

## Förkortningar

Programrelaterade förkortningar:

K	konsonant
Dat	dativ (lokativ)
futpart	futurum particip
POSS	possessivpronomen eller possessiv enklitisk partikel
LF	lexikalisk(a) form(er)
LN	lexikalisk nivå
V	vokal
v	reducerad vokal
YF	ytform(er)
YN	ytnivå
#	ordgräns
+	morfemgräns

Allmänna förkortningar:

ka.	kalmuckiska
mo.	mongoliska
oir.	oiratiska (klassisk)

## Om transkriberingen

I uppsatsen används de gängse markörerna för morfem ({}), grafem (<>), och fonem (/). För läsbarhets skull används för morfemen en normaliserad transkribering, vilken alltså motsvarar den transkribering som i uppsatsen används för PC-Kimmo. Alla transkriberingar återfinns i (5.3).

## 0. Inledning

I föreliggande uppsats undersöks ett urval morfonotaktiskt betingade företeelser i kalmuckisk ordstruktur, vilka analyseras och beskrivs utifrån tvånivåmorfolo­gin för programmet PC-Kimmo. Uppsatsen kan i viss mån ses som en pilotstudie för fortsatt och mer fullständig utveckling av en tvånivåbeskrivning av kalmuckisk morfologi, emedan någon sådan hittills över huvud inte har gjorts. För denna studie har valts ett antal morfonotaktiska "problem", vilka analyserats och erhållit en tvånivåbeskrivning, anpassad till PC-Kimmoomgivning. Spännvidden på materialet blir som synes ganska bred, emedan det så att säga rör sig såväl inom morfologin, som inom fonologin och ortografin. Det fonologiska utgör dock endast bakgrunden, mot vilken de övriga fenomenen uppträder, och det huvudsakliga i uppsatsen rör sig kring morfotaxen och dess grafiska uttryck (programmet skall ju trots allt analysera just skriven text).

För studien har av författaren framtagits en ordlista på ca 200 ord, vilka valts ut för att kunna skapa representativa exempel för de företeelser som tas upp till analys. I uppsatsen presenteras en tvånivåbeskrivning av följande företeelser:

- växling VV/V på ytnivå (3.1)
- vokalharmoni (3.2)
- allomorfer {dv/tv}, {žv/čv}, {hin/in/na} (3.3)
- dubbla stammar (+n och -n) (3.4)
- inskottsljud /h/ mellan vokaler vid morfemgräns (3.5)

Beskrivningen är utförd inom ramarna för programmet PC-Kimmo, vilket förutsätter följande moment: morfotaktisk beskrivning, en uppsättning morfonologiska regler samt ordlista med underlexikon. Beskrivningen utgår snarast från den ortografiska representationen, även om naturligtvis många av resonemangen grundas på fonologisk mark.

Pga. dess begränsning erbjuder föreliggande uppsats ingen färdig grund för användning av den kalmuckiska beskrivningen i praktiska sammanhang, men kan med viss bearbetning och vidareutveckling utnyttjas för morfologisk analys i diverse språkprogram.

I (0.1) och (0.2) beskrivs i grova drag det kalmuckiska språket, samt ges en kortare introduktion till programmet PC-Kimmo och tvånivåmorfolo­gin.

I (1-2) karaktäriseras och analyseras det faktiska arbetsmaterialet. Detta ligger sedan till grund för (3) där själva den tvånivåmorfolo­giska bearbetningen återges och regler presenteras.

Resultaten diskuteras vidare i (4), och de framtagna reglerna m.m. återfinns som bilagor i (5).

## 0.1 Kalmuckiska språket

Kalmuckiska språket hör till den mongoliska språkgruppen, vilken som regel kopplas samman med de turkiska och manchuriska språkgrupperna i en altaisk språkfamilj. Oavsett eventuella faktiska genetiska kopplingar har språkgrupperna ifråga tämligen mycket typologiskt gemensamt, vilket även gäller för de uraliska språken, samt i viss mån japanska och koreanska, och i de mest extrema altaiska teorierna bildar alla dessa grupper tillsammans en stor ural-altaisk språkätt.

Typologiskt är kalmuckiskan ett agglutinerande språk, där den huvudsakliga ordböjningen och -bildningen sker genom tillfogning av ändelser till en mer eller mindre oföränderlig ordstam. Ordföljden är i de allra flesta fall SOV. [3, 4, 7, 12, 14]

Fram till början av 1920-talet använde kalmuckerna en variant av den traditionella ujugurisk-mongoliska skriften. Denna variant kallas ”Den klara skriften” (*Todo bičiq*<sup>1</sup>; [2, 3, 4, 7, 16, 18]), och används i dag av den oiratiska befolkningen i den kinesiska provinsen Xinjiang. Den skrivs lodrätt från vänster till höger. Under första hälften av 1900-talet experimenterade man med ett antal skrifter, utgående än från kyrilliska, än från latinska alfabetet [9, 12]. Den segrande ortografien bygger på kyrilliska alfabetet, med tillägg av sex specifika bokstäver (se vidare 2, samt 5.3).

