

УДК 81'373.47
ББК 81.23

**ЛЕКСИКОГРАФИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ СЛОЖНЫХ СЛОВ
(КОМПОЗИТОВ) В КАЛМЫЦКОМ ЯЗЫКЕ И ИХ ОТРАЖЕНИЕ
В НОРМАТИВНЫХ СЛОВАРЯХ***

В. В. Куканова

Определение лексической единицы — слова — вызывает множество дискуссий, несмотря на то, что уже достаточно давно описаны критерии, которыми оно должно обладать. Нет смысла приводить их в данной работе, тем более, что они хорошо известны [Шанский 1964: 20]. Напомним лишь, что признаков слова насчитывается двенадцать, среди которых можно выделить обязательные и факультативные.

Проблема отдельности слова является до сих пор еще не решенной, о чем свидетельствуют работы последних лет, в частности в области автоматической обработки текстов, например [Мустайоки, Копотев 2004]. Заметим, что делимитация слова в речевой цепи решается и интерпретируется по-разному в зависимости от целей исследования. Например, в прикладной лингвистике используется формальный подход: слово определяется как последовательность знаков, ограниченная пробелами [Гак 1990; Касевич 1977: 57–58]. Однако это не является решением проблемы для ряда языков. Так, посесивные частицы в калмыцком языке пишутся слитно со словом, например *эцкм* ‘мой отец’. Появляется вполне закономерный вопрос: вышеприведенный токен представляет собой сочетание двух слов или все же это одно слово, поскольку графически оно представлено в виде одной линейной последовательности символов, разделенной от других пробелами. Частица *-м* обладает некоторыми признаками, характерными для слова. Например, фонетической оформленностью, лексико-грамматической отнесенностью,

возобновляемостью и другими, но не обладает ударением, цельностью и единоформленностью.

С другой стороны, в языках существуют «сочетания» графических слов, которые, по сути, являются одним словом, несмотря на то, что его компоненты пишутся отдельно или через дефис. Каждый из этих элементов обладает собственным ударением или сочетанием главного и побочного ударений. В целом оно представляет собой одно лексическое значение. Это так называемые сложные слова, или композиты (компаунды — *compound words*), которые изобилуют в тюркских и монгольских, а также в английском, испанском, германском языках. Сложение двух самостоятельных слов как способ образования нового понятия весьма распространено и в калмыцком языке: «два слова входят в синтаксическую связь и образуют новое понятие» [Грамматика калмыцкого языка 1983: 87]. Ср.: словосложение¹ — это «...объединение слов с целью образовать новую лексическую единицу, или модифицировать зна-

¹ Следует различать словосложение и основосложение. В первом случае соединяются слова для образования новых слов (билексем), во втором — основы слов при помощи соединительной гласной, или интерфикса) [Хашимов 2008]. Калмыцкому языку основосложение как способ образования новых слов не свойственно, хотя, конечно, результаты основосложения присутствуют и здесь, но являются заимствованными элементами или калькой слов из русского языка. Например, *электробеш* ‘электропечка’, *электрокөрә* ‘электропила’.

* 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: создание и разработка» (12-04-12047/в).

чение существующих лексических единиц, или выразить грамматическое значение (форму и функцию). При словосложении достигается такое единство сочетающихся слов, что может произойти фонетическая трансформация и структурно-звуковое слияние слов-слагаемых» [Муратов 1962: 23]. Вслед за А. Г. Садыковой, под композитами мы понимаем «структурно-сложные полифункциональные единицы прямой или косвенной номинации, разновидность деривата, образованного по стандартной модели (из двух или более корневых морфем) по аналогии, а также путем номинации (семантической конденсации) синтаксического словосочетания» [Садыкова 2002: 7].

