

О возможностях расширения инвентаря
универсального метаязыка для грамматического
гlossирования текстов на алтайских языках

“Лейпцигский стандарт”: 84 граммемы

Этого инвентаря недостаточно для практических целей, в том числе, при документировании тюркских, монгольских и тунгусо-маньчжурских языков

Расширение: добавлено 350 аббревиатур – имён грамме.

Материал для такого расширения: терминологические указатели к книгам

1. И. А. Мельчук «Курс общей морфологии»
2. В. А. Плунгян «Общая морфология»

А. Части речи и грамматические категории

прилагательные	A.	adjective
наречия	ADV.	adverbial
аспектуальный & вид	ASP.	aspectual & aspect
падеж	C.	case
Союзы	CNJV.	conjunction
словообразовательный, деривационный	DER.	derivative
модальность	MOD.	modality
наклонение	MOOD.	mood

Именные части речи, имена	N.	nominal
числительное	NMR.	numeral
число	NUMB.	number
лицо	PRS.	Person
местоимение	PRON.	pronoun, pronominal
частицы	PTCL.	particle
существительные	S.	substantive
таксис	TXS.	taxis
глаголы	V.	verbal
залог	VOX.	voice

Б. Значения признаков

аллатив		ALL	allative
адмиратив		ADMIR	admirative
совокупность		AGGR	aggregate
отчуждаемость		ALP	alienable possession
предшествование		ANTER	anterior
аорист		AOR	aorist
ассертив утверждение	&	ASS	assertive & assertion
атрибутив		ATTR	attributive

каритив		CAR	caritive
компаратив		COMP	comparative
нереальный, ирреальный, ирреалис		IRRL	irreal
сослагательное конъюнктив	&	CNJ	conjunctive & subjunctive
кунктатив		CUNCT	cunctative
дебитив (=дебитатив)		DEBT	debitive
делимитатив		DELIM	delimitative
уменьшительность	&	DIM	diminutive



Модели тюркских словоформ для автоматического парсинга

Единая порядковая модель для словоформ
разных изменяемых частей речи

При необходимости включаются циклы

Вводится морфонологический компонент:
морфемы берутся в морфонологической
записи и обрабатываются правилами
фонологической сочетаемости

Парсер – конечный автомат

• Древнетюркская словоформа

Table 4: Model of Old Turkic wordform (based on Erdal 2004 and textes in VATEC)

