

Textbooks 1925 – 1930. The Mongolia Society, Inc. 322 Goodbody Hall Indiana University Bloomington, Indiana 47405-2401 USA. 1997.

8. Т. Намжил. «Өөрд моңһлын заң үүлин сойл» – аху эмдрлин боть. – Үрмч: Шинжэнэ ардын кевллин хора, 2010.

9. Хабунова Е. Э. Нулмт. – Элст: ЗАОр «НПП «Джангар». 2005.

10. Тамара Борджанова. Магическая поэзия калмыков. – Элиста: «Калмыцкое книжное издательство», 1999.

11. Х. Бада, Эрднь, Алт-Олгл. «Өөрд түүкин дурсхлмуд». – Үрмч: Шинжэнэ ардын кевллин хора, 1992.

12. «Жацһр» I боть. Нээрүлж кевллд диглж белдснь Тээвэ А, Хасгва Х, Бада Х, Дэмринжав К, Лота Н, Саһа, Бадм Т, Жамца Т, Эрнцэ Ч. – Үрмч: Шинжэнэ ардын кевллин хора, 1986.

13. Т. Намжл. «Өөрд моңһлын заң үүлин сойл» – аху эмдрлин боть. – Үрүмч: Шинжэнэ ардын кевллин хора, 2010.

14. Т. Г. Борджанова. Обрядовая поэзия калмыков. – Элиста: Калм. кн. изд-во», 2007.

15. Цацлин дееж. Составитель Намжавин Содмон. – Элиста: «Калмыцкое книжное издательство», 1997.

16. Саңжин Бося. Жилмүд. – Элста: Калм. кн. изд-во», 1976.

17. Саңжин Бося. Ээжин цэ. – Элст: Хальмг дегтр һарһач, 1980.

18. Бадмин Серэтр. Теегин дүр. – Элист: Хальмг дегтр һарһач, 1978.

19. А. Байн-Цаһан. Гергэ // Өнр цецг. – Көк-Хот: 2004. №. 7.

20. Осорин Утнасун. Күүкн хээр. – Элста: ЗАОр «НПП «Джангар», 2006.

\*\*\*

*Н.Ч. Очирова, В.В. Куканова,  
КИГИ РАН*

### **К ВОПРОСУ О РАЗРАБОТКЕ СЛОВАРНОГО МОДУЛЯ В НАЦИОНАЛЬНОМ КОРПУСЕ КАЛМЫЦКОГО ЯЗЫКА\***

*\*Статья подготовлена при поддержке проекта «Национальный корпус калмыцкого языка» подпрограммы фундаментальных исследований Президиума РАН «Создание и развитие корпусных ресурсов по языкам народов России» программы «Корпусная лингвистика» (2012–2014) и проекта РГНФ «Национальный корпус калмыцкого языка» (12-04-12047, тип «в»)*

Язык является важным средством самоидентификации и адаптированности личности в социальной среде. На сегодняшний день в мире существует большое количество языков, находящихся под угрозой исчезновения, и если сейчас не предпринять никаких действий по их сохранению, то не исключено, что списки десятков тысяч мертвых или вымерших языков могут пополниться. Решение задач по сохранению национального языка зависит только от нас – носителей языка.

Именно поэтому в условиях реальной угрозы исчезновения и с целью развития и исследования родного языка весьма актуальным представляется проект «Лингвистически аннотированный корпус калмыцкого языка» [1], одной из частью которого является «Словарный модуль калмыцкого языка». Надо сказать, что идея его создания появилась после знакомства с отчетом Программы фундаментальных исследований Президиума РАН «Корпусная лингвистика» «Совершенствование разметки и системы выдачи данных в Национальном корпусе русского языка» [2]. Наш проект во многом ориентирован на разработки, предложенные в этой работе.

Его создание необходимо в первую очередь для того, чтобы пополнить грамматический словарь новыми лексическими единицами, восполнить лакуны, появившиеся в результате перехода некоторых слов в пассивный фонд языка или даже утраты в целом их значений в лексической системе языка. Словари, которые были отобраны в качестве источников, относятся к тому периоду, когда калмыцкий язык был единственным средством познания и коммуникации, поэтому предполагается, что в них зафиксирован тот пласт лексики, который уже отсутствует в системе языка на современном этапе. Актуальность создания Словарного модуля в Национальном корпусе калмыцкого языка заключается также и в том, что компьютерные технологии и программное обеспечение поможет достичь самого главного – возможности ускорить работу и тем самым воплотить мечты о создании подобных систем в реальности.

Создание Словарного модуля позволит решить как фундаментальные теоретические и практические задачи, так и частные. Так, впервые будет введен лексикографический материал большого объема и разного времени в научные исследования, а также в процесс компилирования словарей различного формата и характера. Созданная словарная подсистема (или словарный модуль) значительно расширит лексикографическую базу для научно-теоретических исследований и, следовательно, углубит имеющееся описание лексики национального языка.

