.538.

For citation:

Tursunov M. S. Development of software for a Text corpus in Uzbek. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 09.09.2021.

UZBECORPORA.UZ: SOFTWARE OF THE UZBEK LANGUAGE CORPUS

UZBECORPORA.UZ: ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
КОРПУСА УЗБЕКСКОГО ЯЗЫКА

Каршиев Абдували Беркинович, профессор кафедры программной инженерии, Самаркандский филиал Ташкентского Университета информационных технологий имени Мухаммеда аль-Хорезми, Самарканд, Узбекистан.

Karshiev Abduvali Berkinovich, Professor of the Department of Software Engineering Samarkand Branch of the Tashkent University of Information Technologies named after Muhammad al Khwarezmi Samarkand, Uzbekistan.

ORCID: 0000-0001-6121-9928



Турсунов Мухаммадсолих Саъдин угли, ассистент кафедры программной инженерии, Самаркандский филиал Ташкентского Университета информационных технологий имени Мухаммеда аль-Хорезми, Самарканд, Узбекистан.

Tursunov Mukhammadsolikh Sadin ugli, Assistant of the Department of Software Engineering Samarkand Branch of the Tashkent University of Information Technologies named after Muhammad al Khwarezmi, Samarkand, Uzbekistan.

muhammadsolih927@ gmail.com

ORCID: 0000-0002-6485-3630



Холмухамедов Бахтиер Фарходович, доктор филологических наук (PhD), доцент кафедры узбекской лингвистики, Самаркандский государственный университет, Самарканд, Узбекистан

Kholmukhamedov Bakhtiyor Farkhodovich, Doctor of Philology (PhD), Associate Professor of the Department of Uzbek linguistics, Samarkand State University, Samarkand, Uzbekistan.

ORCID: 0000-0002-8546-8192



В статье описывается система *uzbekcorpora.uz*, созданная для узбекского языкового национального корпуса. Программное обеспечение корпуса состоит из таких компонентов, как поиск текста по словам и фразам (согласование), разметка (характеристики слов), леммы, токены и частотные словари. Также имеется административный раздел для редактирования базы данных корпуса. Программа «*Uzbekcorpora.uz*» – это бесплатная онлайн-платформа, позволяющая изучать узбекский язык на любом компьютере с любой точки, где доступен Интернет.

Ключевые слова: *uzbekcorpora.uz*, узбекский язык, токенизатор, лемматизатор, маркировка, согласование, частотные словари, изучение узбекского языка.

This article describes the system *Uzbekcorpora.uz* for the Uzbek language national corpus. The software consists of such components as body search for words and phrases (concordance), markup (word characteristics), lemmas, tokens, and frequency dictionaries. One of the fundamental algorithms of lexical analysis is to divide a given text into tokens. The program that implements this algorithm is called a tokenizer. Tokens usually have the same meaning as word forms. However, the term ‘toke’ is used to describe lexical units, not ‘words’. This is because in some cases, units smaller than a word (a separate morpheme) or units larger than a word (phrases) can be used as tokens. Marking is the process of attaching special tags to text and its components. There are two types of custom tags: linguistic tags and extralinguistic ‘external’ tags. The concordant tool has a number of features that make *uzbekcorpora.uz* an effective tool for Uzbek language learners and researchers: it includes such options (buttons) as admin, word you are looking for, paragraphs, words, statistics. There is also a section to administrate or edit the corpus database. *Uzbekcorpora.uz* software is a free online platform. It allows the learner to work anywhere, on any computer, while researching a language. The construction of the corpus is carried out in strict accordance with the originally received standards. The Uzbek Language Corpus is an important resource for researchers interested in the linguistics of the Uzbek language.

Keywords: *uzbekcorpora.uz*, Uzbek language, tokenizer, lemmatizer, marking, concordance, frequency dictionaries, Uzbek language training.

Introduction

Recently, software has been developed for teaching languages, automating data in obtaining and researching various statistical analyses on languages, and working with large amounts of data. In such situation, the creation of language corpus is very important. The “corpus revolution” took place, as Chappelle described it. Corpus began to be widely used in many fields, including translation, stylistics, grammar, and dictionary creation (Chappelle, 2001). Johns proposed the use of corpora for independent study of how language is used in target contexts in language learning, i.e., an approach to language learning called “data management” (Johns, 1997). Learners will need to use a corpus manager and a corpus search engine to find the information they need when studying language. It allows them to understand the features of the language and work on them. Due to the fact that language characteristics apply to different fields, the above suggestion can be very effective in technical writing lessons (Laurence Anthony, 2005).

1. The main findings and results

Without software that can process data and present results in an understandable way, the language corpus is of no use.

The software consists of such components as concordancing words, phrases, markup, wording, tokens, and frequency dictionaries.

1.1. The main window of the system. To work with the Uzbek National Corpus, visit <http://uzbekcorpora.uz/> using a browser. When you enter the site, the interface shown in Figure 1 is created. The left part of this interface consists of a menu bar that provides general information about the case, project information, and software (Johns, 1997).

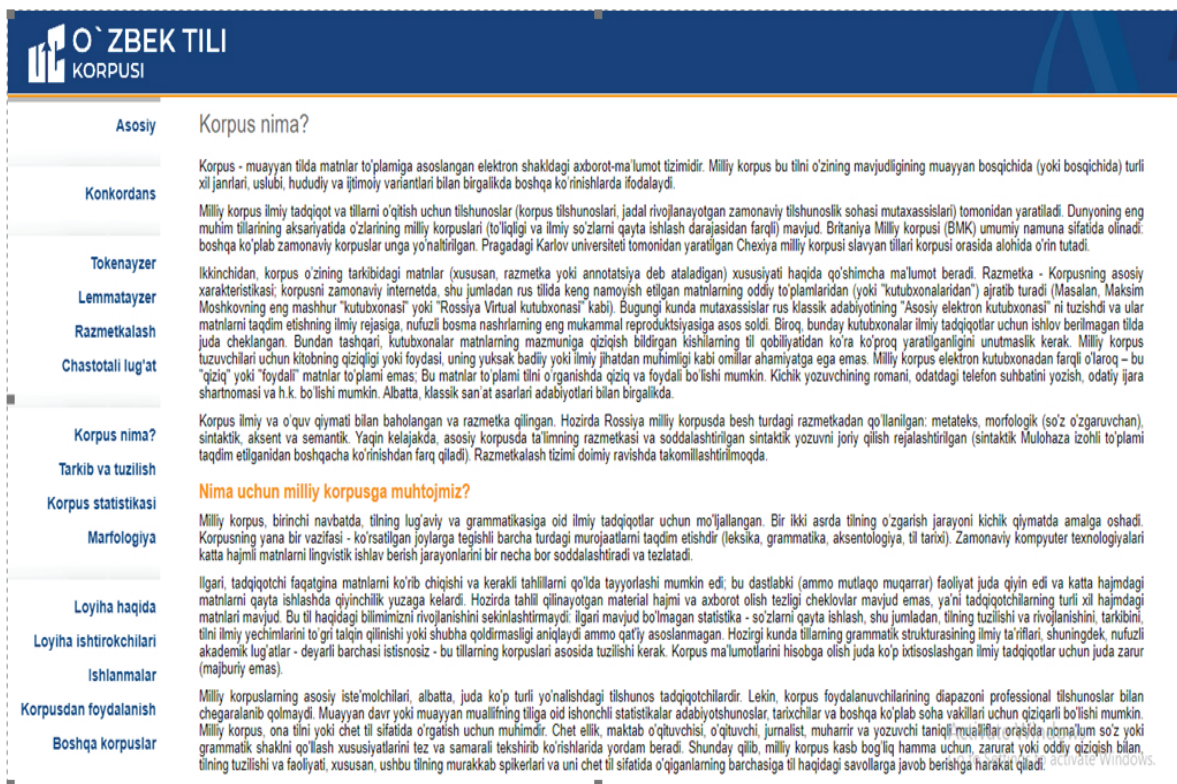


Fig. 1. Main interface

Main - this menu contains information about the national corpus of the Uzbek language;

Concordance - This menu will take you to the search engine of the corpus. This system is described in detail in a paragraph entitled concordance;

Tokenizer - Go to the text token program via this menu. This program is described in detail in the paragraph entitled Tokenizer;

Lemmatizer - through this menu you can go to the text lemma program. This program is described in detail in the paragraph entitled Lemmatizer;

Marking - Go to this program where you can see the marking of words through this menu. This program is described in detail in the section entitled Marking;

Frequency dictionary - through this menu you can go to the program of statistical analysis of texts. This program is described in detail in a paragraph entitled Frequency Dictionary;

What is a corpus? - in this menu there are descriptions of the language corpus, information about the corpus;

Content and structure - this menu shows the structure of the national corpus of the Uzbek language;

Corpus statistics - this menu contains information on the number of texts, the number of sentences, the number of words entered in the language corpus database;

Morphology - this menu contains information on morphological markings and morphological marking standards in the language corpus;

About the project - this menu contains information on the creation of the national corpus of the Uzbek language;

Project participants - this menu contains information about the staff involved in the creation of the national corpus of the Uzbek language;

Developments - this menu contains information on articles published in the process of creating the national corpus of the Uzbek language, software certificates;

Use of the corpus - this menu contains information on the use of the national corpus of the Uzbek language;

Other corpora - this menu contains information about the national corpus of other countries (Laurence Anthony, 2005);

1.2. Tokenizer.

The first step in creating a language corpus is to understand how to distinguish words in a text. To do this, all subsections that do not contain punctuation marks “spaces, punctuation marks, etc.” must be separated from the text. This will be a collection of tokens (Nikolaev, 2016). The input of the algorithm is given a text, and the output is a list of lexical units in the text. One of the fundamental algorithms of lexical analysis is to divide a given text into tokens. The program that implements this algorithm is called a tokenizer. Tokens usually have the same meaning as word forms. However, the term ‘toke’ is used to describe lexical units, not ‘words’. This is because in some cases, units smaller than a word (a separate morpheme) or units larger than a word (phrases) can be used as tokens.

The tokenizer first separates the text based on the spaces between the words, then removes the punctuation marks from the words. Abbreviations “TUIT, UN, CIS, etc.” and date entry “09.04.2018” are also taken as tokens (Tursunov, Garshiev, Karimov).