0	1	2	3	4	5	6		7	8	9	10	11		12	13	14	15
S	Neg	Tense/Mood	Num ₁	Poss ₁	Coord ₁	Case ₁		At ₁	Modif ₁	Num ₂	Poss ₂	Coord ₂	Case ₂		At ₂	Modif ₂	Person
						Simple declension	Possessive declension						Simple declension	Possessive declension			
	Neg -mA	Neg.Indir -mAdUk Neg.Sequ -mAIn Neg.Sequ -mAIl Neg.Praes -mAz Neg.Praes -mAs Neg.Fut. -mAEl Indir -mIš Sequ -(X)p Sequ -(X)p.An Perf -dUk Perf -tUk Res -yUk [Praet -d Praet -t] Praes -(y)Ur Praes -(I)r Praes -(A)r Fut. -gAy Fut. -kAy Fut. -dAEl Fut. -dAEl FutIm -gAllr FutIm -kAllr Inf -mAk PrImpf -(X)gmA PrAct -(X)gIl PrHab -gAn PrHab -kAn PrAuct -gUel PrProsp -sXk	P1 -IAr	Poss1.sg -(X)m (X)n Poss2.sg -(X)g (X)g Poss2.sg -(X)g (X)g Poss3 -(s)l(n-) Poss1.pl -(X)mXz (X)nXz Poss1.pl -(U)mUz Poss2pl -(X)gXz (X)gXz Poss2.pl.Pol -(X)gXz.lAr (X)gXz.lAr Poss2.pl.Pol -(X)gXz.lAr	Coord -Il	Gen -(i)Xy Gen -nXy Gen -nXg Gen -nUy Acc -(X)g Acc -nl Dat -kA Dat -gA Loc -tA Loc -dA Abl -dIn Abl -dIn Abl -dAn Abl -tAn Instr -(X)n Equ -dA Dir -gArU Part -rA Simil -IAjU Simil -eU.IAjU Comit -IXgU	Gen -(i)Xy Gen -(i)Xy Poss3.Acc -nl Poss3.Acc -In Dat -kA Dat -gA Dat -tA Loc -dA Loc -dA Abl -dIn Abl -dIn Abl -dAn Abl -tAn Instr -(X)n Equ -dA Dir -gArU Part -rA Dir -ArU Part -rA Simil -IAjU Simil -eU.IAjU Comit -IXgU	At -kl	Rel -IXg Carit -sXz Comp -dAg Dimin -kyA Dimin -klhA	P1 -IAr	Poss1.sg -(X)m Poss2.sg -(X)g Poss2.sg -(X)g Poss3 -(s)l(n-) Poss1.pl -(X)mXz (X)nXz Poss1.pl -(U)mUz Poss2.pl -(X)gXz (X)gXz Poss2.pl.Pol -(X)gXz.lAr (X)gXz.lAr	Coord -Il	Gen -(i)Xy Gen -nXy Gen -nXg Gen -nUy Acc -(X)g Acc -nl Dat -kA Dat -gA Loc -tA Loc -dA Abl -dIn Abl -dIn Abl -dAn Abl -tAn Instr -(X)n Equ -dA Dir -gArU Part -rA Simil -IAjU Simil -eU.IAjU Comit -IXgU	Gen -(i)Xy Gen -nXy Acc -nl Acc -In Dat -gA Dat -A Loc -tA Loc -dA Abl -dIn Abl -dIn Abl -dAn Abl -tAn Instr -(X)n Equ -dA Dir -ArU Part -rA Simil -IAjU Simil -eU.IAjU Comit -IXgU	At -kl	Rel -IXg Carit -sXz Comp -dAg Dimin -kyA Dimin -klhA	1.sg -mAn 1.sg -bAn 2.sg -sAn 1.pl -bXz 1.pl -mXz 2.pl -sXz 2.pl.Pol -sXz.lAr 3.P1 -IAr Imp.1.Sg -(A)yIn Imp.2.Sg -0 Imp.2.Sg.Emph. -gIl Imp.2.Sg.Emph. -eU Imp.3 -zUn Imp.3 -sUn Imp.3 -eUn Imp.1.P1 -(A)llm Imp.2.P1 -(X)g Imp.2.P1.Pol -(X)g.lAr Imp.3.P1 -dUn.lAr Imp.3.P1 -sUn.lAr Imp.3.P1 -eUn.lAr Praet.1.sg -dXm Praet.1.sg -tXm Praet.2.sg -dXg Praet.2.sg -d(X)g Praet.2.sg -lXg

- Двойной набор именных слотов

män-iŋ-lär (я-Gen-Pl) *ol* ‘они мои’

biz-iŋ-tä-ki-čä (мы-Gen-Loc-Attr-Equ) ‘как в одном из наших’

bar-miš-lar-niŋ-in-da (идти-Indir-Pl-Gen-3pros-Loc) ‘от одного из тех, кто шел’

kut-in-liŋ (доля-3pros-Rel) ‘принадлежащий к его богатству’