## 0.2 PC-Kimmo och tvånivåmorfologi

Den s.k. tvånivåbeskrivningen av morfologi och fonologi kan sägas ha sina teoretiska rötter i Chomskys generativa grammatik [5]. Skillnaden mellan en generativ beskrivning och en tvånivåbeskrivning belyses på följande sätt i manualmaterialet till programmet<sup>2</sup> ((R1) är en generativ regel, och (R2) motsvarande tvånivåregel):

”(R1)  $t \rightarrow c / \_ i$   
(R2)  $t:c \Rightarrow \_ i$

Generative rules have three main characteristics. First, they are transformational rules – they transform or rewrite one symbol into another symbol. Rule R1 states that  $t$  becomes (is changed into)  $c$  when it precedes  $i$ . After rule R1 rewrites  $t$  as  $c$ ,  $t$  no longer "exists." Second, sequentially applied generative rules convert underlying forms to surface forms via any number of intermediate levels of representation; that is, the application of each rule results in the creation of a new intermediate level of representation. Third, generative rules are unidirectional – they can only convert underlying form to surface form, not vice versa.

In contrast, two-level rules are declarative; they state that certain correspondences hold between a lexical (that is, underlying) form and its surface form. Rule R2 states that lexical  $t$  corresponds to surface  $c$  before  $i$ ; it is not changed into  $c$ , and it still exists after the rule is applied. Because two-level rules express a correspondence rather than rewrite symbols, they apply in parallel rather than sequentially. Thus no intermediate levels of representation are created as artifacts of a rewriting process. Only the lexical and surface levels

<sup>1</sup> Detta är transkriberat från skriften ifråga enligt gängse internationella normer. Transkriberat enligt nukalmuckiskans ortografiska regler skrivs det såhär: *tod bičg*.

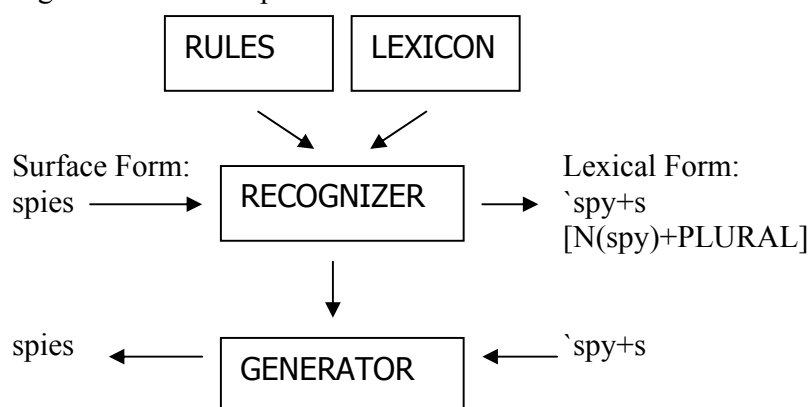
<sup>2</sup> Det citerade stycket är hämtat ur A.1.1. i manualfilen (RULES.PDF). Filen ifråga återfinns i en nedladdningsbar arkivfil (guide.zip) på SIL's webbplats ([http://www.sil.org/pckimmo/v2/pc-kimmo\\_v2.html](http://www.sil.org/pckimmo/v2/pc-kimmo_v2.html)), varifrån även övriga PC-Kimmorelaterade program kan hämtas) eller som html-dokument på ([http://www.sil.org/pckimmo/v2/doc/Rules\\_1.html](http://www.sil.org/pckimmo/v2/doc/Rules_1.html)).

are allowed. It is this aspect of their nature that is emphasized by the name "two-level" rules. Furthermore, because the two-level model is defined as a set of correspondences between lexical and surface representation, two-level rules are bidirectional. Given a lexical form, PC-KIMMO will return the surface form. Given a surface form, PC-KIMMO will return the lexical form."

PC-Kimmo är skrivet för att generera och/eller känna igen (parsa) ord genom denna tvånivåmodell, där orden alltså representeras som en motsvarighet mellan sina lexikala former och ytformerna. Arbetet med PC-Kimmo har pågått sedan 1985<sup>1</sup>, och det togs fram i syfte att möjliggöra test och felsökning av lingvistiska beskrivningar i PC-miljö. PC-Kimmo fungerar som ett skalprogram<sup>2</sup>, till vilket användaren skapar:

- a) en regelfil, som definierar teckenuppsättning och fonologiska eller ortografiska regler;
  - b) en lexikonfil, som innehåller alla stammar och suffix med morfologisk och lexikal information, samt bestämmer det morfotaktiska mönster utifrån vilket programmet kommer att verka.
- Dessa behandlas vidare av programmets "generator" och "analyserare" (generator, recognizer) enligt nedanstående modell:

Figure 1: Main components of PC-KIMMO<sup>3</sup>



Förhållandet mellan ett ords LF och YF behandlas av tvånivåreglerna genom segmentpar. Segmentparet i exempelregeln (R2) ovan är t.ex. t:c, där t är den lexikala symbolen och c ytsymbolen. Varje motsvarighet, varje möjligt segmentpar måste fastställas i beskrivningen, dvs. om en beskrivning behöver redogöra för att LF x har YF y och 0 (noll, se nedan), måste detta först fastställas genom segmentparen x:y *och* x:0.