К рассмотрению сложных слов мы обратились по двум причинам: в Калмыцком институте гуманитарных исследований Российской академии наук ведутся работы по созданию Национального корпуса [Куканова и др. 2012] и Электронного словаря калмыцкого языка со звуковым модулем [Куканова, Бембеев 2012]. Каждый из этих проектов требует настоящего решения проблемы: и здесь, и там необходимо определить статус сложных слов, поскольку для статистического анализа важно выявить единицу описания. Формально мы имеем дело с токенами, т. е. последовательностью знаков между пробелами, но не всегда токен является словом, иначе говоря, он чаще всего совпадает со словом, однако сам может являться частью слова. Но для того чтобы получить достоверную статистику, например частотный список слов (не токенов), по корпусу текстов на калмыцком языке, необходимо учитывать и такую особенность языка, как наличие большого количества сложных слов. Только в этом случае можно получить достоверную картину частотности лексических единиц калмыцкого языка. В плане создания электронного словаря необходимо решить, каким образом представлять сложные слова в словаре: внутри словарной статьи слова (или слов), образующих сложное слово, или как отдельную вокабулу.

В калмыцком языке существует несколько типов компаундов с точки зрения их графической оформленности: одни пишутся через дефис, что несколько облегчает их выделение в линейном потоке; дру-

гие слова — отдельно, т. е. через пробел. С практической точки зрения, и в том, и в другом случае морфологический анализатор разбирает их как отдельные слова, вернее как разные. С теоретической же — выделение сложных слов менее разработано в лексикографии калмыцкого языка², и важно решить эту проблему в самое ближайшее время, поскольку в Республике Калмыкия сложилась ситуация постепенной утраты языка как средства познания и коммуникации. Как известно, если носители того или иного языка перестают осуществлять коммуникацию на своем языке, не говоря и о когнитивных функциях языка, то это приводит, как правило, к исчезновению языка как такового: система перестает быть самодостаточной (в тех или иных коммуникативных ситуациях люди начинают прибегать к использованию средств из другого языка) и уже сам язык не развивается, поскольку нет носителей, которые являются, с одной стороны, «причиной» эволюции языка, с другой стороны, катализатором языковых процессов и, с третьей стороны, речь носителей языка закрепляет результаты этих изменений.

По мнению Ф. А. Ганиева, образование сложных слов в тюркских языках хронологически более поздний процесс по сравнению с сингармонизмом [Ганиев 1976: 35]. Видимо, в монгольских языках это тоже процесс позднего периода развития языка, поскольку среди сложных слов можно найти примеры, демонстрирующие отсутствие уподобления гласных по признаку ряда: *диг-даран* 'порядок', *эдл-ууш* 'продукция', *мек-тах* 'обман'.

Определение сложных слов тесно связано с вопросами о том, на каких основаниях мы выделяем сложные слова, какими критериями они должны обладать, чтобы отнести их к данной группе «слов», какие из них носят обязательный, а какие факультативный характер. В связи с вышесказанным первоначально необходимо сравнить методики выделения подобных единиц в

² Как справедливо отмечает Г. К. Бадмаева, сложные слова фиксируются без указаний критериев их выделения [Бадмаева 2009: 19], хотя проблема сложных слов достаточно разработана в общей лингвистике (см., например, [Ефремова 2012; Маковой 2009; Снегова 2011; Черкасова 2010]).

разных языках, прежде всего в английском, и описать принципы работы применительно к калмыцкому языку.

Само понятие «компаунд» не совсем четко разработанный термин, поскольку ученые до сих пор не пришли к единому мнению о его трактовке и даже статусе. Часть лингвистов, преимущественно зарубежных исследователей, считают, что компоненты, которые образуют сложное слово, являются самостоятельными и соответственно разложимыми [Пауль 1960; Jespersen 1954; Kruisinga 1932]. Остальные наделяют сложные слова самостоятельным статусом, наряду с простым словом ([Арнольд 1986; Жданкин 1986; Муратов 1962; Садыкова 2002] и др.).