Киргизская словоформа

1	0	1	2	3	4	6	7	8	9			10	11
Super	S	Neg	Tense/ Mood	Atr	AbstrNom	Number	Poss	PosNom	Case			Person	Number
									Simple declension	Possessive declension 1-2 Sg	Possessive declension 3&PosNom		
dup+п		Neg - бА	Indicative Present Simple -А/ й	Comp (Y)пААк	лYк	Pl лAp	Poss1.s g -(Y)м	PosNom м нYкY	Gen -∂Yн	Gen -∂Yн	Gen -нYн	1.sg -мYн	Interrogative бY
			Indicative Present Continuous - UU∂А	Com - лUU			Poss2.s g -(Y)н		Acc - ∂Y	Acc - ∂Y	Acc - н	2.sg -сYн	Desiderative чY
			Indicative Past Definite - ∂Y	Car - сYз			Poss2.s g resp - (Y)нYз		Dat -зА	Dat -А	Dat -нА	2.sg resp. -сYз	Suppositional дYр
			Indicative Past Indefinite - зАн Indicative Past Indefinite - зА				Poss3 (с)Y		Abl -∂Ан	Abl -∂Ан	Abl -нАн	1.pl -бYз	миш
			Indicative Past Continuous - чU Indicative Past Continuous - UUчU				Poss1.pl - (Y)бYз		Loc - ∂А	Loc - ∂А	Loc - н∂А	2.pl - сYн. Ap	
			Indicative Past				Poss2pl (Y)уан		Simil - ∂Ай	Simil - ∂Ай	Simil - у∂Ай	2.pl resp. сYз ∂Ан	

Тувинская словоформа

0	1	2	3	4	5	6	7		8	9								
							Distr	Perf			Neg	Tense (Praes, Past,Future) Mood	Num(Pl)	Poss	Case		Person (1, 2)	Ptel
															Simple declension	Possessive declension		
	Distr <i>КЫЛА</i>	Perf (Ы)БЫМ	Neg <i>БА</i>	Praes <i>ыр</i>	Pl <i>ЛАр</i>	1 pos.sg (Ы)М	Gen <i>НЫң</i>	Gen <i>НЫң</i>	1sg.М	Interr (Ы)л								
			Neg.Conv <i>БАйн</i>	Past <i>ТЫ</i>		2 pos.sg (Ы)ң	Dat <i>ГА</i>	Dat <i>нГА</i>	2sg.ң	Add - <i>даа</i>								
			Neg.Fut <i>БАС</i>	PastRes <i>ГАН</i>		3 pos. (з)Ы	Acc <i>НЫ</i>	Acc <i>Н</i>	1pl <i>БЫс</i>	Indir - <i>Тыр</i>								
				PastRtr <i>ЧЫК</i>		1 pos.pl (Ы)БЫс	Loc <i>ТА</i>	Loc <i>нТА</i>	2pl <i>ңАр</i>	Emph - <i>Ла</i>								
				Cont <i>БЫшААН</i>		2 pos.pl (Ы)ңАр	Abl <i>ТАн</i>	Abl <i>нТАн</i>	3pl <i>ЛАр</i>									
				Fut <i>Ыр, Ар</i>		Gen.3Pos <i>НЫЫ</i>	Lat1 <i>Че</i>	Lat1 <i>нЧе</i>	Praes.3sg.у									
				Cunc <i>ГАЛАК</i>			Lat2 <i>ТЫВА</i>	Lat2 <i>нТЫВА</i>	Imp.1sg <i>БЫн, Айн</i>									
				ConvLim <i>ГАЛА</i>			Lat3 <i>КЫДЫ</i>	Lat <i>КЫДЫ</i>	Imp.3sg <i>СЫн</i>									
				ConvPast (Ы)н					Imp.1.Dual <i>ААл(Ы), БЫЛ(Ы)</i>									
				ConvPraes <i>А, Ы, ү</i>					Imp.1.pl <i>ААЫңАр, БЫЛЫңАр,</i>									
				Prosp <i>ГАш</i>					Imp.2.pl (Ы)ңАр									
				ConvLim2 <i>ГЫже</i>					Imp.3.pl <i>СЫн. (нАр)</i>									
				Abstr <i>КЫ</i>					1.sg.Lim <i>м.че</i>									
				Opt. <i>ГАй</i>					2.sg.Lim <i>ң.че</i>									
									1.pl.Lim <i>вис.че</i>									
									2.pl.Lim <i>ңер.же</i>									
									1.sg.Cond <i>СЫ.м.зА</i>									
									2.sg.Cond <i>СЫ.ң.зА</i>									
									3sg.pl Cond <i>зА</i>									
									1.pl.Cond <i>СЫ.вЫс.сА</i>									
									2.pl.Cond <i>СЫ.ңАр.зА</i>									