En tvånivåregel består av tre delar: segmentparet (correspondence, feasible pair), operatorm (operator) och kontexten (environment, context). I (R2) är segmentparet t:c. Operatorm är =>, vilket innebär att den i segmentparet angivna motsvarigheten gäller *endast* i efterföljande kontext. Det finns fyra möjliga operatorer:

=> endast (men inte alltid)

<sup>1</sup> I föreliggande uppsats har använts version 2.1.0. av PC-Kimmo.

<sup>2</sup> en "kommandotolk", vars uppgifter består i att aktivera generatorm eller analyseraren och att manipulera de bakgrundsdata som dessa moduler behöver.

<sup>3</sup> Bild återgiven från 1.3 i INTRO.PDF (se fotnot 2, s. 3). Något modifierad av författaren.

<= alltid (men inte endast)

<=> alltid och endast

/<= aldrig

Kontextdelen (där streck under raden visar aktuell position) definierar i vilken omgivning motsvarigheten förekommer.

Emedan tvånivåreglerna kräver att varje lexikal symbol skall motsvaras av en ytsymbol använder man sig ibland av en nollsymbol (0), vilken också måste omnämnas som del av segmentpar i varje specifikt fall (om man t.ex. skulle beskriva den svenska YF cyklar, skulle man förmodligen ha ett segmentpar e:0 för att kunna skriva en strykningsregel för e). Den lexikala morfemgränsen (+) motsvaras av 0 på ytnivå. Ordgräns (#) markeras endast i kontextdelen.

För att förenkla och förkorta regelskrivandet använder man sig i stor utsträckning av s.k. segmentklasser (subsets), där man samlar symboler med en eller flera regler som gemensam nämnare. Exempel på detta är de i föreliggande uppsats använda segmentklasserna Vfrnt och Vbck, vilka sammanfattar alla främre respektive bakre vokaler och används i regler för vokalharmonin.

En liknande företeelse är de s.k. arkifonemen, vilka dock listas som enskilda symboler i alfabetet och inte som segmentklasser. Arkifonem används för beskrivning av t.ex. allomorfi. I denna uppsats finns t.ex. arkifonemet A, som står för det vokalharmoniska paret a-ä, och används i de lexikaliska representationerna av ändelser med varianter som växlar just enligt vokalharmonin.

För att PC-Kimmo skall kunna behandla reglerna måste dessa först ”översättas” till s.k. ”finite state tables”. Detta moment är mycket tidskrävande (i synnerhet för den dator tekniske lekmannen) att göra ”för hand”, men det finns ett hjälpprogram, Kgen<sup>1</sup>, vilket genererar (mer eller mindre) korrekta regelfiler (.rul) för PC-Kimmo. Jfr t.ex. regeln för h-inskott (se 3.5):

*Ursprunglig tvånivåregel (= indata för Kgen):*

h:0 <=> K:@ +:0 \_\_ @

*Med Kgen genererad finite state table av ovanstående regel (se 5.1.1) (= indata för PC-Kimmo):*

```
RULE " h:0 <=> K:@ +:0 __ @" 5 6
```

	h	h	K	+	@	@
	0	@	@	0	@	@
1:	0	2	2	1	1	1
2:	0	2	2	3	1	1
3:	5	4	2	1	1	1
4:	0	0	0	0	0	1
5:	0	2	2	3	1	0

Såsom nämndes ovan (0.1) så är kalmuckiska språket försett med en relativt rik morfologi med agglutinerings som främsta ”verktyg”, och utgör härav ett lämpligt studieobjekt för tvånivåmorfologin, som ju ursprungligen tillkom med tanke på morfologirika språk som finska, turkiska o.a.

<sup>1</sup> Kgen (skrivet av Nathan Miles, Ohio State University), v. 0.2. Går att hämta från ovan (se fotnot 2, s. 3) angivna webbplats.

## 1. Den fonologiska bakgrunden

Kalmuckiskan har följande uppsättning fonem (se 5.3):

vokaler: /a e y i o u ä ö/

konsonanter: /p b t d t' d' k g h f w s š z x c č ž m n n' ŋ l l' r j/

Alla vokaler har långa allofoner. Kvantitet är dock (enligt de flesta teorier) distinktiv endast i ordets första stavelse – i övriga positioner har det distinktiva särdraget blivit kvalitativt, och man skiljer i dessa positioner på reducerade/oklara och ickereducerade vokaler, vilka som regel alltså icke skiljer sig i kvantitativt hänseende. De reducerade allofonerna är till följd av sin artikulatoriska otydlighet ganska svårbestämbara, men kan i grova drag härledas till tre huvudallofoner (varav två fördelar sig på vokalharmonins främre och bakre grupper, medan den tredje är ett reducerat /i/).

Den uvulära klusilen /h/ realiseras intervokalt som uvulär frikativa. /i/ har i vissa speciella positioner en något mer central allofon, dock inte identisk med t.ex. ryskans <Ы>.