Lexical decomposition is fundamental for automatic analysis of text because it serves as the basis for a number of other algorithms.

Clicking on the tokenizer from the menu bar in Figure 1 creates a tokenizer window (Fig.2).

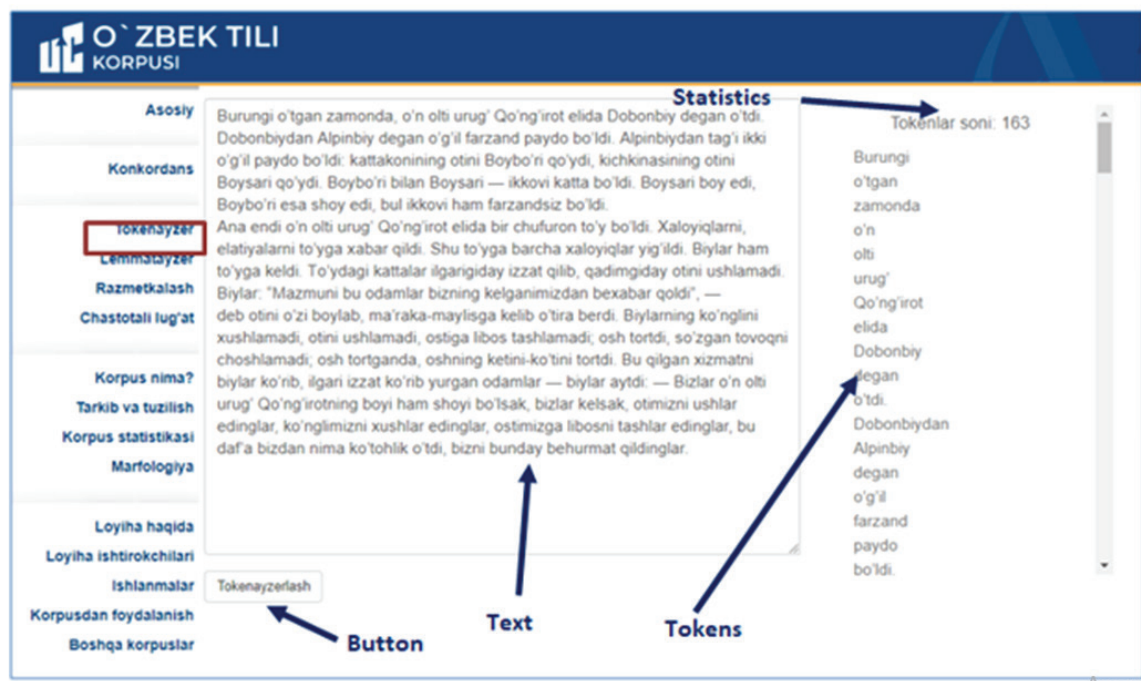


Fig. 2. Tokenizer interface

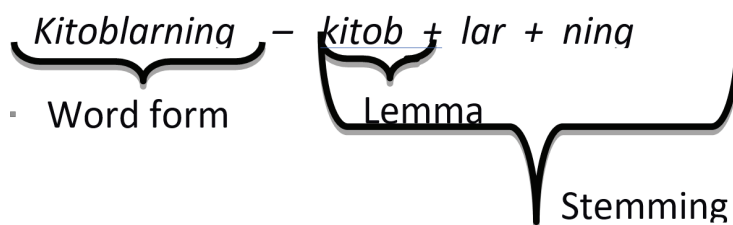
The case includes a tokenizer program as part of the program. This allows the user to token their text. The tokenizer program is divided into four parts, as shown in Figure 4, and they represent the following functions:

- **Text** is the input that the user enters in the field;
- **Button** - this button, when clicked with the mouse, creates the entered text into tokens;
- **Tokens** - this field displays a list of tokens extracted from the text;
- **Statistics** - where the amount of tokens in the text is indicated (Tursunov, Garshiev, Karimov).

1.3. Lemmatayzer.

In morphological designation, the approach that can show the core of a given word form is called lemmatization. While the division of a word form into parts is called stemming, the core part of a separated word form is called a lemma (Nikolaev, 2016).

For example:



From the word form, the program that separates the lemma of a word is called a lemmatizer. Clicking on the Lemmatizer item in the menu bar in Figure 1 will create a Lemmatizer window (Fig. 3).

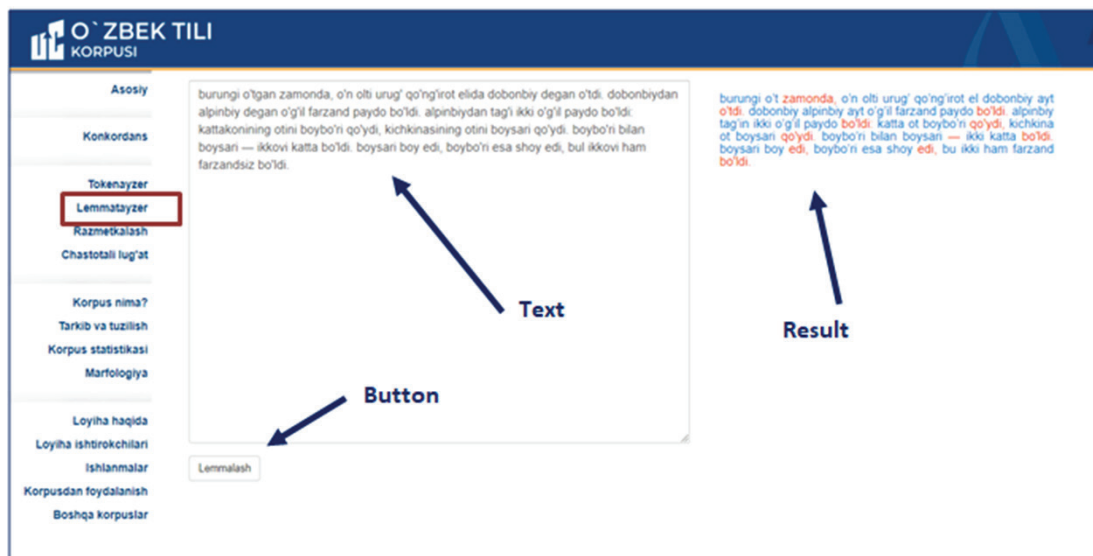


Fig.3. Lemmatizer interface

In addition to identifying the lemma of a word in a search engine, the corpus also includes the lemmatizer program as a subroutine. This allows the user to type their text. The Lemmatizer program is divided into three parts, as shown in Figure 5, and they represent the following functions:

- **Text** is the input that the user enters in the field;
- **Button** - this button identifies the lemmas of the given words when pressed with the mouse;
- **Result** - this is where the text of the program is generated as a result of the program's work. The resulting text shown in Figure 5 is shown in two different colors (blue and red). The blue words are lemma and the red words are not lemma. Therefore, the red word and its lemma should be entered into the program database (Tursunov, Garshiev, Karimov).

1.4. Marking.

Marking is the process of attaching special tags to text and its components. There are two types of custom tags: linguistic tags and extralinguistic 'external' tags. Linguistic tags contain information that describes the lexical, grammatical, and other characteristics of text elements. Extralinguistic tags describe information about the author and the text "author, title, year and place of publication, genre, tenacity, etc." (Nikolaev, 2016).

Marking in proportion to the size of the case is a time-consuming and labor-intensive process. Therefore, it should be done automatically. If any type of marking is not yet automated, then this work has to be done manually (Tursunov, Garshiev, Karimov).

Clicking on the menu item in the menu bar in Figure 1 will create a selection window (Fig. 4).

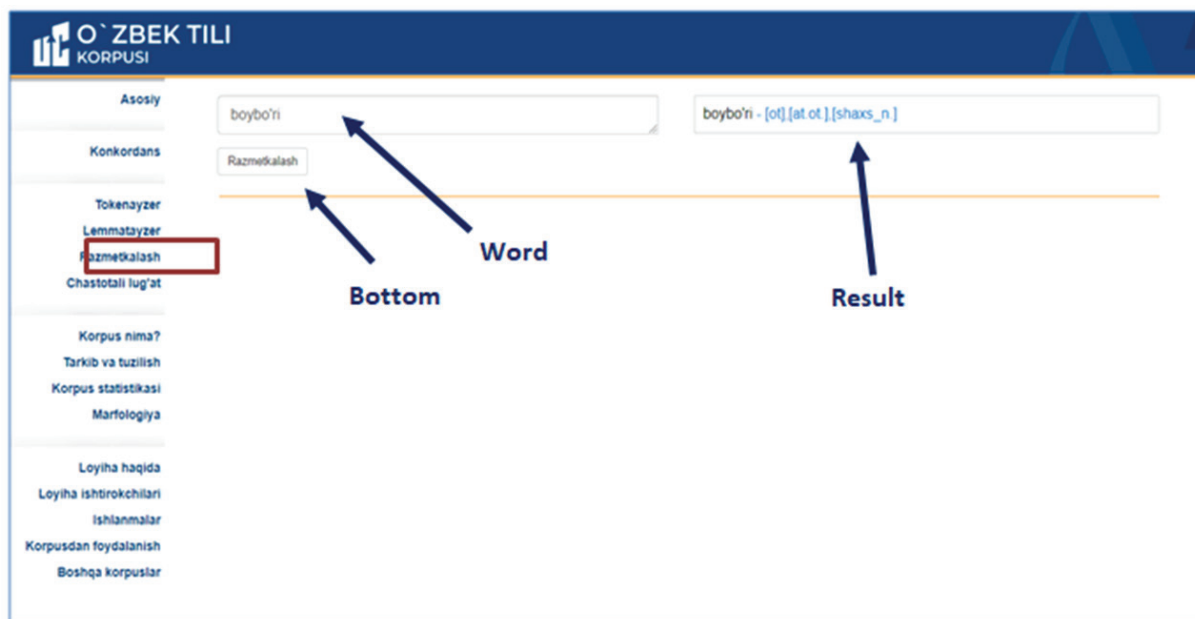


Fig. 4. Marking interface

In addition to identifying the word mark in the search engine, the case also includes the marking program as a partial program. This allows the user to determine the margin of the word they are researching. The layout program is divided into three parts, as shown in Figure 6, and they represent the following functions:

- A word is an incoming piece of information in which a user places a word they are learning;
- Button - identifies the tags of the entered word when the mouse is pressed on it;
- The result is a word on the left side of the field and a check mark for that word on the right. As shown in Figure 6, the mark of the given word consists of three tags (Tursunov, Garshiev, Karimov).