• Хакасская словоформа <http://khakas.altaica.ru/>

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Distr	NF	PtcI1	Perf	Prosp/Dur1	Dur	Neg/Iter	Tense/Mood/Conv2	Transp	PI1	Poss1	Case1	Attr	PI2	Poss2	Case2	Adv	Affirm	Person	PredPI	PtcI3
Distr (K)нА	NF (bl)п	Add TAA	Perf (bl)б.бlc	Dur1 и	Dur чАт Dur.dial чат	Neg ПА iter	причастия	Comit ЛЫГ	Pt: ЛАр (bl)м	1pos.sg	Gen: НЫң	Attr Кы	Pz: ЛАр	1pos.sg (bl)м	обычные	Assz OK	Affirm ЧЫК	полные	PredPI ЛАр	Assz OK
	NF 0	Cont: LA	Perf0 (bl)с	Dur1 ир		A.dblr	Past ГА	Adv Ли		2pos.sg (bl)ң	Loc: ТА			2pos.sg (bl)ң	Gen НЫң	Contz LA		1.sg ПЫн		Фос ТЫр Прес ТАК
	NF.Neg Пин NF.Neg.sh ПААн	Assz OK		Prosp АК		iter A.dbl	Past ГАН			3pos (з)bl 1pos.pl (bl)ббlc 2pos.pl (bl)ңАр	All: САр			3pos (з)bl 1pos.pl (bl)ббlc 2pos.pl (bl)ңАр	Instr нАң			2.sg СЫң		Assz.dial ox
	NF.Neg.sh ПАан			Dur1.кас Ат			Fut А				Loc: (н)ТА				Dat ГА			1.pl ПЫс		
				Dur1.sag ит			Fut Ар				All: (н)САр				Acc НЫ			2.pl САр		
							Neg.Fut ПА.с								Gen.3pos Ни	Loc ТА		смешанные		
							Наб ЧА								Gen.3pos.dial Ди	Abl ДАң		1.sg.mix м		
							Наб ЧАң									All САр		2.sg.mix СЫң		
							Assum ГАдАГ									ProI ЧА		1.pl.mix ПЫс		
							Cunc ГАлАК									Delib ДАңАр		2.pl.mix САр		
							Simul ААчыК									Comp ТАГ		краткие		
							финитные									Abl1 тын		1.sg.br м		
							Pres чА									посессивные		2.sg.br ң		
							Pres1 чАдыр									Dat (н)А		1.pl.br ПЫс		
							Pres1 чАды									Acc НЫ		2.pl.br (bl)ңАр		
							Cond СА									Acc Н		императивные		
							RPast ТЫ									Loc (н)ТА		mp.1.sg им		
							Opt ГАй									Abl нАң		mp.1.pl и.бЫс		
							Indir тыр									All (н)САр		mp.Incl Аң		
							Indir Тыр									ProI (н)ЧА		mp.Incl.pl АңАр		
							деепричастия									Delib нАңАр		mp.2.pl (bl)ңАр		
							Convп (bl)п									(н)ТАГ		mp.3 СЫн		
							Neg.Conv Пин									Gen.dial ДЫң		Диалектные		
							Neg.Conv.Abl Пин.Аң									Acc.dial ДЫ		mp.Incl.pl.dial АлАр		
							ConvA А									Instr.dial ПАң		1.pl.dial СЫбЫс		
							Lim ГАли									Instr.dial мАң		2.pl.dial СЫң.Ар		
							Диалектные									Instr.dial ПлАң		1.sg.dial СЫм		
							PresPt чАн									Instr.dial Пл/		mp.Incl.dial АК		
							Convпас А.бАс									Instr.dial ПЫ/		mp.1.pl.dial АлЫс		
							Pres.kyz тур									Instr.dial мн/				
							Pres.kyz ту									All.dial САрЫ				
							Pres.sh чАр									All.dial СА				
																Delib.dial нАңАрЫ				
																посессивные				
																All.dial (н)САрЫ				