Stavelsestrukturen i kalmuckiskan är relativt enkel. I ordbörjan tillåts inte mer än en konsonant, medan det inuti ord tillåts två. För ordinledande stavelse gäller alltså:

V	/ä/ ljud
VC	/ig/ garnnystan
VCC	/uls/ folk
CV	/sö/ natt
CVC	/sän/ god
CVCC	/gils/ blänk

och för ickeinledande stavelser i tillägg:

CCVC	/mal.tmvl/ mineral
------	--------------------

### 1.1 Vokallängd

”Inte bara korta vokaler i andra än första stavelsen har utsatts för reduktion, utan i en rad andra fall gäller detta även långa vokaler och diftonger... I nukalmuckiskan har långa vokaler bevarat sin fonematiska längd praktiskt taget enbart i första stavelsen av flerstaviga ord, ibland i mitten av ord. I alla övriga fall har de långa vokalerna reducerats både kvantitativt och kvalitativt, och uppträder numera som nyanser av motsvarande korta vokalfonem” [3:33-4]. Detta innebär att den kvantitativa distinktionen är förpassad inte bara till första stavelsen, utan där även till flerstaviga ord.

Reduceringen kan exemplifieras på följande sätt (med former ur klassisk oiratiska först, varpå följer ett (hypotetiskt) mellanled, och sist den moderna formen):

*sayin* → *sään* → *sän*, men *sayin-iyēr* → *säänäär* → *säänär*  
*öür* → *üür* → *ür*, men *öüriyin* → *üüriin* → *üürin*

*niyul* → *nüül* → *nül*, men *niyul-tai* → *nüültää* → *nüültä*

På så vis får man alltså ordpar som *när* 'fest' - *näärin* [GEN], *när* 'samtycke' - *närin* [GEN].

## 1.2 Vokalharmoni

Kalmuckiskan har palatal vokalharmoni, vilken delar upp vokalismen i följande segmentklasser:

främre - /e, ä, y, ö/

bakre - /ī, a, u, o/

neutrala - /i (ä)/

Det bör påpekas att paret /ī/-i/ inte är något fullödigt par i vokalharmoniskt hänseende - /i/ förekommer visserligen bara i ord med bakre vokaler, men endast i suffix, medan /ī/ är helt neutral och förekommer i alla positioner.

/ä/-fonemets status i kalmuckiskans system för vokalharmoni vacklar. I äldre talspråk (och förmodligen i en stor del av kalmuckiskans språkvarianter) är det neutralt, men har i modernt standardspråk och i gällande ortografi förpassats till den främre gruppen. Dess neutrala karaktär ser man i bruket av en mängd ändelser, vilka bevarar /ä/ oavsett vilken grupp stammen tillhör, t.ex. den sammansatta presensändelsen *-čänä/-žänä* (vilken av en del beskrivs som en självständig presensändelse; den består av gerundieändelsen *-č/-ž-* och *-änä-* (reducerat hjälpverb *bää-* med den "enkla" presensändelsen *-nä*)); komitativändelsen *-tä*; preteritumändelserna *-lä* och *-wä* m.fl. Dessa återgår på *-či/-ži bayinai*; *-tai/-tei*; *-lā/-lē* resp. *-bai/-bei*, dvs. varianten med /ä/ ter sig som en utjämnande nyskapelse. I standardspråk realiseras ovannämnda ändelser som respektive *<-čänä/-čana, -žänä/-žana; -tä/-ta; -lā/-la; -wä/-wa>*.

Den beskrivning som presenteras i denna uppsats följer den moderna ortografin och delar in alla berörda affix enligt vokalharmonins regler.

## 2. Den ortografiska bakgrunden.

Såsom ovan nämndes använde kalmuckerna fram till 1930-talet en variant av den traditionella ujugurisk-mongoliska skriften, kallad *todo bičiq*. Under första hälften av 1900-talet experimenterade man med ett antal skrifter, utgående än från kyrilliska, än från latinska alfabetet. Nukalmuckiskans skrift utgår från kyrilliska alfabetet, med tillägg av sex specifika bokstäver (se 5.3). Den ortografiska grunden för skriften har vissa egenartade drag, av vilka det mest iögonfallande är att man inte sätter ut de reducerade vokalerna i skrift. Härav följer att det i vissa ord kan uppstå anhopningar på upp till 10 konsonanter (förmodligen går det att komma på ord med bortåt 15, även om exempel på dylika ”skriftliga tungvrickare” blir ganska krystade):

<ümsldxšwdnl> /ymseldexšywdenel/<sup>1</sup>.

**ümsld x š wdn l**  
pussas [fut.part.] [NEG] [PRO.1p.pl.] ju

Som regel blir det dock inte mer än 4-5 konsonanter, och även om det försvårar läsningen något i början, så gör det nog inte kalmuckiskan svårare att läsa än t.ex. engelska. Under en längre tid har det titt som tätt framkommit förslag till skriftreformer, av vilka alla har gått ut på att införa någon form av beteckningar för de reducerade vokalerna. I slutet av 90-talet genomfördes en reform, som gav upphov till en begränsad mängd tryckmaterial (däribland läroböcker och ordlistor), men som dock ganska snart avbröts, varefter man återgick till den gamla, ”vokallösa” skriften.