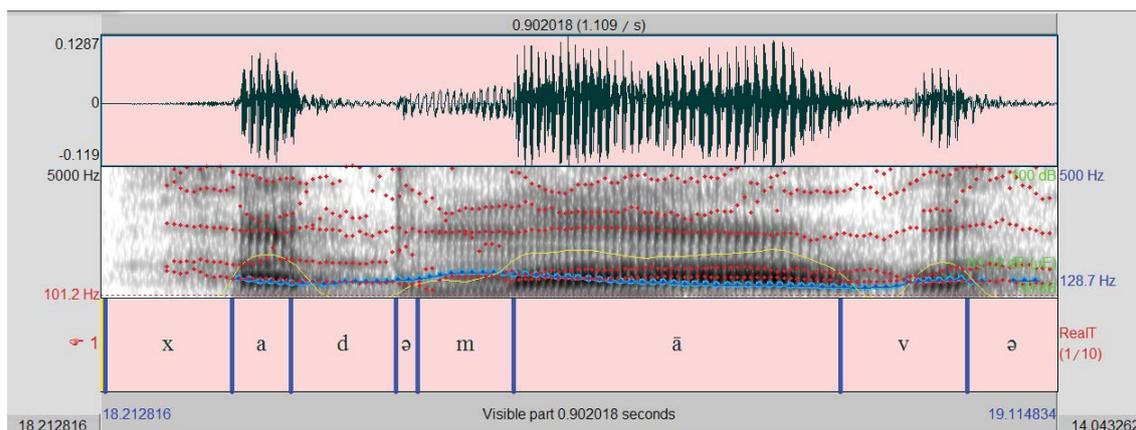
Следующим спорным моментом является вопрос о границах сложных слов: одни подчеркивают, что компаунды — это только комбинация двух и более «свободных» слов, т. е. равных по отношению к друг другу (например, «A compound word is usually understood to be the result of the (fixed) combination of two free forms, or words that have an otherwise independent existence... (Adams, 1973: 30)»). Другие же расширяют рамки этого понятия, утверждая, что они являются соединением двух любых слов, в том числе и зависимого компонента с главным («When two (or more) elements which could potentially be used as stems are combined to form another stem, the form is said to be a compound... (Bauer, 1983: 28)»). Следующим спорным моментом является длина компаунда: большая часть ученых относит к компаундам комбинацию только двух компонентов («...a compound is a word that consists of two elements, the first of which is either a root, a word, a phrase, the second of which is either a root or a word... (Plag, 2003 :135)»), другие же подчеркивают, что в языке существуют компаунды из более двух слов («A compound is a word which consists of two or more words... (Fabb, 1998: 66)»; «Compounding is a process by which a compound lexeme is derived from two or more simpler lexemes... (Matthews, 1991: 82)» [цит. по: Dottorato 2006: 2].

Другой проблемой, с которой сталкиваются все лингвисты, является проблема разграничения сложного слова и словосо-

четания, до сих пор еще не получившая в силу разных причин однозначного решения. В языках разных структур необходимость и острота разграничения композитов и словосочетаний зависит от их структурно-типологических особенностей: «...критерии дифференциации сложного слова и синтаксического словосочетания обычно проявляют большую зависимость от конкретного языка и его грамматического строя. В языках флективного типа, например, распознавание сложных слов не представляет собой трудности, напротив, в так называемых аналитических языках грани между словами и словосочетаниями не всегда достаточно отчетливы» [Кубрякова 1965: 27]. Возьмем к примеру сложные слова в русском языке: существуют специальные материально выраженные маркеры сложного слова (интерфиксы, дефисное написание). Калмыцкий язык принадлежит к группе монгольских языков, которые по своей структуре являются агглютинативными. Здесь большинство композитов образуется примыканием не основ, а отдельных слов. Структурная модель сложного слова в калмыцком языке аналогична структурной модели синтаксического словосочетания, что затрудняет их разграничение.

Тем не менее, несмотря на многочисленные споры, ученым удалось выделить критерии, позволяющие отнести комбинацию слов к одной лексической единице [Панфилов 1958: 107]. Рассмотрим некоторые из них:

1) фонетический критерий: наличие ударения. Некоторые слова могут получать, помимо основного, еще и побочное ударение, как, например, в русском языке. В калмыцком языке до сих пор еще выяснена природа самого ударения. Однако, по наблюдениям за речью носителей, в некоторых сложных словах выделяются два слога, а в других — один слог. В связи с вышесказанным данный критерий является нерелевантным для выделения сложных слов в калмыцком языке. На осциллограмме, которая приведена ниже, видно, что между конститuentами компаунда отсутствует пауза: слово произносится как одно.