Page 1

Page 3

Page 5

Page 7

- Именные и глагольные словоформы:
- 1. Позиции 1–8 могут заполняться только у слов с пометой *Verbum*.
- 2. Заполнение позиций 9-17 (падежно-посессивный блок) может происходить: 1) либо сразу после позиции 0, если слово имеет помету *Nomen*, 2) либо непосредственно после позиции 7, заполненной одним из причастных показателей, если слово имеет помету *Verbum*
- *хайди тогынып үгрен-гле-п-чет-кен-нер-ін чоохтып пиреңер*
(учиться-Distr-Conv-Dur-Past-P1-Poss3.Acc)
Расскажите, как они работают, учатся?

Күн таг назына түс-чет-кен (спускаться-Dur-Past)

Солнце спускалось на вершину горы.

- **Двойной набор именных слотов**

алт-ым-зар-хы (верх-1pos-All-Attr) 'тот, что сверху от меня'

20 чыл-лар-да-гы (год-Pl-Loc-Attr) *халыхчылар* 'просветители 20-х годов'

Ол өрiн парган осхас уучаның иб-де-зи-не. (дом-Loc-3pos-Dat)
'Она, видно, обрадовалась тому, что бабушка дома.'

чүс=наз-ы-лар-ы-ның (сто=голова-3pos-Pl-3pos-Gen) *ікізін*
позынзар хыгыртыбызын 'призвав двоих из сотников'

паба-м-ни-наң (отец-1Sg.pos-Gen.3pos-Abl) - 'из принадлежащего моему отцу'

стол-зар-тын (стол-All-Abl) 'со стороны стола'

- Грамматикализованные и синтезированные в разное время сериальные конструкции
- *Піс парчатхан чурт иледе аал ортызынзар тур-чат-тыр* (стоять-Dur-Indir) 'Жилище, в которое мы направлялись, было расположено посреди обширного улуса'
- *Пір ізіг күнде ол ибінде тынанып чат-чат-хан* (лежать-Dur-Past) 'Однажды в жаркий день он, отдыхая в своем доме, лежал'
- *Суз амыр, эрінңек ах-чат-ча* (течь-Dur-Praes) 'Вода течет спокойно, лениво'
- *Түгее саңнары астал кил-ир* (приходить-Dur) 'Доносился звон колокольчиков, подвязанных к дуге' = Тур. *gel-i-yor* 'приходит', **jor*- 'двигаться', хак. *чораах* 'бродячий'

• Неполная синтезация

- *Анаң нинче-де минута назынаң оларның табыстары исти=л-бин-ібіс-кен-нер* (слышать=Pass-Conv.Neg-Perf-Past-PredPl)
'Потом, через несколько минут, голоса их стихли';
- *тооз-ып-таа-быс-ты-лар* (кончить-Conv_{II}-Add-Perf-RPast-PredPl) 'и почти закончили'
- но

Пасха соорлар аралары ачыхти тохта-гла-б-ог-ыс-хан-нар
(останавливаться-Distr-Conv_{II}-Ass-Perf₀-Past-PredPl)

'И другие сани также остановились на некотором расстоянии друг от друга'

(*-Хb *id-, Conv + пускать, хак. *ыс*- 'посылать, отправлять')

- "Сдвиг" позиции отрицания

- *тогыстаң сала-сала ла сыгардыр-быс-па-ды-лар*

(увольнять-Perf-Neg-Rpast-P1)

Их ведь чуть не уволили с работы

- *наңмыр айыл-бин-ыбыс-ты*

(прекращаться-Conv.Neg-Perf-Rpast)

Дождь не прекращается.

- *пол=ын-мин-ыбыс-ты-м*

(мочь=Refl-Conv.Neg-Perf-Rpast-1Sg)

Мне неможется (=Я не в состоянии)

- **Спасибо за внимание!**