En annan följd av de reducerade vokalernas frånvaro är att det i skrift inte går att skilja enstaviga ord från tvåstaviga ord med reducerad vokal i utljud. Ett skrivet ord <(C)VC> kan representera både /(C)VC/ och /(C)VCv/. Detta <(C)VC> kan vid utvidgning av stavelsestrukturen förvandlas till <(C)VVC->. Detta är den grafiska återspeglingsen av de i föregående kapitel omnämnda fonetiska företeelserna, dvs. att etymologiskt och lexikaliskt långa vokaler i enstaviga ord (samt i sista stavelsen av flerstaviga ord) i nukalmuckiskan numera uppträder som nyanser av motsvarande korta vokalfonem. Så länge ett ord består av en stavelse kan det inte ha dubbeltecknat vokaltecken, men så fort det genom modifiering får ytterligare en stavelse så måste vokaltecknet dubbeltecknas. Detta är en konsekvent ortografisk regel i kalmuckiskan.

Vi har t.ex. ordet <ul> ’fnöske’, som i utvidgad form sammanfaller med <uul> ’berg’:

<ul> ’fnöske’            <uuld> (dat-lok.)  
<uul> ’berg’            <uuld> (dat-lok.)

Eller ’skimmel’ och ’stäpp’:

<ker> ’skimmel’            <keerin> (gen.)  
<keer> ’stäpp’            <keerin> (gen.)

En stor grupp bildar också de enstaviga verbstammarna, vilka får ett lika regelbundet växlingsmönster som nominalstammarna:

LF xaa 'stänga'

---

<sup>1</sup> Med *todo bičiq* skulle anförda exempel skrivas så här: *ümsüldekeüšü bidan ele*.



imp.2p. <xa>, imp.3p. <xag>, fut.part. <xaax>, mom.ger. <xan> osv.

uu 'driicka'

imp.2p. <u>, imp.3p. <ug>, fut.part. <uux>, mom.ger. <un> osv.

ää 'frukta'

imp.2p. <ä>, imp.3p. <äg>, fut.part. <ääx>, mom.ger. <än> osv.

### 3. Behandling av ämnet inom ramarna för programmet PC-Kimmo

Till analys uppställdes följande företeelser:

- växling VV/V på ytnivå (3.1)
- vokalharmoni (3.2)
- allomorfer {dv/tv}, {žv/čv}, {hin/in/na/nä} (3.3)
- dubbla stammar (+n och -n) (3.4)
- inskottsljud /h/ mellan vokaler vid morfemgräns (3.5)

#### 3.1 Växling VV/V på ytnivå

Såsom förklarades i (1) återgår denna växling på följande: lång vokal är möjlig endast i första stavelsen av flerstaviga ord. Detta ger till följd att alla enstaviga ord med etymologiskt lång vokal i nukalmuckiskan uppträder med kort vokal, men att dessa vid utvidgning av stavelsestrukturen (böjning) uppvisar "återupprättad" lång vokal.

Frågan var vilka LF man skulle tillskriva en- och tvåstaviga ord med lexikaliskt lång vokal (som t.ex. <ker, ul> (med kasusformer som <keerig, uulta>) och <keer, uul>)? Den tämligen logiska lösningen torde vara att gå tillbaka till den fonotaktiska verkligheten som ligger bakom de ortografiska reglerna. Som visades ovan (1.2, 1.3), så har de ortografiska reglerna uppkommit som återspeglning av den kvantitativa reducering som långa vokaler i sista stavelsen (innefattande enstaviga ord, såsom utgörande samtidigt första och sista stavelse) genomgått i kalmuckiskan.

Genom en kortare analys kommer man så fram till slutsatsen att ett av de fonotaktiska särdragen mellan ovannämnda en- och tvåstavningar är den reducerade vokalen (dess från- och närvaro). Den lösning som erbjuds i föreliggande uppsats är att på LN införa en (i förhållande till reellt uttal något abstrakt) reducerad vokal, som här representeras med tecknet *v*. På så sätt får vi följande LF:

keer [skimmel]  
keerv [stäpp]

Regeln som skrevs för att beskriva detta ser ut på följande vis:

Vlng:Vshrt <=> (K:@) \_\_ [ (K)# | (K) +:0 K# ]

Segmentparet består av två segmentklasser: lexikalisk Vlng med motsvarighet Vshrt, dvs. "lång vokal motsvarar kort vokal". Operatorm fastställer att omnämnda förhållande gäller alltid och endast i efterföljande kontext. Kontexten är i det här fallet trefaldig: en föregående del där det kan stå vilken konsonant som helst, och en sammansatt efterföljande del, bestående av två med lodrätt streck skilda uttryck. På så vis undviker man två varandra motsägende regler (med operatorerna alltid och endast). De två efterföljande delarna möjliggör förekomsten av konsonant före ordslut, eller eventuell konsonant ihop med ännu en konsonant efter morfemgräns före ordslut. Här bör påpekas att den reducerade vokalen <*v*> är placerad bland konsonanterna i segmentklass K, och alltså innefattas av det avslutande K.

*Exempel:* LF keer [skimmel] = YF ker; LF ää [ljud] = YF ä; LF keer + m [poss] = YF kerm; LF ää + n' [poss] = YF än'; LF däa [krig] + d [pl] = YF däd; LF däa + d + Dv [dat] = YF dääd; LF aawv [farfar/morfar], aahv [kopp] = YF aaw, aah.