Рис. Осциллограмма сложного слова *хадм аав*

- 2) **графический критерий:** орфографическая запись. Как уже было сказано выше, отсутствует унификация орфографического оформления сложных слов³: *ааһ-сав* ‘посуда’, *бүрн-төгс* ‘полный; полностью’, *хот-хол* ‘пища’; *хар модн* ‘дуб’, *хөн бөөлжрһн* ‘малина’, *темән чикн* ‘щавель’. Признак является нестабильным;
- 3) **морфологический критерий:** отнесенность к той или иной части речи и стирание грамматического значения одного из компонентов, чаще всего первого. Более релевантный критерий по сравнению с двумя предыдущими. Как правило, достаточно легко определяется частеречная принадлежность сложного слова: *...хойр ик күргн хадм эцкүр ирэд, уурлжана...* ‘...два старших зятя, придя к свекру, ругают...’⁴ [Буутан С.];
- 4) **синтаксический критерий:** синтаксическая роль в предложении и фиксированный порядок компонентов сложного слова. Это также релевантный признак. Сложное слово невозможно разделить на составляющие его компоненты без потери смысла: *Терүнд бер, көвүн, хадм-эцк, эк, гижгтә күүкд - цуһар бәэдг...* ‘Там все были: невестка, сын, **свекор**, свекровь, незамужние девушки...’ [Доржин Б.];
- 5) **ономасиологический критерий:** семантическая цельность сложного слова. Наиболее «работающий» критерий, позволяющий дифференцировать сложное слово от словосочетания. Так, например, слово *хадм эцк* обозначает один денотат.

Сложные слова в калмыцком языке традиционно делятся на две группы⁵: парные (копулятивные) и составные (детерминативные). К первой группе относятся слова, которые объединяются в одно целое на равных правах, при помощи сочинительной связи между компонентами. Составные же слова связаны между собой по принципу примыкания или управления, например, их структура может выглядеть следующим образом: Adj+Noun, Noun(Gen/Ass)+Noun и др.

Рассмотрим пример оформления словарной статьи в Калмыцко-русском словаре под редакцией Б. Д. Муниева [Калмыцко-русский словарь 1977: 17].

⁵ Автор работы предпочитает отличать сложные слова от словосражений, считая, что слова типа *эндр* ‘сегодня’ (*эне* ‘этот’ + *өдөр* ‘день’) не являются сложным словом, а являются результатом словосращения, произошедшего в силу разных причин (в данном случае сращение обусловлено частотностью использования этого сочетания в речи, а также вследствие действия принципа экономии речевых усилий [Мартине 1960], в результате фонетических процессов — выпадения гласных звуков — появилось новое слово, морфемный состав которого можно установить только путем этимологического анализа).

³ См. также примеры, которые приведены ниже. В них по-разному зафиксировано одно и то же сложное слово *хадм эцк* (через пробел и дефис).

⁴ Примеры из фольклорных и художественных произведений, приведенные в данной работе, взяты из материалов Национального корпуса калмыцкого языка.

аав [аавъ] 1) отец, папа; **аав үкхлэ** — **ааһ герэсн**, **акт үкхлэ** — **арсн герэсн** *полс.* когда отец умирает, на память остается его чашка, когда конь гибнет, на память остается его шкура; **аавиннь эмдд кү тань**, **актинь сээнд һазр тань** *полс.* узнавай людей, пока жив отец, узнавай просторы, пока резв конь; **хадм аав** [здесь и далее подчеркивание автора. — В. К.] свекор; 2) дед, бабушка; **көгшин аав** дед (*по отцовской линии*); **наһц аав** дед (*по материнской линии*)

