

Федеральное государственное бюджетное учреждение науки
«Калмыцкий научный центр Российской академии наук»

**Ежегодные научные чтения
Калмыцкого научного центра
Российской академии наук – VI
(г. Элиста, 22-23.12.2016)**

Элиста, 2017

УДК 94+81+31
ББК 63.3(2Рос=Калм)+81.2(2Рос=Калм)+60.5
Е 35

Утверждено к печати Ученым советом
Федерального государственного бюджетного учреждения науки
«Калмыцкий научный центр Российской академии наук»

Редколлегия:
д-р ист. наук *Э. П. Бакаева*, канд. фил. наук *В. В. Куканова*,
д-р ист. наук *Е. Н. Бадмаева*

Составитель:
канд. фил. наук *Е. В. Бембеев*

Е 35 **Ежегодные научные чтения Калмыцкого научного центра РАН – VI**
[электронное издание] / ред. Э. П. Бакаева, В. В. Куканова, Е. Н. Бадмаева;
сост. Е. В. Бембеев. – Элиста: КалмНЦ РАН, 2017. – 160 с.

В настоящем издании представлены тезисы докладов и сообщения, которые прозвучали на ежегодной научной сессии Калмыцкого научного центра Российской академии наук по итогам научно-исследовательской работы в 2016 г.

ISBN 978-5-906881-16-8
(e-pub)

© ФГБУН «Калмыцкий
научный центр РАН», 2017
© Коллектив авторов, 2017

- Сабрукова С. С. Текстологический анализ двух монгольских переводов «Бодхи-чарьяаватарь» *Bodisud nom-un yabudal dur oroqu neretü sastir orosibai* // *Mongolica*-VIII. СПб.: Петербургское востоковедение, 2008. С. 86–92.
- Сазыкин А. Г. Монгольские переводы Манджушри нама сангити. Сост. А. Г. Сазыкин. Киото, 2006. 265 р.
- Цендина А. Д. Два монгольских перевода тибетского сочинения «Книга сына» // *Mongolica* V. СПб.: Петербургское востоковедение, 2001. С. 54–74;
- Пюрбеев Г. Ц. Согласование в числе как грамматический прием в истории монгольских языков // V Международный конгресс монголоведов. Доклады советской делегации. М.: Наука, 1987. С. 118–126.
- Яхонтова Н. С. Ойратский литературный язык XVII в. М.: Вост. лит-ра, 1996. 152 с.

К ВОПРОСУ О СОЗДАНИИ ФОРМАЛИЗОВАННОГО ОПИСАНИЯ МОРФОЛОГИИ СТАРОКАЛМЫЦКОГО ЯЗЫКА В ЦЕЛЯХ РАЗРАБОТКИ СТАРОКАЛМЫЦКОГО ПОДКОРПУСА В НАЦИОНАЛЬНОМ КОРПУСЕ КАЛМЫЦКОГО ЯЗЫКА

В. В. Куканова¹, Е. В. Бембеев²

¹ кандидат филологических наук, директор КалмНЦ РАН

² кандидат филологических наук, заместитель директора по научной работе КалмНЦ РАН

Одной из актуальных задач калмыцких лингвистов является разработка старокалмыцкого подкорпуса, в прежнем названии подкорпуса «ранних» текстов в Национальном корпусе калмыцкого языка [Куканова и др. 2012]. Сейчас, однако, мы намеренно отказываемся от последнего названия, поскольку, во-первых, период, которым мы ограничиваем данный подкорпус, значительно шире: это отрезок времени с момента создания старокалмыцкой письменности (*тодо бичиг* «ясное письмо») до официальной ее отмены, т. е. до 1924 г., который можно условно разделить на два этапа: со второй половины XVII в. до начала XVIII в. и с начала XVIII в. до 1924 г. Мы не берем во внимание другие этапы развития ойратского языка [Павлов 2000: 20–21], поскольку считаем своей первой задачей оцифровку рукописей на *тодо бичиг* и введение их в корпус, дабы получить полную картину развития языка, который зафиксирован в одной графической системе, тем самым формируя репрезентативную базу данных однородных материалов.

Старокалмыцкий язык по своей графической системе значительно отличается от современного калмыцкого языка, хотя в целом современное словоизменение осталось практически в том же самом виде, что и было в период XIX и начала XX в. (за исключением ряда единичных случаев) в сравнении с морфологией, например, русского языка, когда кардинально перестроилась морфологическая система.

Объектом в нашем исследовании является линейная последовательность букв на *тодо бичиг*, транслитерированная на латиницу, в том виде, в котором она используется в рукописи или литографированном издании без какого-либо вмешательства со стороны ученого, проводящего транслитерацию [Куканова и др. 2013]. За основу взят словарь, подготовленный А.М. Позднеевым [1911], по причинам, которые были разъяснены в [Бембеев 2015].