1.5. Concordance. Like other buildings, one of the main functions of the system *uzbekcorpora.uz* is a search system throughout the building. It provides information about the properties of the word being searched. When the concord button is pressed from the menu bar of the main window (see Figure 1), the interface in Figure 2 opens. A simple interface has been developed to make it easier for users to use the system, and all its functions are displayed in a single window (Tursunov, Garshiev, Karimov).



Fig. 5. Concordance interface

The concordant tool shown in Figure 5 has a number of features that make *uzbekcorpora.uz* an effective tool for Uzbek language learners and researchers.

These features can be summarized as follows:

1. **Admin** - this button can be used to access the database by the system administrator or users authorized by the administrator. It has the ability to manage the database (add, edit, delete data).

2. **The word you are looking for** is the first task in using the field search engine; enter the word you are looking for in this field and click the “View result” button. The “delete” button clears the field of the word you are looking for.

3. **Paragraphs** - when a word is entered in the data entry field and the “see result” button is pressed, as a result, the paragraphs with the word under study appear. Each paragraph is highlighted.

4. **Words** - When you hover the mouse over the words in the resulting paragraphs, a line will appear at the bottom of the word on which the mouse is above. In this case, when the mouse is over the word, the word is tapped, and when the mouse is clicked on the word, the properties of the word (word lemma and mark) are formed, as shown in Figure 6. The word being studied appears in a different color than other words.

5. **Statistics** - when the word under study is entered in the field and the “see result” button is pressed, two statistical analyzes appear in the analysis field:

- The number of paragraphs found as a result of the search, in which the word under study is involved;
- The amount of the word under study in the texts available in the database (Baranov, 2001).

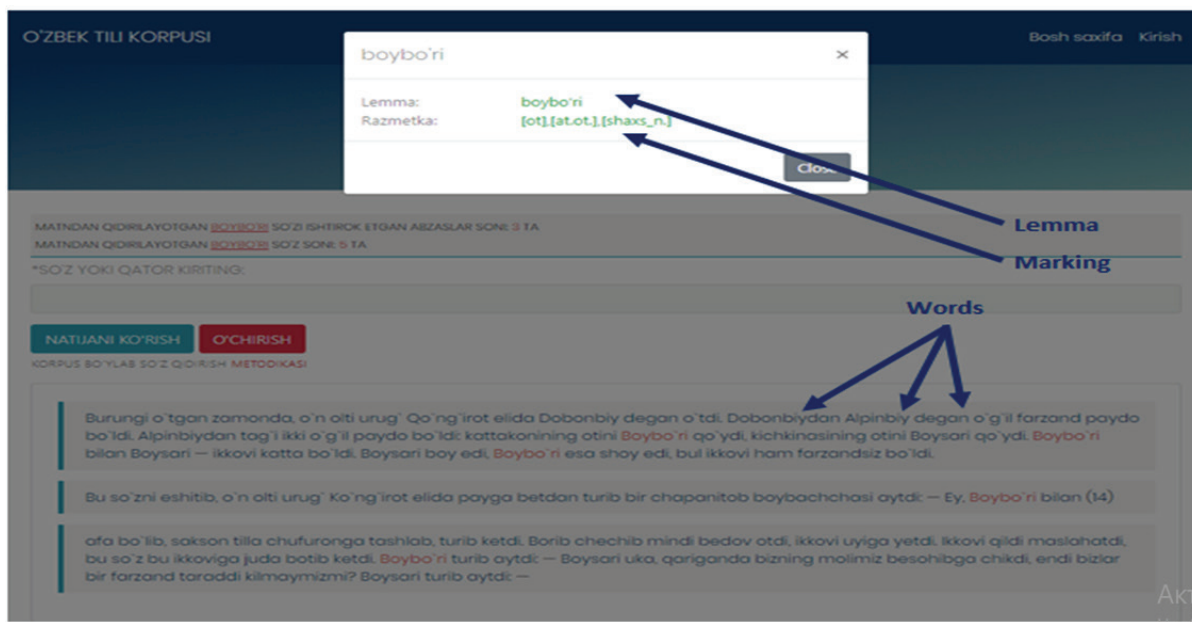


Fig. 6. The lemma and mark of the word.

The program is shown in Figure 5 as a result of the work. The word *boybo'ri* is entered in the field. When the “see result” button is clicked, the paragraphs containing the word being studied from the database (*boybo'ri*) are generated as a result. When you hover the mouse over the word *boybo'ri* and click on the word, the information in Figure 6 is created. In Figure 6, the lemma and mark of the word *boybo'ri* are formed (Gvishiani, 2008).

Table 1. Result.

| Searched word | The lemma of the word | Word markup |
|---------------|-----------------------|--|
| boybo'ri | boybo'ri | (ot),(at.ot),(shaxs_n.) (noun), (at.ot), (person_n.) |

The mark of the word under study (*boybo'ri*) consists of 3 tags, which are named as follows:

(*noun*) - this tag is part of the word group of the word under study;

(*at.ot*) - this tag indicates that the word under study belongs to a group of famous horses in the horse word category;

(*person_n.*) - this tag indicates that one of the characters present in the group of nouns in the word group of the word under study is the name of the person (McEnery, Hardie, 2012).

Conclusion

The construction of the corpus is carried out in strict accordance with the originally received standards. The Uzbek Language Corpus is an important resource for researchers interested in the linguistics of the Uzbek language.

References

1. Laurence Anthony, AntConc: Design and Development of a Freeware Corpus Analysis Toolkit for the Technical Writing Classroom, 2005 IEEE International Professional Communication Conference Proceedings, pp.729-737

2. Chapelle, C. A. Computer applications in second language acquisition: Foundations for teaching, testing, and research. Cambridge, England: Cambridge University Press, 2001.

3. Johns, T. "Contexts: the Background, Development and Trialling of a Concordancebased CALL Program," in Teaching and Language Corpora, London, England: Longman, 1997, pp. 100-115

4. Nikolaev I. S., Mitrenina O. V., Lando T. M. (ed.) (2016) Applied and Computational Linguistics, Lenand. – p. 320. - ISBN 978-5-9710-3472-8 (Николаев И.С., Митренина О.В., Ландо Т.М. (ред.) Прикладная и компьютерная лингвистика, Ленанд, 2016. 320 с. ISBN 978-5-9710-3472-8)

5. Tursunov M.S., Garshiev A.B., Karimov S.A. Development of a modern corpus of computational linguistics. Available at: https://www.scopus.com/record/display.uri?eid=2-s2.0-85102147010&origin=AuthorNamesList&tx_Gid=13070f5e69ea0839e04674499810efe1, DOI: 10.1109/ICISCT50599.20.20.9351376 (accessed: 01.06.2021.).

6. Baranov A.N. (2001) Corpus linguistics. Baranov A.N. Introduction to Applied Linguistics. Moscow. pp.112-137. (Баранов А.Н., Корпусная лингвистика // Баранов А. Н. Введение в прикладную лингвистику. Moscow, 2001. С.112–137.)

7. Gvishiani, N.B. (2008) Workshop on corpus linguistics: Textbook. English manual. Moscow. (Гвишиани Н.Б., Практикум по корпусной лингвистике : Учеб. пособие по английскому языку. Москва : 2008.)

8. McEnery, T., Hardie A. (2012) Corpus linguistics: method, theory and practice. Cambridge University Press.

For citation:

Karshiev A. B., Tursunov M. S., Kholmukhamedov B. UZBEKCORPORA.UZ: software of the Uzbek language corpus. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 09.09.2021.

СИНОНИМАЙЗЕР (СИНОНИМИЗАТОР) В ОБРАЗОВАТЕЛЬНОМ КОРПУСЕ УЗБЕКСКОГО ЯЗЫКА

SYNONYMIZER IN THE EDUCATIONAL CORPUS OF THE UZBEK LANGUAGE

Абжалова Манзура Абдурашетовна, кандидат наук по филологии (PhD). Ташкентский государственный университет узбекского языка и литературы имени Алишера Навои, Узбекистан.

Abjalova Manzura Abdurashetovna, Doctor of Philosophy in Philology (PhD). Tashkent State University of Uzbek Language and Literature named after Alisher Navoi, Uzbekistan.

abjalova.manzura@gmail.com
ORCID: 0000-0002-1927-2669



Сегодня изучение и преподавание языков с использованием цифровых технологий является актуальным вопросом. В связи с этим создаются лингводидактические платформы, такие как различные словари, мобильные приложения и компьютерные программы. Образовательный корпус узбекского языка, созданный с учетом потребности в современных технологических средствах языкового образования, решает такие задачи, как повышение эффективности обучения, предоставление удобных дидактических возможностей, содействие развитию словарного запаса при изучении языка. То, что это корпус служит единой образовательной лингво-информационной системой, определяется наличием учебников в его базе данных. В базу данных также входит словарь синонимов, что позволяет корпусу выступать в качестве синонимизатора. В данной статье рассматривается синонимайзер, принцип его работы и начальный процесс работы синонимайзера в узбекском образовательном корпусе.

Ключевые слова: узбекский язык, учебный корпус, синонимайзер, синонимизатор, синонимы.

Today, learning and teaching languages is a global problem. Therefore, linguodidactic platforms are being created, such as various dictionaries, mobile applications and computer programs. The educational corpus of the Uzbek language, created taking into account the need for modern technological means of language education, solves such problems as increasing the effectiveness of learning, providing convenient didactic opportunities, promoting the development of vocabulary in learning the language. The fact that this corpus serves as a unified educational linguistic information system is determined by the presence of textbooks in its database. The database also includes a synonym dictionary, which allows the corpus to act as a synonymizer. This article discusses the synonymizer, the principle of its operation and the initial process of the synonymizer operation in the Uzbek educational building.

Key words: Uzbek language, educational corpus, synonymizer, synonyms.