Как видно, выделенные слова в данной словарной статье пишутся через пробел. Лексические единицы *хадм аав* ‘свекор’ (‘родня по жене’+‘отец’), *көгшин аав* ‘дед по отцовской линии’ (‘старый; старик, старушка’+‘отец’), *наһц аав* ‘дед по материнской линии’ (‘родные по материнской линии’+‘отец’) имеют общий компонент *аав*, при этом при словосложении лексическое значение слова модифицируется: сема “отец” является объединяющей, а первые компоненты композитов создают дифференцирующие семы в значении⁶. Фонетический и графический критерии не являются точными при их отличии от сочетаний слов в данных случаях. Морфологический и ономаσιологический признаки наиболее ценны здесь, так как во всех случаях каждое сложное существительное обозначает свой собственный денотат. Ср.: «единственным достаточно надежным критерием отличения сложной основы, в частности, атрибутивного или парного основосложения от аналогичных по внешней структуре свободных синтаксических сочетаний слов, остается семантическое единство сложной основы, которое внешне проявляется в использовании ее в качестве исходной основы при словообразовании» [Наделяев 1988: 37].

Напомним, что выделенные выше (подчеркнутые) слова пишутся через пробел, и соответственно для того, чтобы найти перевод того или иного сложного слова, необходимо понять, какой из конституентов является главным, если таковой имеется, и проверить словарную статью на данное слово или проверить две статьи на наличие перевода, если элементы, образующие сложное слово, являются равнозначными по отношению к друг другу. Все это занимает достаточно много времени.

⁶ Сложные слова построены по принципу «целое + часть».

Целесообразно ввести правило орфографического оформления сложных слов: либо писать через дефис, либо вместе, без пробела, как, например, в английском. Возможно, следует дифференцировать написание сложных слов в зависимости от их типа. Например, аппозиционные и биномные сложные слова следует писать через дефис, а остальные графически оформлять как одно слово. Представляется, что этот фактор поможет быстрее читателю ориентироваться в словаре, что немало важно в аспекте сохранения языка.

Для решения проблемы в корпусе текстов на калмыцком языке необходимо создать словарь сложных слов, содержащий все возможные варианты написания. По нашему мнению, анализатор должен сначала обработать массив текстов с целью приписывания леммы каждой словоформе, затем при помощи отдельного модуля и статистических обработок текстов на рекуррентность элементов в текстах с применением 2-х и более n-граммного подхода и опорой на словарь найти сложные слова и связать их в единое целое. Только в этом случае мы сможем получить «чистую» статистику по корпусу калмыцкого языка.

Проблема анализа сложных слов в калмыцком языке требует специального изучения с привлечением большого количества фактического материала. Данная работа призвана лишь выявить первоочередные проблемы исследования сложных слов в лексикографическом аспекте калмыцкого языкознания. Однако уже очевидна необходимость выделения сложных слов в виде отдельной вокабулы в словаре, а в статьях элементов, образующих сложное слово, делать ссылки на данную запись, где оно будет истолковано и будет дан переводной эквивалент (см., например: [Куканова, Бембеев 2012]).