Работа продвигалась по двум направлениям: 1) создание банка основ, где указывается частеречная принадлежность; 2) создание банка аффиксов и их последовательностей. Алгоритм, по своей сути, не отличается от алгоритма анализа, разработанного для парсера современных калмыцких текстов [Куканова, Каджиев 2014]. На наш взгляд, значительное отличие состоит в том, что на *тодо бичиг* отображаются так называемые в настоящее время «неясные» гласные, которые внесли некую «путанность» в определении словоизменительных типов современного калмыцкого языка, которые необходимы для уровня проверки результатов автоматического анализа в целях снятия омонимии разборов. Опыт работы над разработкой программного обеспечения для современного калмыцкого языка показал, что для успешной работы парсера необходима минимальная информация: основы с определениями частей речи, комплекс аффиксов, присоединяющихся к слову. Словоизменительные типы нужны для снятия омонимии, как было уже сказано выше.

Что касается банка данных по аффиксам и основам, то здесь мы столкнулись со следующей проблемой, имеющей место при создании любого корпуса рукописных или старопечатных книг, текстов: отсутствие единообразия как такового, другими словами, орфографической кодификации, что приводит к различному написанию тех или иных единиц в текстах или тексте даже у одного и того же автора (переписчика). К тому же мы берем тексты достаточно большого периода (почти три столетия), которые содержат в себе следы развития языка. К примеру, современный глагол *баах* 'быть' выглядел следующим образом *bauixu*, от которого отсоединили аффикс *-xu*, в результате мы имеем основы: *baui-*, *bi-*.

Другой проблемой является отсутствие инвентаря аффиксов, которые еще выявляются и описываются (см, например: [Гедесева 2012]). Надо сказать, что отсутствует полное описание системы старокалмыцкого языка, что в несколько раз усложняет работу над созданием корпуса.

При разработке формализованного описания морфологии старописьменного языка мы предпочли двигаться от текста к словарю основ и аффиксов, дополняя материал из указанного выше словаря, поскольку, имея некодифицированный материал корпуса, было проще начинать работу с постепенного ввода лексических единиц, имеющих в тексте.

Список слов, из которого дополняли словарь для парсера, состоит из 6 640 лексем. В качестве формальных признаков в определении части речи были использованы алгоритмы, связанные с финалями слов. Приведем несколько правил, которые применялись в этой работе.

1) если слово заканчивается на -kü, то это чаще всего глагол, за исключением ряда слов, таких как, например, börkü ‘сазан’, küŋkü ‘опытный’;

2) если слово заканчивается на -хи, то это чаще всего глагол, за исключением ряда слов, таких как, например, adxu ‘горсть’, tsaxu ‘козлы; крючок, сошки’, šandūgaxu ‘иначе’, šarxu ‘палевый’;

3) если слово заканчивается на -či, то это чаще всего существительное, за исключением ряда слов, таких как, например, ögönci ‘щедрый’.

Как видно, использовались формальные свойства слова для определения части речи, но, тем не менее, затем вручную проверялись списки слов на предмет выявления ошибки. На данном этапе были выявлены стеммы лексических единиц, которые в целом схожи со системой основ современного калмыцкого языка: dun ‘звук, голос; песня’ может иметь несколько основ: dūn-, dū-; to — to-, tō- и т. д. Кроме того, как мы отмечали выше, на письме *тодо бичиг* фиксировались «неясные» гласные, этот фактор существенно повлиял на выявление основ, поскольку, если в современном языке у лексем, которые имели одну основу, на конце имеется «неясный», не имеющий буквенного выражения, то в старокалмыцком языке такие лексеммы будут иметь две основы — с «неясным» гласным и без него, поскольку происходит чередование при словоизменении. Например: у лексемы sara ‘месяц, луна’ встречаются две основы: полная и «усеченная» — sar- и sara-:

1. (2)¹ zaqıbei dayin kelen albei tini ʏorban jılıyın (3) ʏöčin zurʏan sar-du erke biši nülgeji (4) ire zaqād balaman ulan beyiren ürün ʏazaran [Сказки на калмыцком языке]

¹ В круглых скобках указаны номера строк на листе.

2. (9) γazartu, Dalai Blamayin gegen zunai cagtu zalaradag bayiſing, Norbo lingқа gezi nerēdedeg bayinai. Uryātai (10) sad modon dotor barigsan örgö žigen. Zunai teriün qonin saradu tende zalarād, namuriyin dundu yaqai (11) saradu Bodalayin ordo qarſi-du zalaradag yuuman žigen. Mön ödöge tere Norbo lingқа gegči zusalanggiyin [Сказание о хождении в Тибетскую страну малодербетовского Бааза-багши]

В данных тезисах мы постарались кратко описать проделанную работу и проблемы, с которыми мы встретились, а также пути их решения. Замечания, высказанные здесь, носят предварительный характер, поскольку в ходе работы банк данных пополняется новой информацией, которая корректирует уже имеющуюся.

