

*Электронный морфемно-орфографический
словарь как база для создания
морфологического анализатора для
Электронного корпуса тувинского языка.*

Бавуу-Сюрюн Мира Викторовна.
г.Кызыл, Тувинский госуниверситет

Электронный морфемно-орфографический словарь был создан на базе Орфографического словаря тувинского языка (1967), был дополнен современными материалами. Содержит около 30 тыс. словоформ, разбитых на морфемы.

Работа была начата еще в конце 90-х годов по настоянию первого ректора ТувГУ, выпускника мехмата МГУ Бузур-оола Орлана Базыровича.

Для последующего автоматизированного поиска
нужных форм были введены следующие знаки:
= перед словообразовательными аффиксами
(в словарях тувинского языка глагол дается в форме на
-ар)

Ажыл=да+ар ‘работать’

/ - перед формообразовательными аффиксами (залога, вида, отрицания, лично-оценочных (уменьшительно-ласкательных), степеней сравнения прилагательных)


Тур/гус ‘заставить встать’

номчу/гула ‘почитывать’

тур/ба ‘не вставай’

оол/ак ‘мальчонка’

ак/сымаар ‘белесоватый’



+ - перед формоизменятельными аффиксами (лица, числа, падежа, времени, наклонения, причастия, деепричастия).

* - перед основами, где возможно произошли исторические изменения в их морфемном составе.

Знак # указывает, что имеется частица со слитным написанием (перед частицей): *ийик#не* ‘так ведь’

“ перед дополнительными звуками, звуковыми изменениями в составе основы: *ном+у‘н+да* ‘в книге его’

Знак // указывает, что имеется чередование звуков:

Ат+ы > ады т// д ‘имя его’

& или @ - сложное слово с отдельным написанием

&уран чүүл ‘искусство’

Имеется целый ряд случаев, когда одна и та же основа может иметь несколько иное графическое оформление. Некоторые основы при прибавлении аффиксов принадлежности могут видоизменяться, что обусловлены историческими процессами:

А) аас +ы > акс=ы, оол +у > огл=у,

Б) эгин +и > экт=и, оол +у > огл=у

В) орун +у > орн=у, чарын +ы > чарн=ы.

Чаш+ы чажы чередование ш//ж


ук+у угу чередование к//г, но в 2-сложных словах
выпадает согласный идик+и – иди+и

идик

Следует также брать во внимание случаи разного обозначения долгих гласных. Как правило, долгие гласные обозначаются а) повторением одной и той же графемы


Долгий [e:] обозначается графемой э:

сири=лэ+эр а в основе слова краткий гласный обозначается буквой е сири=ле, но прибавлении аффикса с инициальной гласной е возникает долгий гласный, который обозначается двумя ээ: сири=лэ+эр ‘дрожать’.



Орфографические правила таковы, что у одного и того же корня может получиться разное графическое оформление. Например, при обозначении фарингализации в 9 словах *аът, эът, оът, каът, дүъш, аъш-чем, чаъс, чөъп, чүък*, из них в 4 словах *аът, эът, оът, каът* пишется постоянно, в остальных случаях в производных от них основах Ъ не пишется *дүштеки, дүштээр* и т.д

В этом случае основы могут быть распознаны как разные.



При прибавлении к основам, оканчивающимся на согласный Й аффиксов с начальным гласным появляется так называемые йотированные гласные: сарба+яр (> сарбай+ар). Они в словаре даны с пояснениями.



европа (европа чурттары)

европей=жи

информас=тыг

информация

В словарь не входят словоформы, не представляющие трудности в своем правописании. Но с точки зрения морфемной структуры они заслуживают внимания. Такова форма древнего звательного падежа в определенном списке слов, т.е. в терминах родства и свойства: ава+й, угба+й, акы+й. В то же время в слове *ире+й произошло опрощение, в современном литературном языке выступает как неразложимое слово.

Заключение

- * Нужно прежде всего очень тщательно исследовать морфонологию тувинского языка.
- * Пополнить базы данных, на основе которых составить наиболее полный глоссарий.