Литература

- Арнольд И. В. Лексикология современного английского языка: учеб. для ин-тов и фак. иностр. яз. 3-е изд., перераб. и доп. М.: Высш. шк., 1986. 295 с.
- Бадмаева Г. К. Критерии определения сложных слов в калмыцком языке (в сопоставлении с английским) // Гуманитарные исследования. 2009. № 4(32). С. 19–22.
- [Буутан С.] — Буутан Санжин туульс (Сказки Санджи Бутаева). Записи 1971–1978 годов: в 2-х кн. Кн. 1. Сер.: Өвкнрин зөөр («Сокровища предков»). Элиста: КИГИ РАН, 2008. 308 с.
- Гак В. Г. Слово // Лингвистический энциклопедический словарь / под ред. В. Н. Ярцевой. М.: Советская энциклопедия, 1990 [электронный ресурс] // URL: <http://tapemark.naod.ru/les/464c.html> (дата обращения: 07.03.2012).
- Ганиев Ф. А. К вопросу о сложном слове в современном татарском языке // Советская тюркология. 1976. № 4. С. 31–37.
- Грамматика калмыцкого языка: фонетика и морфология / Калмыцкий институт языка, литературы и истории. Элиста: Калм. кн. изд-во, 1983. 336 с.
- [Доржцин Б.] — Доржцин Б. Теегин цолд. Элст: Барин гер “Герл”, 2008. 814 х.
- Ефремова Е. Субстантивные модели многокомпонентных композитов в современном английском языке // Вестник Московского государственного областного университета. Серия «Лингвистика». 2012. № 3. С. 54–58.
- Жданкин И. В. К вопросу о сложном слове без морфологических признаков словесности (на материале китайского языка) // Вопросы языкознания. 1986. № 1. С. 99–111.
- Калмыцко-русский словарь / под ред. Б. Д. Муниева. М.: Русский язык, 1977. 768 с.
- Касевич В. Б. Элементы общей лингвистики. М.: Наука, ГРВЛ, 1977. 177 с.
- Кубрякова Е. С. Что такое словообразование. М.: Наука, 1965. 78 с.
- Куканова В. В., Бембеев Е. В. К проблеме электронных словарей калмыцкого языка // Информационные технологии и письменное наследие. E1’Manuscript-2012: Мат-лы IV Междунар. науч. конф. (Петрозаводск, 3–8 сентября 2012 г.). Петрозаводск; Ижевск, 2012. С. 146–149.
- 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.
- Маковой Р. Г. Соотношение сложного слова и словосочетания // Вестник Харьковского национального автомобильно-дорожного университета. 2009. № 45. С. 14–18.
- Мартине А. Принцип экономии в фонетических изменениях (Проблемы диахронической фонологии). М.: Иностран. лит-ра, 1960. 260 с.
- Монраев М. У. Калмыцкие личные имена (семантика). Изд. 2-е, перераб. и доп. Элиста: Калм. кн. изд-во, 2007. 224 с.
- Муратов С. Н. Устойчивые словосочетания в тюркских языках. М.: Вост. лит., 1962. 130 с.
- Мустайоки А., Копотев М. К вопросу о статусе эквивалентов слова типа *потому что*, в зависимости от, к сожалению // Вопросы языкознания. 2004. № 3. С. 88–107.
- Наделяев В. М. Современный монгольский язык. Морфология. Новосибирск: Наука, 1988. 112 с.
- Панфилов В. З. Сложные существительные в нивхском языке и их отличие от словосочетаний (к проблеме слова) // Вопросы языкознания. 1958. № 1. С. 105–111.
- Пауль Г. Принципы истории языка. М.: Иностран. лит., 1960. 500 с.
- Садькова А. Г. Система субстантивного словосложения в тюркских и германских языках в сравнительно-типологическом аспекте: на материале татарского, киргизского, туркменского, турецкого, английского, немецкого языков: дис. д-ра филол. наук. Казань, 2002. 492 с.
- Снегова Е. П. О лексикографическом представлении сложносоставных слов с иноязычным компонентом // Вестник Нижегородского университета им. Н. И. Лобачевского. 2011. № 6(2). С. 643–646.
- Хашимов Р. И. Билексема как особая единица языка // Вопросы филологии. 2008. № 4. С. 30–38.
- Черкасова А. А. О некоторых грамматических признаках билексемы как особой единицы языка // Вестник Нижегородского университета им. Н. И. Лобачевского. 2010. № 3(1). С. 347–350.
- Шанский Н. М. Лексикология современного русского языка. М.: Просвещение, 1964. 315 с.
- Dottorato di T. Compound Adjectives in English: dissertation. CICLO, 2006. VI+259 p.
- Kruisinga E. A Handbook by Present-day English. P. II. Groningen, 1932. 360 p.
- Jespersen O. A. Modern English Grammar on Historical Principles. Part IV. Morphology. London: G. Allen and Unwin., Copenhagen: K. Munbesgaerd, 1954. 570 p.