### 3.2 Vokalharmoni

Kalmuckiskan har palatal vokalharmoni, vilken delar upp vokalismen i följande segmentklasser:

främre - /e, ä, y, ö/

bakre - /ĩ, a, u, o/

neutrala - /i/

Som nämndes ovan (1.2) är paret /i/-/i/ inte något fullödigt par i vokalharmoniskt hänseende – /i/ förekommer visserligen bara i ord med bakre vokaler, men endast i suffix, medan /i/ är helt neutral och förekommer i alla positioner.

Emedan användandet av /i/ därtill är något oregelbundet (och tämligen marginellt) förs de berörda orden och suffixen till speciella underlexikon (NOUN3, CN3). För suffixens del så gäller ju dessutom att det bara rör sig om varianter med bakre vokaler, varför ingen speciell regel för i-variationen behövs.

Vokalharmonin beskrivs här på följande sätt:

A:a <=> Vbck @\* + (@) \_\_

A:E <=> Vfrnt @\* + (@) \_\_

U:u <=> Vbck @\* + (@) \_\_

U:y <=> Vfrnt @\* + (@) \_\_

Här används två arkifonem A och U för att beteckna respektive varianter. Dessa definieras med operatoren ”alltid och endast”. Samma resultat hade kunnat uppnås med en ”aldrig-regel” för ett av paren (U:u <=> Vbck ..., U:y /<= Vbck ...).

*Exempel:* aawv [farfar/morfar] + hAn [poss] = aawan; aahv [kopp] + s [pl] + hAn [poss] = YF aahsan; LF keer [skimmel] + mUd [pl] + hAn [poss] = YF keermydän; sal [flotte] + ĩn = salĩn.

### 3.3 Allomorfer {dv/tv}, {žv/čv}, {hin/ĩn/na/nä}

En mindre grupp affix har positionsbetingade varianter<sup>1</sup>. I denna uppsats behandlas följande:

dativ-lokativ {dv} <d> {tv} <t>

---

<sup>1</sup> Obs. dock att övriga allomorfer inom böjningen behandlas av reglerna för vokalharmoni samt h-inskott.

preteritum {žv} <ž> {čv} <č>

De kontextuella kraven för dessa allomorfier är följande: <t> och <č> uppträder efter <r s g w>, i alla andra fall <d> och <ž>. En ”biefekt” av införandet av den reducerade vokalen (se 3.1) är att reglerna för dessa allomorfier kraftigt förenklas, emedan det annars skulle krävas en mängd kompletterande regler för ord som slutar på [”otillåten konsonant”]+[red.vok].

Programmet måste alltså veta att <bär-> ’hålla’ skall ha gerundieändelsen <ž> och inte <č>, att <saw> ’kär!’ i dativ skall vara <sawd> och inte <\*sawt>, osv. Detta blir möjligt när ordens LF är *bärv*, *sawv* osv.

Reglerna är formulerade såhär:

D:t <=> [ s | r | g | w ] +:0 \_\_

q:C <=> [ s | r | g | w ] +:0 \_\_

Detta vill säga att d- respektive ž-allomorferna förutsätts i alla andra fall, utom de här specificerade. Här används arkifonemen D och q för de bägge paren. Jfr även användandet av lodrätt streck för sammansatta kontexter med (3.1).

För varianten {in}, se ovan (3.2). För inskotts-h vid tillfogning av {hin}, se nedan (3.5). Na/nä-varianten hör till problemet med de dubbla stammarna (3.4).

*Exempel:* LF aawv [farfar/morfar], aahv [kopp] + Dv [dat] = YF aawd, aahd; aawv + nvr [pl] + Dv = YF aawnrt; LF aahv + s [pl] + Dv = YF aahst; LF sur [fråga] + qv [pret] = YF surč; LF bärv [hålla] + qv = YF bärž; LF zoog [skämt] + Dv = YF zoogt.

### 3.4 Dubbla stammar (+n och -n)

En tämligen stor grupp nomina i kalmuckiskan är utrustade med två stammar. Den ena (som i föreliggande uppsats kallas andrastam) skiljer sig åt från den andra (som följaktligen kallas förstastam) genom utljud på nasalen /n/ (i vissa fall även /ŋ/). Denna företeelse kännetecknar de flesta mongoliska språk. I xalxamongoliskan är användandet av andrastammen något mer obestämt än i kalmuckiskan, i burjätiskan ungefär detsamma som i kalmuckiskan. I burjätiskan utgör andrastammen den s.k. nominativstammen (som är identisk med nominativformen), medan förstastammen är den oblika stammen (identisk med formen för direkt objekt) [13:32 ff.], och ungefär detsamma kan man alltså säga om förhållandet i kalmuckiskan. T.ex. heter ’sång; röst’ i nominativ <du> (med lexikaliskt lång vokal), men som omarkerat direkt objekt <du>, t.ex. <du duulx> ’att sjunga en sång’, och som markerat direkt objekt <duug>.