Введение

Узбекский язык – богатый язык. Возникает вопрос: измеряется ли богатство языка количеством слов или его исторической значимостью? Может быть, есть другой критерий? Конечно, есть. Богатство языка заключается не только в большом количестве слов, но и в разнообразии их значений. Новые смысловые оттенки придают языку гибкость, живость и выразительность. В языке много различных омонимов, синонимов, антонимов, паронимов, которые делают речь образной, красноречивой, экспрессивной, помогают избежать повторения одних и тех же слов, позволяют эмоционально выразить мысль. Каждое слово имеет альтернативное значение в зависимости от функции стиля речи (разговорный / диалектный, художественный, официальный, научный, публицистический). Функциональные стили речи – разновидности литературного языка, которые обусловлены различными сферами общения. Они создаются в зависимости от целей и задач общения и соответственно различаются отбором языковых средств. Знание синонимических богатств родного языка – необходимое условие речевой культуры человека.

Целью статьи является представление принципа работы и процесса работы синонимайзера в Узбекском образовательном корпусе. Материалом статьи послужили материалы Узбекского образовательного корпуса.

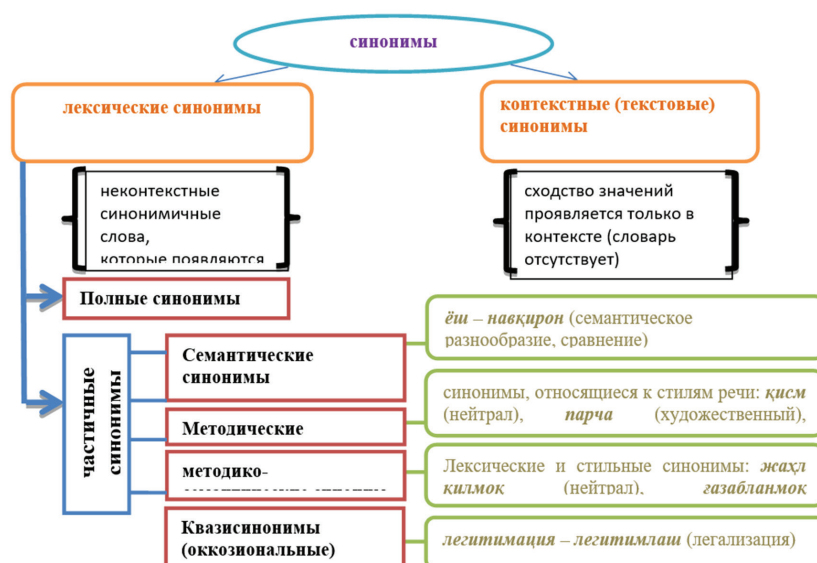
Использование синонимайзера в образовательном корпусе узбекского языка

Синонимия в лингвистике является универсальным языковым явлением. Синонимический ряд является связующим звеном, коммуникативным центром между разными речевыми сферами, разными стилями речи (Корсакова 2010: 105-106).

Известно, что существует два основных типа синонимов: лексические синонимы и контекстные синонимы.

Лексические синонимы, в свою очередь, имеют полные (абсолютные) и частичные (семантические, методологические и методолого-семантические) синонимы. В то время как лексические синонимы формируют синонимию в результате близости и согласованности в семантике слов, а контекстуальные синонимы формируются благодаря лексической согласованности, которая возникает только в контексте текста. Хотя типы синонимов классифицируются в источниках по-разному, в целом можно выделить следующие типы:

Схема 1. Типы синонимов.



Контекстные или текстовые синонимы также называются авторскими синонимами²⁸, и они обычно создают синонимию только в контексте определенного контекста и не могут быть синонимами вне контекста. Их главная задача заключается в описании события, а не в его констатации. Поэтому составление словаря текстовых синонимов – сложный процесс²⁹. К типу лексических синонимов относятся также квазисинонимы в узбекском языке, которые еще не изучены глубоко. Квазисинонимы (ложные синонимы, частичные синонимы³⁰ – это слова, которые относительно похожи по значению, но во всех случаях не являются взаимозаменяемыми. Известно, что лексические синонимы – это в основном лексические единицы со схожими значениями, принадлежащие одной и той же группе слов. Квазисинонимы также образуются синонимами слов, принадлежащие разным группам слов. Например: лемматизация (анализ слова до основной морфемы; определение основы) – (N – существительное) ↔ лемматизациялаш – (Vh – Глаголь название действия); регенерация (N – существительное) ↔ қайта тикланиш (Vh – Глаголь название действия); дахшат (N – существительное) ↔ zo‘r (Adj – прилагательное) (в стиле устного общения).

Как правило, синонимы используются в речи для следующих целей, а в некоторых местах даже необходимы:

- в выражении идеи более четко и точно (для сравнения: *чет эл* и *хориж*);
- в более ярком выражении мысли, эмоциональной окраске;
- во избежание тавтологии (повторение - многократное употребление слова в предложении).

Синонимы эффективно используются для соединения смежных предложений в тексте. Учитывая эти возможности, в учебном корпусе узбекского языка был создан синонимайзер. Синонимайзер – это программа, которая заменяет фразы или слова в тексте синонимами (находящимися в базе данных программы)³¹. В некоторых источниках его называют синонимизатором, синониматором. Этот термин впервые был придуман копирайтером по имени Кевиндарк на форуме SEO для описания программы, которая может упростить переписывание текстов.

Синонимайзер служит для изменения уникальности текста. Это может быть программа-синонимизатор или веб-сайт в Интернете. В мировой компьютерной лингвистике синонимизаторы в основном используются при переписывании (рерайтинг) и копирайтинге для создания уникального контента для веб-сайтов.

Принцип работы синонимайзеров

Синонимайзеры можно разделить на автоматические и полуавтоматические. Полуавтоматические синонимайзеры предлагают пользователю выбрать подходящий из списка синонимов, соответствующий слову, которое он определил. Полностью автоматические синонимы автоматически заменяют слова в тексте значениями используемых символов.

Существует два типа текстовых заменителей, специфичных для автоматических синонимизаторов: неморфологический (линейный) и морфологический (лексический)³².

При неморфологической (линейной) замене учитывается программный бэкэнд заданного слова в тексте, т. е. процесс замены в базе данных слова, содержащего с обеих сторон неисчислимы символы букв, на его аналог по значению, соответствующий той же форме (Абжалова, 2021, с. 12-17). Точнее говоря, этот процесс опирается только на толковый словарь синонимов (Ходжиев, 1974; Отабеков, Бойматова, 2007, с. 49) и вспомогательные словари (Толковый словарь, 2006; Бободжонов, Исламов,

²⁸ <http://rusolimp.kopeisk.ru/leksik/?file=111>; синонимы неполные – это... Что такое синонимы неполные? (academic.ru)

²⁹ <https://synonymonline.ru/synonyms.html>

³⁰ <https://ru.wiktionary.org/wiki/quasi>

³¹ <https://xn----7sbbaqh1km9ah9aiq.net/news-new/sinonimizator-ru.html>

³² <https://animatika.ru/info/gloss/synonymizer.html#:~:text=>

2007, с. 38; Курбанова, Абжалова и др. 2021, с. 988). Форма слова не учитывается. Например, вместо слова “*agar*” предложены его синонимы “*мабодо*”, “*магар*”, “*башарти*”, “*бордию*” (см. рис. 1).

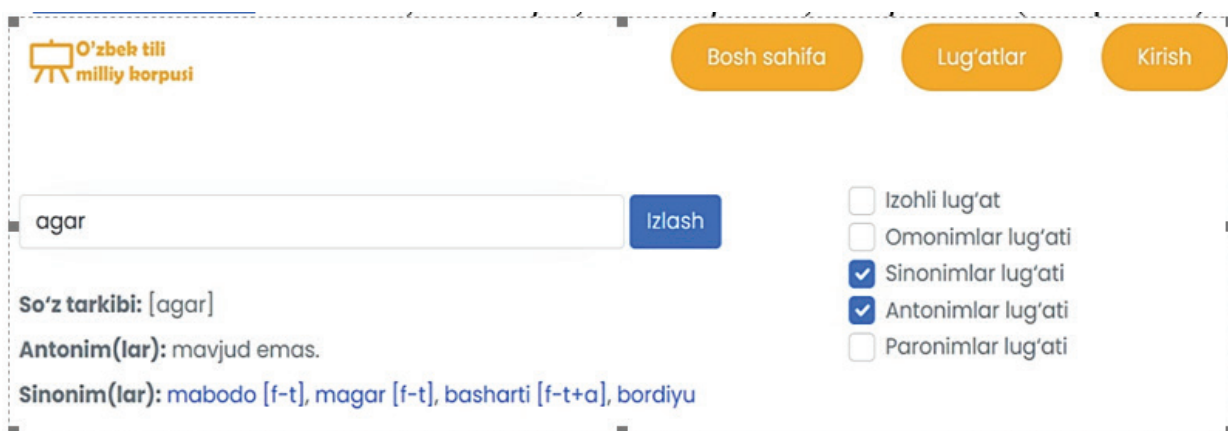


Рис. 1. В синонимайзере представлены синонимы слова «агар» («если»).

Picture 1. The synonymizer presents synonyms for the word "agar" ("if").

Однако, если вы добавите аффикс местного падежа и наберете словоформу “агарда” в строке поиска, информация не будет предоставлена (см. рис. 2).

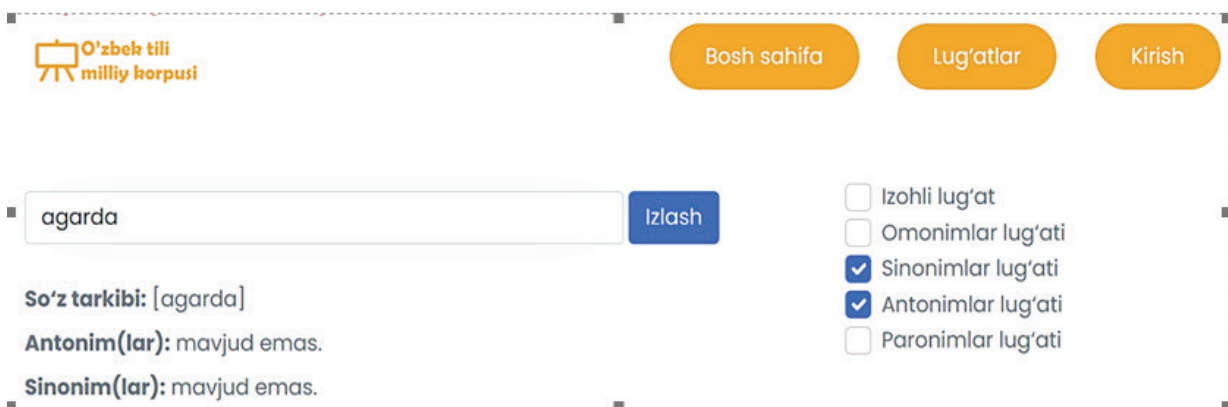


Рис. 2. Синонимизатор не может предоставить синонимы слова «агарда» («а если»).