Источники

Национальный корпус калмыцкого языка [электронный ресурс] // URL: <http://kalmcorpора.ru> (дата обращения: 22.12.2016).

Литература

- Бембеев Е.В.* Сравнительный анализ «ранних» лексикографических памятников старописьменного калмыцкого языка в свете разработки морфологического анализа // Вестник Тувинского государственного университета. Педагогические науки. 2015. № 4 (27). С. 107–115.
- Гедеева Д.Б.* О падежном показателе -А: -О: -Е: в калмыцком письменном языке XVIII в. (на материале деловых документов) // Вестник Калмыцкого института гуманитарных исследований РАН. 2012. № 4. С. 97–99.
- Куканова В.В.* Морфологическая модель калмыцкого языка в свете автоматической обработки текстов: предварительные замечания // Информационные технологии и письменное наследие. ЕІ Manuscript-2012: Материалы IV международной научной конференции (Петрозаводск, 3–8 сентября 2012 г.). Петрозаводск, Ижевск, 2012. С. 142–146.
- Куканова В.В., Бембеев Е.В., Мулаева Н.М., Очирова Н.Ч.* Национальный корпус калмыцкого языка: архитектура и возможности использования // Вестник Калмыцкого института гуманитарных исследований РАН. 2012. № 3. С. 138–150.
- Куканова В.В., Каджиев А.Ю.* Алгоритм работы морфологического парсера калмыцкого языка // Писменото наследство и информационните технологии: Материалы от Умеждународна науч. конф. (Варна, 15–20 септември 2014 г.) / отг. ред. В.А. Баранов, В. Желязкова, А.М. Лаврентьев. София; Ижевск, 2014. С. 116–119.
- Куканова В. В., Бембеев Е. В., Музраева Д. Н.* К вопросу о кодировке UNICODE графической системы «тодо бичиг» и создании базы данных текстов на старокалмыцком языке // Вестник Калмыцкого университета. № 3. Элиста: КалмГУ, 2013. С. 65–75.

Павлов Д. А. Формирование и развитие калмыцкого национального литературного языка // Павлов Д. А. Вопросы истории и строя калмыцкого литературного языка. Сб. науч. ст. Изд. 2. Элиста: Калм. гос. ун-т, 2000. С. 17–37.

Позднеев А. М. Калмыцко-русский словарь в пособие к изучению русского языка в калмыцких начальных школах. СПб.: Типография Императорской академии наук, 1911. 312 с.

О СЮЖЕТАХ БОГАТЫРСКИХ СКАЗОК КАЛМЫКОВ В СВОДЕ ФОЛЬКЛОРА

*Б. Б. Манджиева*¹

¹ кандидат филологических наук, зав. отделом фольклора КалмНЦ РАН

В Калмыцком научном центре РАН авторским коллективом фольклористов: Т. Г. Басанговой [2007, 2013, 2014], Э. Б. Оваловым, Б. Б. Манджиевой [2015], Б. Б. Горяевой [2010, 2011, 2012], И. С. Надбитовой [2011, 2012], Д. В. Убушиевой [2010, 2011, 2013], Б. Х. Борлыковой [2014], Ц. Б. Селевой [2014, 2015], редактором Т. А. Михалевой подготовлен ряд томов многотомного «Свод калмыцкого фольклора», который представляет собой научную публикацию фольклорных произведений, имеющих огромное значение для сохранения, возрождения и обогащения культурного наследия народа. Целью Свода являются широкое введение в научный оборот неизданных материалов, раскрытие историко-культурной значимости, художественных особенностей, жанровой специфики публикуемых текстов.

В рамках фундаментального проекта «Свод калмыцкого фольклора» фольклористами продолжается научно-исследовательская работа по составлению томов: «Калмыцкие волшебные сказки», «Калмыцкие богатырские сказки», «Калмыцкие сказки о животных, бытовые, кумулятивные и небылицы».

Изучение особенностей отдельных жанровых разновидностей сказок в фольклористике имеет большое значение, поэтому выяснение специфики каждой из них способствует пониманию сказки как жанра. Богатырские сказки характеризуются многообразием сюжетов и персонажей, богатством изобразительно-выразительных средств, что делает их важной составной частью духовной культуры калмыков. До настоящего времени богатырская сказка калмыков еще не исследована в полной мере, между тем она представляет собой значимую часть калмыцкого сказочного фонда.