Egenskapen att ha två stammar återfinns hos en relativt stor grupp nomina. Hos nomina med två stammar växlar dessa tämligen regelbundet i böjningen (så att t.ex. dativ och genitiv alltid kräver n-stam).

Vad gäller de tvåstammiga orden tedde det sig lämpligast att föra dessa till ett särskilt underlexikon (NOUN2), där n-utvidgningen ses som en speciell egenskap hos dessa nomina. Således har det varit nödvändigt att även skapa särskilt underlexikon med kasusändelser,

kopplade till n-stammarna. Även om det kanske verkar lite "bakvänt", ansågs det lämpligare att i samband med detta införa stam-n:et som inledande för suffixen ifråga. På så vis får man denna typ av förhållande:

(LEXICON Noun2) dää 'krig', vilket genom en fortsättningsregel kopplas till

(LEXICON Case2) nAs [abl]

Dessutom skrevs en strykingsregel för n, för att erhålla korrekta genitivformer av t.ex. perf.part.:  
n:0 <=> n +:0 (@:0) \_\_

*Exempel:* YF däänäs [krig (abl)] = LF dää + nAs; YF däänd [krig (dat)] = LF dää + ndv; YF sursnas [fråga (perf.part. abl.)] = LF sur + svn + nAs.

### 3.5 Inskottsljud /h/ mellan vokaler vid morfemgräns

Vid tillfogning av suffix med anljud på vokal till stam med utljud på vokal (förutom reducerad sådan) uppkommer ett inskotts-h mellan stam och suffix. Detta gäller följande (här behandlade) ändelser:

ablativ (h)As

possessiv 3p. (h)An

genitiv (h)in

Ex.: keer+hin → keerin, keer+hAs+hAn → keerasän

Reglerna är:

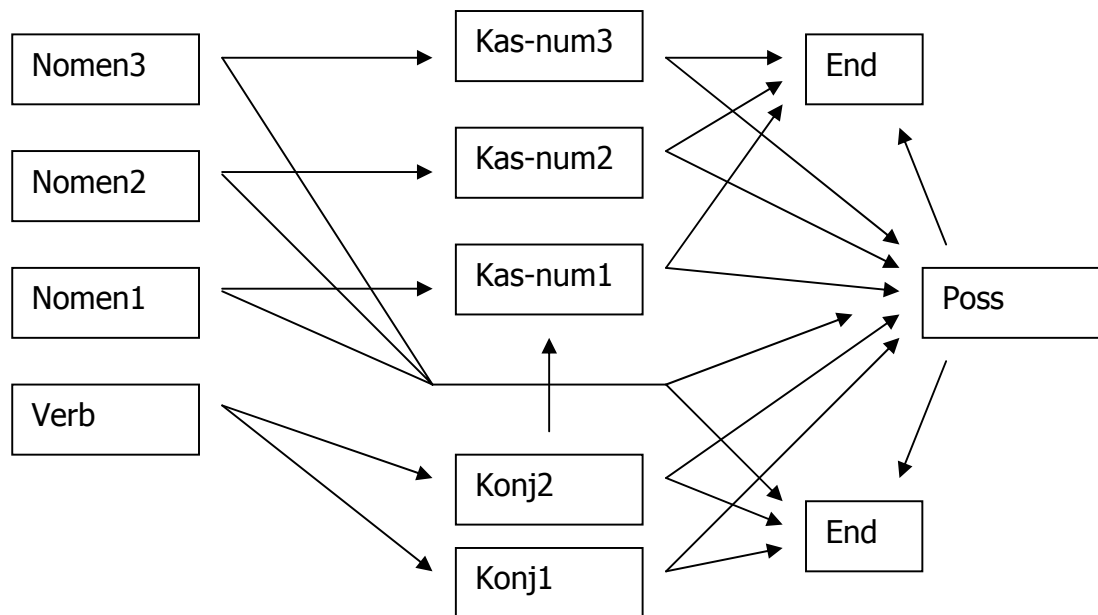
h:0 <=> K:@ +:0 \_\_ @

I denna regel fastställs att lexikaliskt h alltid och endast representeras som 0 efter konsonant och morfemgräns.

*Exempel:* LF aawv + hin = YF aawin; LF ää + hin = YF äähin; LF keer + mUd + Dv + hAn = YF keermyddän.

### 3.6 Morfotaxen

Morfotaktiska förhållanden beskrivs i PC-Kimmo genom stam- och suffixlexikon som kopplas till varandra genom s.k. fortsättningsregler (eng. *continuation rules*). I denna uppsats tillämpas följande nominala och verbala morfotax (observera att det av rent grafiska skäl fanns nödvändigt representera "End" på två ställen, emedan pilmängden annars skulle göra bilden alltför svårtydbar):



Nomen1 = nomina med endast förstastam

Nomen2 = nomina med andrastam

Nomen3 = nomina som tar i-suffix

Kas-num1 = kasus- och numerusändelser för förstastammar

Kas-num2 = kasus- och numerusändelser för andrastammar

Kas-num3 = "i-varianter" av kasus- och numerusändelser

Konj1 = verbändelser (här: imperativ, preteritum, gerundium)

Konj2 = participändelser m.m.