Picture 2. The synonymizer cannot provide synonyms for the word "agarda" ("what if")

Морфологические (лексические) замены также учитывают грамматическую форму слова, и любая словоформа заменяется соответствующей ей формой. Некоторые программы могут изменять не только слова, но и фразы длиной до десяти слов. Некоторые синониматоры могут выполнять оба типа замены. Если программа не может различить морфологию, в результате возможности леммирования или стем различных словоформ в базе данных вручную или с помощью различных утилит, появляется возможность представления синонимов к словоформам или сопоставления их по смыслу. Неморфологический тип синонимайзера был запущен в Национальном корпусе узбекского языка (см. рис. 3), который является первым синонимайзером, созданным для узбекского языка (Абжалова, 2021, с. 38-41). Ведется работа по совершенствованию базы данных, чтобы можно было заменять словоформы по их морфологическому типу. На рисунке 3 словом «катта», введенное в поисковую строку столбца «Словари» Национального корпуса узбекского языка, были указаны синонимы по корпусу и даже слова с противоречивыми значениями (антонимы).



Рис. 3. Неморфологический тип синонимайзера в Национальном корпусе узбекского языка.

Picture 3. Non-morphological type of synonymizer in the Uzbek National Corpus.

Проблема с синонимайзерами. Основные проблемы заключаются в следующем:

- выбор синонимов должен зависеть от контекста;
- инверсия предложения усложняет работу синтаксического анализатора и, как следствие, теряется возможность замены словосочетаний их синонимами;
- если новые слова (термины) не включены в базу данных, неологизмы не будут представлены пользователю.

Чтобы синонимизированный текстовый стиль был свободным, программа должна учитывать закон Зипфа, то есть активные слова должны быть заменены активными синонимами, а редкие слова должны быть заменены редкими синонимами.

Заключение

Обработка синонимизированного текста требует дополнительной грамматической, синтаксической и семантической обработки в других лингвистических компьютерных программах. В процессе обработки тексты сравниваются между собой, после чего выбирается наиболее подходящий вариант.

Литература

Абжалова, М. Вопрос создания синонимайзера или синонимизатора в Национальном Корпусе узбекского языка // Теоретические и практические вопросы создания узбекского национально-просветительского корпуса / Материалы Международной научно-практической конференции. 7 мая 2021 года. Электронное издание / электронная книга. Ташкент : ТошГУУЯЛ, 2021. С. 38-41.

Абжалова, М. А. Возможности поиска словоформ в национальном корпусе узбекского языка на основе лексикографической базы. // Компьютерная лингвистика : проблемы, решения, перспективы / Материалы республиканской научно-технической конференции. Электронное издание. Ташкент : Ташкентский государственный университет узбекского языка и литературы, 23.04.2021. С.12-17.

Бободжонов, Ш., Исламов, И. Словарный классификатор узбекского языка : для школьников. Ташкент : Янги аср авлоди, 2007. 38 с.

Корсакова, Ю. С. Синонимия – универсальное языковое явление // Филологические науки. Вопросы теории и практики, № 1 (5) 2010, часть 2. Тамбов, 2010. С. 103-106.

Курбанова, М., Абжалова, М. [и др.]. Акцентный словарь заимствованных слов узбекского языка : учебно-методический словарь. Ташкент : Нодирабегим, 2021. 988 с.

Отабеков, Ш., Бойматова, Б. Словарь синонимов узбекского языка : для школьников. Ташкент : Янги аср авлоди, 2007. С. 49.

Толковый словарь узбекского языка : более 80 000 слов и словосочетаний (под ред. А. Мадвалиева).

5 томов. Ташкент : Национальная энциклопедия Узбекистана, 2006.

Ходжиев, А. Толковый словарь синонимов узбекского языка. Ташкент : Учитель, 1974. 308 с.

References

Abjalova, M. The issue of creating a synonymizer or synonymizer in the National Corpus of the Uzbek language // Theoretical and practical issues of creating an Uzbek national educational corpus / Materials of the International Scientific and Practical Conference. May 7, 2021. Electronic edition / e-book. Tashkent : TSUULL, 2021 . pp. 38-41. (in Russ.).

Abjalova, M. Possibilities of searching for word forms in the national corpus of the Uzbek language based on the lexicographic base. // Computational linguistics: problems, solutions, prospects / Materials of the republican scientific and technical conference. Electronic edition. Tashkent : Tashkent State University of Uzbek Language and Literature, 04/23/2021. pp.12-17. (in Russ.).

Bobojonov, Sh., Islamov I. Dictionary classifier of the Uzbek language: For schoolchildren. – Tashkent: Yangi asr avlodi, 2007. – 38 p.

Korsakova, Yu. S. Sinonimiya - universal'noe yazy'kovo yavlenie // Filologicheskie nauki. Voprosy` teorii i praktiki, № 1 (5) 2010, chast` 2. Tambov, 2010. pp.103-106. (in Russ.).

Kurbanova, M., Abjalova M. and others. Accent dictionary of borrowed words of the Uzbek language: [Text]. educational-methodical dictionary / M. Kurbanova, M. Abdzhalova, N. Akhmedov, R. Tulaboeva. Tashkent : Nodirabegim, 2021. 988 p. (in Russ.).

Otabekov Sh., Boymatova B. Dictionary of synonyms of the Uzbek language: For schoolchildren. Tashkent : Yangi asr avlodi, 2007. 49 p. (in Russ.).

Explanatory dictionary of the Uzbek language: more than 80,000 words and phrases (edited by A. Madvaliev). 5 volumes. Tashkent : National Encyclopedia of Uzbekistan, 2006. (in Russ.).

Khodzhiyev A. Explanatory dictionary of synonyms of the Uzbek language. Tashkent : Teacher, 1974. 308 p. (in Russ.).

Для цитирования:

Абжалова М. А. Синонимайзер (синонимизатор) в образовательном корпусе узбекского языка [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Abjalova M. A. Synonymizer in the educational corpus of the Uzbek language. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 06.09.2021.

ЦИФРОВИЗАЦИЯ КУЛЬТУРНОГО НАСЛЕДИЯ ТУВЫ: СОСТОЯНИЕ, ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

DIGITALIZATION OF TIVA'S CULTURAL HERITAGE: STATUS, PROBLEMS AND PROSPECTS

Тарыма Алдынсай Константиновна, кандидат педагогических наук, доцент кафедры информатики физико-математического факультета, Тувинский государственный университет.

Taryma Aldynsay Konstantinovna, Candidate of Pedagogical Sciences, Associate Professor of the Department of Computer Science of the Faculty of Physics and Mathematics, Tuvan State University.

taryma_ak@mail.ru



Ооржак Байлак Чаш-ооловна, доктор филологических наук, руководитель научно-образовательного центра «Тюркология», Тувинский государственный университет.

Oorzhak Baylak Chash-oolovna, Doctor of Philology, Head of the Scientific and Educational Center of Turkology, Tuvan State University.

oorzhak.baylak@mail.ru

ORCID ID: 0000-0001-8373-9107



Салчак Аэлиита Яковлевна, кандидат филологических наук, доцент кафедры тувинской филологии и общего языкознания, Тувинский государственный университет.

Salchak Aelita Yakovlevna, Candidate of Philological Sciences, Associate Professor of the Department of Tuvan Philology and General Linguistics, Tuvan State University.

aelita_74@mail.ru

ORCID ID: 0000-0003-2304-6901





Хертек Аржаана Борисовна, кандидат филологических наук, старший научный сотрудник научно-образовательного центра «Туркология», Тувинский государственный университет.

Khertek Arzhaana Borisovna, Candidate of Philological Sciences, Senior Researcher of the Scientific and Educational Center of Turkology, Tuvan State University.

khertek.ab@yandex.ru

ORCID ID: 0000-0002-0896-8468



Бавуу-Сюрюн Мира Викторовна, доктор филологических наук, профессор кафедры тувинской филологии и общего языкознания, Тувинский государственный университет. mira.bavuu_surun@mail.ru

Bavuu-Syuryun Mira Viktorovna, Doctor of Philology, Professor of the Department of Tuvan Philology and General Linguistics, Tuvan State University.

mira.bavuu_surun@mail.ru

ORCID ID: 0000-0002-7456-7339



Ондар Валентина Сувановна, кандидат филологических наук, доцент кафедры русского языка и литературы, Тувинский государственный университет.

Ondar Valentina Suvanovna, Candidate of Philological Sciences, Associate Professor of the Department of Russian Language and Literature, Tuvan State University.

barys-hoov@mail.ru

ORCID ID: 0000-0002-6882-7720

Монгуш Чодураа Михайловна, кандидат физико-математических наук, доцент кафедры информатики физико-математического факультета, Тувинский государственный университет.

Mongush Choduraa Mikhailovna, Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Computer Science of the Faculty of Physics and Mathematics, Tuvan State University.

multik_10bv@mail.ru

ORCID ID: 0000-0002-0006-4835



В статье обосновывается необходимость цифровизации культурного наследия Тувы – одного из важнейших факторов сохранения и развития тувинского языка и культуры; обозначаются проблемы и перспективы работы в данном направлении. Даются сведения о содержании «Электронного корпуса текстов тувинского языка», который послужил основой заявки нового Международного проекта по улучшению образовательных программ в области искусства и гуманитарных наук с помощью методов и технологий STEM. Тувинским госуниверситетом разработан обновлен учебный план магистратуры направления 44.04.01 Педагогическое образование, профиля «Цифровые технологии в гуманитарном образовании», реализация которого позволит подготовить специалистов, которые будут владеть методами цифровизации и компьютерной обработки данных в области культурного наследия тувинцев.

The article substantiates the need to digitalize the cultural heritage of Tuva – one of the most important factors in the preservation and development of the Tuvan language and culture; the problems and prospects of work in this direction are outlined. Information is given about the content of the "Electronic Corpus of Texts of the Tuvan language", which served as the basis for the application of a new International project to improve educational programs in the field of arts and humanities using STEM methods and technologies. Tuva State University has developed and updated the curriculum of the master's degree in the direction of 44.04.01 Pedagogical education, the profile "Digital technologies in humanitarian education", the implementation of which will allow training specialists who will master the methods of digitalization and computer processing of data in the field of cultural heritage of Tuvinians.