Теоретическая новизна заявленной темы заключается в использовании современных компьютерных технологий, а их активное внедрение в лингвистические исследования позволит поднять калмыцкое языкознание на более высокий уровень его развития, отвечающий требованиям времени. Ирактическая значи-

мость результатов по созданию Словарного модуля заключается в том, что будет осуществлена цифровая обработка ранних лексикографических источников калмыцкого языка; подготовлена электронная лексикографическая база данных, которая будет доступна любому пользователю сети Интернет; введены в оборот слова, отсутствующие в современных словарях калмыцкого языка; решена проблема эквивалентного написания слов в калмыцком языке и т. д. Разработка Словарного модуля может стать прекрасным материалом для обучения калмыцкого языка в дошкольных и средних образовательных учреждениях.

Словарная подсистема в Национальном корпусе калмыцкого языка развивается по двум направлениям:

1) по разметке и системе выдачи данных словарного модуля (структура модуля, интерфейс, поиск по заголовочному слову и расширенный поиск по зонам внутри словарной статьи, если такие имеются (в перспективе));

2) по интегрированию всех существующих словарей в единую поисковую систему, с удобным интерфейсом для поиска и восприятия информации.

В Словарный модуль на данный момент войдут лексикографические источники XVIII и начала XX вв., тексты которых будут интересны для изучения языковых особенностей национальной калмыцкой письменности «тодо бичиг» («ясное письмо»). Принято было решение в качестве первых источников использовать следующие словари калмыцкого языка:

- 1) П. Смирнов «Краткий русско-калмыцкий словарь» (1857 г.);
- 2) «Русско-калмыцкий словарь анонимного автора» (XVIII в.)
- 3) Г.А. Цвик «Handbuch dler Westmongolischen» (1853 г.);
- 4) К.Ф. Голстунский «Русско-калмыцкий словарь» (1860 г.);
- 5) М. Львовский «Калмыцко-русский словарь» (1893 г.);
- 6) А.М. Позднеев «Калмыцко-русский словарь» (1911 г.);
- 7) G.J. Ramstedt «Kalmückisches Wörterbuch» (1935 г.).

Для того чтобы реализовать поставленные задачи, необходимо оцифровать указанные словари, привести их к единообразию в оформлении словарной статьи. Тексты словарей будут представлены в специальном HTML-формате, которым выделяются элементы словарной статьи. Некоторые словари из вышеперечисленных являются переводными (калмыцко-немецкими), поэтому необходимо осуществить перевод словарных статей на русский язык, чтобы словари были доступны любому пользователю. Словарная подсистема в перспективе будет включать в себя следующие виды словарей:

- 1) толковые – общие и специальные (терминологические);
- 2) грамматические, морфемные и словообразовательные;
- 3) фразеологические;
- 4) синонимов;
- 5) антропонимов, топонимов и т.п.;
- 6) переводные (калмыцко-немецкий и др.).

Поскольку ранние словари калмыцкого языка написаны на старописьменном языке («тодо бичиг»), их обработка в цифровой формат и разметка в HTML затрудняются, к примеру «Краткий русско-калмыцкий словарь» П. Смирнова (1857 г.), обработка которого строится на следующих принципах (примеры см. ниже):

1) изображение слова на старописьменном калмыцком языке;

2) транслитерация на латиницу и кириллицу в соответствии с UNICODE, при этом их написание будет полностью соответствовать написанию на «тодо бичиг»;

3) современное написание лексических единиц (в соответствии с современной калмыцкой орфографией, т. е. без обозначения «неясных» гласных, долгих гласных и т. п. Слова, которые отсутствуют в калмыцко-русском словаре под ред. Б. Д. Муниева [3], будут помечаться специальным знаком «\*»;

4) написание авторского перевода П. Смирнова (с нашими уточнениями, для внесения ясности некоторым значениям слов);

ᠰᠤᠳᠤᠷᠢᠶᠢᠨ ᠭᠡᠷ	<i>suduriyin ger</i>	<i>судурийин гэр</i>	<i>судрин гэр</i>	‘библиотека’
ᠶᠷᠭᠢ	<i>urgi</i>	<i>өрги</i>	<i>өрг*</i>	‘анст’
ᠵᠠᠬ᠎ᠠ	<i>zaqa</i>	<i>заха</i>	<i>зах*</i>	‘берег’
ᠬᠣᠰᠤ	<i>xosu</i>	<i>хосу</i>	<i>хос*</i>	‘береза’
ᠮᠢᠰᠢᠶᠢᠯ	<i>mišiił</i>	<i>мишиүл</i>	<i>мишүл</i>	‘бечева’

5) разметка текста словаря в HTML-формате (формат предоставляет целый ряд преимуществ в использовании, упрощает справочно-поисковую систему так, что любому пользователю будет доступна вся информация из словарной статьи).