Ключевые слова: *цифровизация, культурное наследие Тувы, Электронный корпус текстов тувинского языка, STEM, учебная программа магистратуры.*

Keywords: *digitalization, cultural heritage of Tuva, Electronic corpus of Tuvan language texts, STEM, Master's degree curriculum.*

Введение

Культурное наследие народа – это материальная и духовная ценность, созданная предыдущими поколениями, пронесенная через века и выдержавшая испытания временем, сохраняющаяся и почитаемая носителями и передающаяся следующим поколениям. Культурное наследие является объединяющей силой народа, неотъемлемой частью его самосознания на протяжении его истории и развитии.

Туву по праву можно назвать одной из самых богатых республик в плане культурного наследия. Нематериальное культурное наследие тувинцев включает тувинский язык, тувинскую музыку, фольклор, танцы, праздники, обряды и пр. Поклонники музыкального искусства хорошо знакомы с уникальным вокальным искусством тувинцев – горловым пением (хөөмей).

На небольшой территории Тувы сосредоточено огромное количество самых разнообразных природных, исторических, археологических и культурных памятников. Всего на государственной охране Республики Тыва находятся 860 объектов культурного наследия, большинство из которых является объектами археологического наследия — 787 ед. Среди них – объект Всемирного наследия ЮНЕСКО биосферный заповедник «Убсунурская котловина»; памятник федерального значения – древнеуйгурская крепость «Пор-Бажын»; историко-культурный комплекс «Долина царей»; руины разрушенного храма «Устуу-Хурээ»; ниша с изображением Будды, высеченная в скале на одном из берегов Саяно-Шушенского водохранилища, являющаяся единственной в России. На территории Республики Тыва расположены многочисленные памятники наскального искусства, древнетюркские изваяния, памятники древнетюркской рунической письменности и т. д. Приведенными примерами далеко не исчерпывается перечень уникальных памятников истории и культуры на этой территории.

Понятие «культурное наследие тувинцев» охватывает также и памятники письма, хранящиеся в библиотеках и архивных фондах – материалы на старомонгольском, тувинском латинизированном и кириллическом алфавитах.

Состояние, проблемы и перспективы цифровизации культурного наследия Тувы

Культурное достояние тувинского народа как один из элементов сохранения самобытности народа и явление общемировой культуры должно быть зафиксировано для последующих поколений средствами современных информационных технологий.

Научными сотрудниками научно-образовательного центра «Тюркология», преподавателями и студентами филологического факультета оцифрованы тексты художественной литературы на тувинском языке. Переведенные в цифровой формат тексты после проверки и корректировки размещаются на сайт «Электронный корпус текстов тувинского языка» (ЭКТТЯ). Подготовлены некоторые базы данных по тувинскому языку. Начинаются работы по созданию подкорпуса тувинского языка на латинице.

Базы ЭКТТЯ послужили основой для создания научно-методического и образовательного контента «Тываның чогаалчылары – Писатели Тувы» <http://pisатели-tuvy.ru/>, который содержит тексты произведений тувинских писателей 1970-90-х годов в электронном виде, автобиографические сведения о писателях, фотографии, критические статьи по их произведениям. Материал сайта широко используется учителями тувинского языка и литературы республики.

Работа по пополнению баз ЭКТТЯ продолжилась как один из важных пунктов Государственной программы «Развитие тувинского языка в Тувинском государственном университете на 2021-2024 годы».

В настоящее время ЭКТТЯ включает:

- 1) подкорпус текстов художественной литературы на тувинском языке советского и современного периодов;
- 2) подкорпус параллельных тувинско-русских, русско-тувинских, монгольско-тувинских текстов с выравниванием текстов на уровне предложений;
- 3) подкорпус произведений фольклора разных жанров;
- 4) диалектный подкорпус с интерактивными лингвистическими картами по диалектным явлениям в тувинском языке.

Разработаны базы данных:

- 1) Лексико-грамматические разряды и семантические группы имен существительных тувинского языка
- 2) Лексико-грамматические разряды и семантические группы имен прилагательных тувинского языка
- 3) Лексико-грамматические разряды и семантические группы местоимений тувинского языка
- 4) Лексико-семантические группы глаголов тувинского языка: сочетаемость с именами

существительными и наречиями;

- 5) Клише и стандарты в тувинском героическом эпосе
- 6) Антропонимы в тувинском героическом эпосе
- 7) Зачины тувинских сказок
- 8) Концовки тувинских сказок.
- 9) Гидронимия Тувы;
- 10) Лексика животноводства в цэнгэльском говоре тувинского языка;
- 11) Монгольско-тувинские параллельные тексты;
- 12) Русско-тувинские параллельные тексты.

Для ЭКТТЯ преподавателями кафедры информатики физико-математического факультета разработаны пробные программы по поиску морфем, словоформ, аналитических скреп в тексте на тувинском языке. В настоящее время занятием собственной программы поиска в ЭКТТЯ используется популярная бесплатная, мультиплатформенная программа Antconc, разработанная профессором Лоуренсом Антони. Благодаря AntConc сегодня материалы ЭКТТЯ используются в исследованиях, практике преподавания вузовских курсов по тувинской филологии; разрабатываются учебно-методические материалы по курсу «Родной (тувинский) язык». При всем удобстве использования этой программы нужно признать, что для масштабных исследований и прикладных работ по тувинскому языку она, конечно же, не может полностью удовлетворять запросы тувинских филологов. Поэтому перед нами стоит вопрос разработки программного обеспечения ЭКТТЯ.

При назревшей необходимости работы по программному обеспечению текстов ЭКТТЯ и цифровизации объектов культурного наследия Тувы очевидна необходимость подготовки специалистов, профессионально владеющих методами создания программных продуктов, обладающих знаниями по тувинскому языку, материальной и нематериальной культуре, способных фиксировать и обрабатывать данную информацию.

Тувинский государственный университет принял участие в Международном конкурсе Европейского Союза Erasmus+ и стал в 2020 г. обладателем гранта «Улучшение образовательных программ в области искусства и гуманитарных наук с помощью методов и технологий STEM / ARTEST». В данном проекте принимают участие ряд вузов России, Германии, Греции, Кипра и Монголии.

Проект направлен на подготовку магистерской программы по подготовке профессиональных кадров в области гуманитарных наук, владеющих цифровыми методами в области искусства и гуманитарных наук. Гуманитарный аспект этой магистерской программы усилен методами и инструментами STEM и по усилению STEM-образования и планом действий по цифровому образованию. STEM-образование (Science, Technology, Engineering and Mathematics) должно удовлетворять требованиям по приобретению знаний на основе простой и доступной визуализации научных явлений, что обусловит облегчать изучаемый материал. Поэтому новая магистерская программа направлена на подготовку специалистов-гуманитариев, владеющих STEM-инструментами – современными инженерными технологиями, интегрирующихся с гуманитарными знаниями.

Сегодня STEM специалисты – самые востребованные люди на мировом рынке труда и в России потребность в STEM образовании растет. Основное направление развития STEM – цифровизация и анализ данных, поэтому компетенции в IT и Data Science необходимы практически каждому специалисту (см. подробнее на РБК: <https://trends.rbc.ru/trends/education/5f6399a69a79471ec02bfe4f>), т.к. применение методов искусственного интеллекта и машинного обучения оптимизируют работу во всех областях, в том числе и в гуманитарных исследованиях и прикладных науках в области филологии.

Согласно разработанному Тувинским госуниверситетом учебному плану направления 44.04.01 Педагогическое образование, профиль «Цифровые технологии в гуманитарном образовании» выпускники магистратуры приобретают следующие профессиональные компетенции: 1) применять современные инновационные цифровые инструменты и приемы STEAM для создания различных учебных проектов и материалов этнокультурной составляющей гуманитарного образования в профессиональной деятельности; 2) разрабатывать учебно-методические материалы, электронные образовательные ресурсы,

необходимые для организации учебной деятельности в области гуманитарного образования с применением цифровых технологий; 3) самостоятельно осуществлять научно-исследовательскую деятельность с использованием современных методов исследования и информационно-коммуникационных технологий в своей профессиональной области и в междисциплинарных исследованиях; 4) владения основными методами программирования и трёхмерного моделирования для обработки информации в области культурного наследия в целях использования в научно-исследовательской и образовательной деятельности.

При подготовке магистрантов предусмотрены следующие специальные дисциплины: «Корпусная лингвистика», «Web-программирование и дизайн», «Цифровые методы в сохранении и презентации культурного наследия», «Программирование на языке Python», «Введение в анализ данных», «Основы графического дизайна и 3D-моделирования», «Технология создания цифровых культурных объектов», «Информационные системы и базы данных в гуманитарных областях», «Основы SQL и работы с базами данных» и др.

Цифровая гуманитаристика (Digital Humanities) направлена на междисциплинарное решение задач в современной гуманитарной науке в контексте цифрового поворота. Необходимо изучение актуальной литературы, междисциплинарных подходов, современного инструментария нового направления «цифровых гуманитарных наук» применительно к задачам исторического исследования в условиях «цифрового поворота».

Наряду с дисциплинами по компьютерной инженерии вводятся дисциплины, призванные углубить компетенции в гуманитарной области знания: «Актуальные проблемы филологических исследований», «Ареальные исследования в гуманитарных областях», «История и археология Центральной Азии», «Фольклор народов Центральной Азии», «Культура народов Центральной Азии».

Непосредственному применению выпускниками магистратуры современных цифровых технологий в обучении филологических дисциплин в образовательных учреждениях будут способствовать знания, полученные на курсах «Современные методы обучения родному языку в школе», «Цифровые технологии в обучении филологических дисциплин в школе».

Большое внимание будет уделяться научно-исследовательской работе магистрантов, которая будет состоять в разработке программных продуктов по цифровизации культурного наследия тувинцев. Особо будет обращать внимание на изучение английского языка, оно будет охватывать весь период обучения в магистратуре. Качественное овладение английским языком необходимо для дальнейшего профессионального развития выпускника.