Итак, в словаре П. Смирнова размечаются следующие зоны:

1) заголовочное слово на современном калмыцком языке (в этом случае информация будет доступна для всех, в особенности и для тех, кто не владеет «тодо бичиг», к тому же в ранних словарях отсутствует единообразие в графическом оформлении слов) – <title>...</title>

2) написание на тодо бичиг— `<span class=“Todobichig”>...</span>`

3) написание на латинице – `<span class=“Latin”>...</span>`

4) написание на кириллице – `<span class=“Cyrillic”>...</span>`

5) переводной эквивалент – `<span class=“TE”>...</span>`

б) ссылки на другие статьи – `<span class=ref>...</span>`, `<a href="...">...</a>`

Образец размеченной статьи из словаря П.Смирнова

```
<p>
<title> холмтха</title>
<span class= Todobichig>ᠬᠣᠯᠮᠠᠬ᠎ᠠ</span> —
  <span class=Latin> xolomtoḡi</span> —
  <span class=Cyrillic >холомтоһи</span> —
  <span class=TE>брезгливый</span> —
  <span class=ref>см. также ...</span>, <a href="...">...</a>
</p>
```

Приведенный выше образец словарной статьи в HTML-формате – самый простой по своей структуре, разметка будет варьироваться в зависимости от словаря и наших задач. Конечной целью в создании Словарного модуля является объединение всех словарей в одну большую лексикографическую базу данных по калмыцкому языку, что во многом пополнит лексическую систему языка.

Таким образом, данный подход позволит обнаружить лакуны в калмыцких словарях, создание списка слов, отсутствующих в современных словаре, является важным дополнением к калмыцкой лексикографии. С другой стороны, пополнение словаря грамматического анализатора позволит усовершенствовать разметку корпуса и уточнить поиск по корпусу калмыцкого языка.

### Литература

1. Куканова В. В. Общая структура и перспективы использования Национального корпуса калмыцкого языка в свете проблемы репрезентативности // Материалы XL Международной филологической конференции (г. Санкт-Петербург, 21–25 марта 2011 г.). Вып. 24. СПб.: Изд-во С.-Петербург. ун-та, 2011. С. 125–137.
2. Совершенствование разметки и системы выдачи данных в Национальном корпусе русского языка: отчет по Программе фундаментальных исследований Президиума РАН «Корпусная лингвистика» [Электронный ресурс] // URL: [www.corpling-ran.ru/files/I/shajkevich.pdf](http://www.corpling-ran.ru/files/I/shajkevich.pdf) (дата обращения 20.12.2011).
3. Калмыцко-русский словарь / под ред. Б. Д. Муниева. М.: Изд-во «Русский язык», 1977.

\*\*\*

*А.В. Радионов,  
Калмыцкий государственный университет*

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПОСЛОВИЦ И ПОГОВОРОК В КАЛМЫЦКОМ И АНГЛИЙСКОМ ЯЗЫКАХ, ВЫРАЖАЮЩИХ БЛИЗКИЕ РОДСТВЕННЫЕ ОТНОШЕНИЯ

Паремиологический фонд языка – совокупность различных и территориальных по происхождению народных изречений. Паремиологические единицы способны к воссозданию круга наиболее важных для народной культуры понятий, составляющих ядро нормативно ценностной картины мира. Г. Л. Пермяков в паремиологическом фонде насчитывает 24 единицы [6; 255], среди которых можно выделить афоризмы, веллеризмы, загадки, пословицы; последние в данном списке стали объектом настоящего исследования. Однозначного определения пословицы в современной науке нет, равно как и нерешенным остается ряд иных вопросов. Мы согласны с дефиницией, предложенной А. В. Куниним, который определял пословицы как «афористически сжатые изречения с назидательным смыслом в ритмически организованной форме» [2; 431].

Вопросами функционирования пословиц в калмыцком языке занимались видные ученые-монголоведы Калмыкии: Т. А. Бертагаев, Г. Ц. Пюрбеев, Б. Б. Оконов и др. Пословичный фонд рассматривался с позиции структуры, семантики, этнографического анализа. Актуальность нынешней работы заключена в современном когнитивном подходе к анализу паремий, обусловившем их распределение по когнитивным (смысловым) группам, а также в применении квантитативного метода. Когнитивная структура есть схема репрезентации знания, опыта, реконструируемая на основе анализа ряда языковых единиц, в данном случае – пословиц. Основной функцией современной науки в целом и лингвистики, в частности, является обслуживание потребностей человека, ее антропоцентрическая направленность, что идеально репрезентировано материалом исследования – пословицами с включением терминов родства, которые выражают межличностные и внутрисемейные отношения.

Материал отбирался методом сплошной выборки по данным словарей калмыцкого [4; 7] и английского языков [1; 3; 9; 10; 11; 12]. Всего было выявлено 143 пословиц и их вариантов в калмыцком и 109 в английском языках. В количественном отношении пословицы, выражающие семейные связи, распределились следующим образом: родители parents (2), дети children (78), отец father (10), мать mother (12), сын son, дочь daughter; родители аав ээж, ээж аав, эцк эк, эк эцк (10); отец эцк (23); мать эк, ээж (23); дети үрн, көвүд, күүкд (37); сын көвүн, көвүн күн (14); дочь күүкн (25); дядя, племянник наһцх, зе (8); двоюродные братья и сестры бөл (3).