Значительное место в программе уделяется обязательным общенаучным дисциплинам: «Философские проблемы науки», «Современные проблемы науки и инновационные процессы в образовании», «Методология и методы научного исследования». Вводятся также курсы: «Государственная политика в сфере образования и охрана культурного наследия», «Музееведение» и др.

По окончании магистратуры 44.04.01 Педагогическое образование по профилю «Цифровые технологии в гуманитарном образовании» выпускники получают актуальное образование и возможность трудоустройства сразу после выпуска в образовательные и научные организации. Выпускники должны уметь использовать информационно-справочную систему по тувинскому языку; владеть методиками проведения каталогизации текстов и записей на тувинском языке; уметь работать с корпусными словарями различных типов в том числе и иллюстрированными; иметь навыки использования компьютерных программ для оцифровки текстов; иметь представление о сборе и составлении материалов для мультимедийного, диалектного, устного, словарного корпусов; иметь представление о составлении информации экстралингвистического характера; создавать и работать с программами автоматической обработки текста.

Выпускник должен знать проектирование, разработку и использование онтологических информационных систем для цифровизации культурного наследия на основе открытых данных для творческой, технической и научной деятельности; проектирование систем учета и долгосрочного хранения цифровых и аналоговых данных; электронную визуализацию и изобразительное искусство;

работу с проектными архивами (например, с музеями); визуализацию прошлого, посвященную музейным и выставочным технологиям; цифровую археологию крупных городских и подземных инфраструктур; археологическую разведку; 3D раскопки Geodata и ГИС; цифровое архивирование; цифровизацию археологического наследия; визуализацию гипотез в цифровых 3D-моделях; создавать и использовать проекты виртуальных музеев.

В рамках этого проекта на филологическом факультете начинает работать цифровая лаборатория. Начали действовать несколько студенческих проектов по оцифровке и представлению объектов культурного наследия. В проекте «Цифровизация литературного наследия Тувы» (рук. А. Хертек, В. Ондар) студентами 1-го курса ведется работа по оцифровке текстов детской художественной литературы, переведенных на тувинский язык: басни И. Крылова. В проекте «Традиционные мужские и женские занятия» (рук. А. Салчак, Б. Ооржак) магистранты и студенты работают над подготовкой видеоматериалов, которые иллюстрируют традиционные ремесла, занятия тувинцев.

В 2021 году исполнители проекта провели ряд рабочих встреч участников консорциума. Ход и результаты работы по цифровизации культурного наследия Тувы были доложены на VI Форуме «Тыва дыл болгаш амгы үе», посвященном Дню тувинского языка 09 ноября 2021 г. Ооржак Б.Ч. выступила с докладом «Тыва дыл компьютерлиг технологиялар үезинде».

Важным событием для развития сотрудничества в области компьютерной лингвистики и дальнейшей работы по ЭКТТЯ является IX Международная конференция по компьютерной обработке тюркских языков «TurkLang 2021» (21–23 сентября 2021 г.) <http://tuvsu.ru/news/6908/>, <http://tuvsu.ru/news/6918/>, <http://tuvsu.ru/news/6941/>, <http://tuvsu.ru/news/6928/>. Конференция прошла в этот раз на площадке Тувинского государственного университета. В работе конференции приняли участие более 100 ученых в очной и дистанционной формах на платформе ZOOM. География участников конференции охватила ближнее и дальнее зарубежье, регионы и города России: Тува, Татарстан, Башкортостан, Чувашия, Саха-Якутия, Горный Алтай, Хакасия, Москва, Новосибирск, Томск, Омск. Сотрудники университета выступили с докладами на темы: «Создание базы данных текстов, написанных на латинской графике, для Электронного корпуса тувинского языка» (Хертек А. Б.), «Создание базы данных для подкорпуса периодической литературы» (Хертек А.Б., Саая Ч. О.), «Цифровизация культурного наследия в программе подготовки магистров в рамках проекта Эразмус+» (Ооржак Б.Ч., Тарыма А. К., Бавуу-Сюрюн М. В., Салчак А.Я., Хертек А.Б., Ондар В.С.).

Необходимо продолжить совместную работу с коллегами по развитию электронных корпусов тюркских языков и унификации систем разметок, а также созданию единых многоязычных лингвистических ресурсов и сервисов для тюркских языков. И работа по пополнению и развитию ЭКТТЯ будет продолжена как важная часть работы по цифровизации культурного наследия.

Для цитирования:

Тарыма А. К., Ооржак Б. Ч., Салчак А. Я., Хертек А. Б., Бавуу-Сюрюн М. В., Ондар В. С., Монгуш Ч. М. Цифровизация культурного наследия Тувы: состояние, проблемы и перспективы [Электронный ресурс] // Родные языки и культуры в современном изменяющемся мире. 2022. №1. URL.: (дата обращения: дд.мм.гг.). DOI:

For citation:

Taryma A. K., Oorzhak B. Ch., Salchak A. Ja., Hertek A. B., Bavuu-Sjurjun M. V., Ondar V. S., Mongush Ch. M. Digitalization of Tuva's cultural heritage: status, problems and prospects. Native languages and cultures in the modern changing world. 2022. №1. [online]. Available at (access date ...). DOI:

Дата поступления/Submission date: 08.12.2021.

ТРЕБОВАНИЯ К ИНФОРМАЦИИ И СТАТЬЯМ, ПРЕДСТАВЛЯЕМЫМ ДЛЯ ПУБЛИКАЦИИ (СОКРАЩЕННАЯ ВЕРСИЯ)

Редакция принимает только оригинальные, неопубликованные ранее научные статьи и иные материалы научного характера. Тематика статьи должна соответствовать содержанию журнала, а также одной из основных рубрик. В редакцию журнала предоставляются:

1. Авторский оригинал статьи (на русском языке)

Оригинал статьи предоставляется в электронной форме, содержащей текст в формате Microsoft Word, через систему электронной редакции на сайте издания.

Весь текст набирается шрифтом Times New Roman Cyr, кеглем 12 pt, с полуторным междустрочным интервалом. Объем статьи — не более 18—25 тысяч знаков с пробелами (без учета реферата, ключевых слов, примечаний, списка источников).

Структура текста:

Сведения об авторе/авторах — имя, отчество, фамилия, место работы (учебы), в том числе подразделение (при наличии), должность, почтовый адрес организации, ученая степень, ученое звание, ORCID, адрес электронной почты автора — размещаются перед названием статьи в указанной выше последовательности и будут опубликованы вместе со статьей. Контактная информация: почтовый адрес для передачи корреспонденции, телефоны (рабочий, домашний, мобильный) — может быть использована только для переговоров между автором и редакцией и не подлежит опубликованию.

Индекс УДК, раскрывающие тематическое содержание статьи.

Цветная фотография автора (отдельным файлом).

Название статьи.

Сведения об источнике финансирования исследования/публикации (в случае наличия) оформляется в виде сноски «звездочка» к названию статьи и представляется в подстрочнике на 1 странице.

Реферат — краткое изложение статьи по следующей структуре: актуальность проблематики и новизна решения, главные содержательные аспекты. Объем — 200—250 слов. Размещается после названия статьи.

Ключевые слова по содержанию статьи (8—10 слов) размещаются после реферата.

Основной текст статьи желательно разбить на подразделы (с подзаголовками).

Инициалы в тексте набираются через неразрывный пробел с фамилией (одновременное нажатие клавиш «Ctrl» + «Shift» + «пробел». Между инициалами пробелов нет.

В тексте используются кавычки «...», если встречаются внутренние и внешние кавычки, то внешними выступают «елочки», внутренними «лапки» — «...“...”».

В тексте используется длинное тире (—), получаемое путем одновременного нажатия клавиш «Ctrl» + «Alt» + «-» (на цифровом блоке клавиатуры), а также дефис (-).

В списках и перечнях не используется автоматическая нумерация текстового редактора.

Инфографика, в том числе таблицы, схемы, рисунки и формулы в тексте должны нумероваться; схемы и таблицы должны иметь заголовки, размещенные над схемой или полем таблицы, а каждый рисунок — подрисуночную подпись.

Список источников оформляется как затекстовые библиографические ссылки в соответствии с принятым стандартом (ГОСТ Р 7.0.5—2008 «Библиографическая ссылка»), выносятся в конец статьи. Источники даются в алфавитном порядке. Отсылки к списку в основном тексте даются в круглых скобках (Фамилия, год, страница) (например: (Иванов, 1980, с. 10).

Заголовки (и названия журналов) во второй и последующих ссылках нельзя заменять их словесными эквивалентами «Его же», «Он же», «Там же» и т. д.

При оформлении списка источников автоматическая нумерация текстового редактора не используется, порядковый номер отделяется от текста ссылки знаком табуляции. Знаки «точка» и

«тире», разделяющие области библиографического описания, заменяются точкой. Во всех библиографических ссылках на электронные ресурсы обязательно указывается дата обращения.

При цитировании интернет-ресурсов необходимо максимально полно, насколько это возможно указывать ФИО автора (авторов), точное название работы, название ресурса, опубликовавшего описываемый документ, а также точную ссылку на веб-страницу-первоисточник (а не на главную страницу сайта, с которого взят материал).

Любое упоминание в тексте статьи имен ученых и исследователей должно сопровождаться отсылками на их работы. Все источники, представленные в списке источников, должны иметь отсылки в тексте статьи. В одной отсылке не рекомендуется размещать более двух источников.

Редакция рекомендует отслеживать и использовать публикации по теме статьи, вышедшие в предыдущих номерах журнала «...», а также в других научных журналах по тематике авторской рукописи, в целях обеспечения преемственности и подтверждения актуальности темы.

Примечания нумеруются арабскими цифрами (с использованием кнопки меню текстового редактора «надстрочный знак» — a1, a2). При оформлении библиографических источников, примечаний и ссылок автоматические «сноски» текстового редактора не используются. Текст примечаний размещается после текста статьи с заголовком «Примечания».

Подписуемые подписи оформляются по схеме: название/номер иллюстрации — пояснения к ней (что/кто изображен, где; для изображений обложек книг и их содержимого — библиографическое описание; и т. п.). Имена файлов в списке должны соответствовать названиям/номерам предоставляемых фотоматериалов.

2. Материалы на английском языке

Информация об авторе/авторах — имя, инициал отчества (если имеется), фамилия, место работы (учебы), почтовый адрес организации, адрес электронной почты автора, название статьи,

реферат, ключевые слова (в том же объеме и порядке, как в русском тексте), сведения об источнике финансирования — в электронной форме (отдельный файл) через систему электронной редакции как дополнительные материалы, содержащей текст в формате Microsoft Word.

Отдельным файлом предоставляется список источников в транслитерации, с переводом на английский язык (согласно Правилам, принятым в редакции). Нумерация источников должна соответствовать нумерации в авторском оригинале на русском языке.

3. Иллюстративные материалы

Предоставляются в электронной форме (фотография автора обязательна, иллюстрации) — отдельными файлами через систему электронной редакции как дополнительные материалы в форматах TIFF/JPG разрешением не менее 300 dpi одновременно с авторским оригиналом статьи. Не допускается предоставление иллюстраций, импортированных в Microsoft Word, а также их ксерокопий.

Таблицы, схемы, диаграммы и графики предоставляются в файле в формате Microsoft Word или Microsoft Excel, включающем исходные данные, в котором возможно редактирование графических материалов.

Иллюстративный материал и инфографика должны быть адаптированы для черно-белой печати высокого качества.

Ко всем изображениям автором предоставляются подписуемые подписи (включаются в файл с авторским текстом).

4. Распечатанная и подписанная Авторская оферта

Правовые вопросы, связанные с публикацией в журнале, включая обязательства сторон (автора и издателя), регулируются на основе подписанной Авторской оферты к Приглашению делать оферты. Предоставляя свои материалы в журнал, авторы гарантируют, что они обладают исключительными правами на передаваемый для публикации материал, который является их оригинальным,

нигде ранее не публиковавшимся произведением.

Авторская оферта должна быть предоставлена в распечатанном виде на бумажном носителе, подписана автором (соавторами) собственноручно шариковой ручкой с синими чернилами.

Для удобства можно воспользоваться подготовленными образцами Авторской оферты (или Авторской оферты для статей в соавторстве).

Оригинал акцепта нужно выслать по адресу: 667000, Российская Федерация, Республика Тыва, г. Кызыл, ул. Ленина, д.36. Редакция журнала.

Авторы статей несут ответственность за содержание статей и за сам факт их публикации. Редакция не всегда разделяет мнения авторов и не несет ответственности за недостоверность публикуемых данных. Редакция журнала не несет никакой ответственности перед авторами и/или третьими лицами и организациями за возможный ущерб, вызванный публикацией статьи.

Статьи и другие предоставленные материалы не возвращаются.

Статьи, оформленные без учета вышеизложенных требований, к публикации не принимаются.

Требования составлены с учетом требований, изложенных Приказе Минобрнауки России от 25 июля 2014 г. № 793 «Об утверждении правил формирования в уведомительном порядке перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук и требований к рецензируемым научным изданиям для включения в перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук» с изменениями, утвержденными Приказом Минобрнауки России от 03 июня 2015 № 560 (подробнее на сайте: <http://vak.ed.gov.ru/>).

РЕКОМЕНДАЦИИ ПО ТРАНСЛИТЕРАЦИИ СПИСКА ЛИТЕРАТУРЫ (REFERENCES)

Перевод русскоязычной библиографической ссылки на английский язык состоит из следующих частей, **отделяемых точками**:

- транслитерация списка авторов;
- транслитерация названия книги, статьи или автореферата и перевод названия на английский язык в квадратных скобках;
- для статьи: транслитерация названия журнала (научной конференции) и в квадратных скобках перевод названия журнала (научной конференции). Название издания выделяется курсивом, отделяется точкой;
- выходные данные (место издания, год, номер выпуска, количество страниц) приводятся на английском языке. Необходимо указать официальную английскую версию названия издательства (обычно указывается на сайте журнала), если ее нет, то приводится транслитерация.

Если в русском или иностранном источнике

автор – иностранец (англичанин, немец, француз и т. п.) то нельзя транслитерировать его имя и фамилию, нужно найти их в оригинальном написании.

Для описания книг (монографий, учебников и т. п.) указывается, количество страниц (235 p.), для статей – диапазон страниц (pp. 220–222).

На сайте <http://translit.ru/> можно бесплатно воспользоваться программой транслитерации русского текста в латиницу. Последовательность действий:

- На сайте выбираем стандарт BGN. До помещения текста в рабочее поле и до начала транслитерации убедитесь, что выбрали именно стандарт BGN.
- Помещаем библиографические ссылки из русскоязычного списка в рабочее поле и нажимаем кнопку «в транслит». Копируем получившийся текст и работаем дальше с ним – вносим необходимые дополнения и изменения.

Принятые сокращения

| | |
|--|--|
| Номер | No. (no.) |
| Страницы | P. (p.) (если указана одна страница или количество страниц в книге) – page |
| Рр. (pp.) (если указан диапазон страниц) | |
| Том | Vol. (vol.) – volume |
| В трех томах | In 3 vols. |
| Часть | Pt (pt) – part |
| Выпуск | Iss. (iss.) – issue |

Абумова Ольга Дмитриевна, канд. филол. наук, доцент кафедры русской филологии и журналистики Хакасского государственного университета им. Н. Ф. Катанова.

Анна́й Элла́да Кан-ооловна, канд. филол. наук, научный сотрудник Тувинского института гуманитарных и прикладных социально-экономических исследований при правительстве Республики Тыва.

Байыр-оол Азияна Витальевна, канд. филол. наук, старший научный сотрудник Института филологии СО РАН.

Дамдынчап Вера Монгушевна, канд. истор. наук, доцент, зав. кафедрой всеобщей истории, археологии и документооборота Тувинского государственного университета.

Добринина Альбина Альбертовна, канд. филол. наук, старший научный сотрудник Института филологии СО РАН.

Ондар Валентина Сувановна, канд. филол. наук, доцент кафедры русского языка и литературы Тувинского государственного университета.

Прокопьева Алена Кирилловна, канд. филол. наук, доцент кафедры якутского языка Института языков и культур народов Северо-Востока Российской Федерации Северо-Восточного федерального университета им. М. К. Аммосова.

Салчак Аэлита Яковлевна, к. филол. н., начальник отдела организации научных исследований Тувинского государственного университета.

Санчай Чойган Херел-ооловна, канд. культурологии, преподаватель кафедры философии Тувинского государственного университета.

Соян Айланмаа Мылдыргыновна, канд. филол. наук, доцент кафедры тувинской филологии и общего языкознания Тувинского государственного университета.

Тазранова Алена Робертовна, канд. филол. наук, старший научный сотрудник Института филологии СО РАН.

Тюнтешева Елена Валерьевна, канд. филол. наук, старший научный сотрудник Института филологии СО РАН.

Уртегешев Николай Сергеевич, д-р филол. н., ведущий научный сотрудник Института филологии СО РАН.

Хертек Аржаана Борисовна, канд. филол. наук, старший научный сотрудник, научно-образовательного центра «Тюркология» Тувинского государственного университета.

Чемчиева Аржана Петровна, канд. филос. наук, научный сотрудник Института археологии и этнографии СО РАН.

Шагдурова Ольга Юрьевна, канд. филол. наук, старший научный сотрудник Института филологии СО РАН.

Abumova Ol'ga Dmitrievna, Candidate of Philology, Associate Professor of the Department of Russian Philology and Journalism of the Khakass State University named N. F. Katanov.

Annaj Jellada Kan-oolovna, Candidate of Philology, Researcher at the Tuvan Institute for Humanitarian and Applied Socio-Economic Research under the Government of the Tuva Republic.

Bajyr-ool Azijana Vital'evna, Candidate of Philology, Senior Researcher at the Institute of Philology SB RAS.

Damdynchap Vera Mongushevna, Candidate of History, Associate Professor, Head of the Department of General History, Archeology and Documentation of the Tuvan State University.

Dobrinina Al'bina Al'bertovna, Candidate of Philology, Senior Researcher of the Institute of Philology SB RAS.

Ondar Valentina Suvanovna, Candidate of Philology, Associate Professor of the Department of Russian Language and Literature of the Tuvan State University.

Prokop'eva Alena Kirillovna, Candidate of Philology, Associate Professor of the Department of the Yakut language of the Institute of Languages and Cultures of the Peoples of the North-East of the North-Eastern Federal University named M. K. Ammosov.

Salchak Ajelita Jakovlevna, Candidate of Philology, Associate Professor of the Department of Tuvan Philology and General Linguistics of the Tuvan State University.

Sanchaj Chojgan Herel-oolovna, Candidate of Cultural Studies, lecturer of the Department of Philosophy of the Tuvan State University.

Sojan Ajlanmaa Myldyrgynovna, Candidate of Philology, Associate Professor of the Department of Tuvan Philology and General Linguistics of the Tuvan State University.

Tazranova Alena Robertovna, Candidate of Philology, Senior Researcher at the Institute of Philology SB RAS.

Tjuntешева Elena Valer'evna, Candidate of Philology, Senior Researcher at the Institute of Philology SB RAS.

Urtegeshev Nikolaj Sergeevich, Doctor of Philology, Leading Researcher at the Institute of Philology SB RAS.

Hertek Arzhaana Borisovna, Candidate of Philology, Senior Researcher at the Scientific and Educational Center «Turkology» of Tuvan State University.

Chemchieva Arzhana Petrovna, Candidate of Philosophy, Researcher at the Institute of Archaeology and Ethnography SB RAS.

Shagdurova Ol'ga Jur'evna, Candidate of Philology, Senior Researcher of the Institute of Philology SB RAS.

**Родные языки и культуры в современном
изменяющемся мире**

Научный журнал
№ 1 ● 2022

**Native Languages and Cultures
in the Modern Changing World**

Scientific Journal
№ 1 ● 2022

Главный редактор *Б. Ч. Ооржак*
Библиографическое редактирование *А. Дандыр*
Перевод *Ч. Б. Донгак*
Компьютерная верстка *Л. А. Ооржак*

Сдано в набор 01.02.2022. Формат 60×84 ¹/₈.

Издание цифровое. Усл. печ. л. 12,3. Уч.-изд. л. 11,3. Заказ № 1.

Учредитель:
ФГБОУ ВО «Тувинский государственный университет»
667000 Российская Федерация, Республика Тыва, г. Кызыл, ул. Ленина, д.36.
Тел/факс приемная ректора: 8(39422) 2-19-69; e-mail: tgu@tuvsu.ru