Poss = possessiva enklitika

#### 4. Slutledningar

Materialet för denna uppsats hämtades ur kalmuckiska språket. Målet var att skapa en tvånivåmorfologisk beskrivning av materialet för programmet PC-Kimmo. De företeelser som togs upp till behandling var följande:

- växling VV/V på ytnivå
- vokalharmoni
- allomorfer {dv/tv}, {žv/čv}, {hin/in/na/nä}
- dubbla stammar (+n och -n)
- inskottsljud /h/ mellan vokaler vid morfemgräns

I (0-2) skapas en hållbar teoretisk grund för bearbetning av materialet utifrån tvånivåmorfologin och PC-Kimmo genom analys av 1) fonotaktiska (fonologiska) bakgrunden; 2) ortografiska bakgrunden och 3) PC-Kimmo och tvånivåmorfologin som praktiskt ramverk.

I (3) beskrevs och argumenterades för lösningar på de belysta problemen:

(3.1)

Vlng:Vshrt  $\Leftrightarrow$  (K:@) \_\_ [ (K)# | (K) +:0 @#]

Segmentparet består av två segmentklasser: lexikalisk Vlng med motsvarighet Vshrt, dvs. ”lång vokal motsvarar kort vokal”. Operatorm fastställer att omnämnda förhållande gäller alltid och endast i efterföljande kontext. Kontexten är i det här fallet trefaldig: en föregående del där det kan stå vilken konsonant som helst, och en sammansatt efterföljande del, bestående av två med lodrätt streck skilda uttryck. På så vis undviker man två varandra motsägande regler (med operatorerna alltid och endast). De två efterföljande delarna möjliggör förekomsten av konsonant före ordslut, eller eventuell konsonant ihop med ännu en konsonant efter morfemgräns före ordslut. Här bör påpekas att den reducerade vokalen <v> är placerad bland konsonanterna i segmentklass K, och alltså innefattas av det avslutande K.

(3.2)

A:a  $\Leftrightarrow$  Vbck @\* + (@) \_\_

A:E  $\Leftrightarrow$  Vfnt @\* + (@) \_\_

U:u  $\Leftrightarrow$  Vbck @\* + (@) \_\_

U:y  $\Leftrightarrow$  Vfnt @\* + (@) \_\_

Här används två arkifonem A och U för att beteckna respektive varianter. Dessa definieras med operatorm ”alltid och endast”.

(3.3)

D:t  $\Leftrightarrow$  [ s | r | g | w ] +:0 (@:0) \_\_

q:C  $\Leftrightarrow$  [ s | r | g | w ] +:0 (@:0) \_\_

Detta vill säga att d- respektive ž-allomorferna förutsetts i alla andra fall, utom de här specificerade. Här används arkifonemen D och q för de bägge paren.

(3.5)

h:0  $\Leftrightarrow$  K:@ +:0 \_\_ @

I denna regel fastställs att lexikaliskt h alltid och endast representeras som 0 efter konsonant (inkl. reducerad vokal) och morfemgräns.

Vad gäller de tvåstammiga orden (3.4) tedde det sig lämpligast att föra dessa till ett särskilt underlexikon (NOUN2), där n-utvidgningen ses som en speciell egenskap hos dessa nomina. Således har det varit nödvändigt att även skapa särskilt underlexikon med kasusändelser, kopplade till n-stammarna. Detta gjordes även för ord som tar "i-ändelser". Även om det kanske verkar lite "bakvänt", ansågs det lämpligare att i samband med detta införa stam-n:et som inledande för suffixen ifråga. På så vis får man denna typ av förhållande:

(LEXICON Noun2) dää 'krig', vilket genom en fortsättningsregel kopplas till  
(LEXICON Case2) nAs [abl]

Dessutom skrevs en strykningsregel för n:

n:0 <=> n +:0 (@:0) \_\_

Detta beror på att om man för t.ex. tvåstamsnomina kan anföra en grundform utan n (vilket dock skiljer sig från bruket i ordböcker), så går inte detta för t.ex. perf.part. på <-sn>. Och eftersom stam-n:et listas ihop med kasusformerna för n-stammar i ett särskilt underlexikon så får vi alltså följande resultat om vi försöker generera perf.part.gen.: sur+svn+nA > sursna. Med n-strykningsregeln blir resultatet korrekt: sursna.

Eventuellt skulle det kunna gå att utöka systemet med underlexikon för att behandla de dubbla stammarna (stam1 > kasus1/-n- > kasus2), och på så vis komma ifrån n-strykningsregeln. För denna uppsats syfte ansågs dock ovanstående tillräckligt.

Genitivens varianter beskrevs dels genom reglerna för inskotts-h, dels genom lösningen för orden med två stammar.

Även om just föreliggande arbete inte självständigt skulle kunna ligga till grund för utnyttjande i någon form av programvara för språkhantering, kan det mycket väl i utvidgad och utökad form komma att bli analyskomponent i t.ex. stavnings- och grammatikkontroll, sökmotor e.d.

Intressant vore även att skapa en omvandlare från klassisk oiratiska till modern kalmuckiska. Detta torde gå relativt lätt, då merparten av skillnaderna är regelbundna.



## Bilagor

### 5.1 Regelfil och lexikonfil

#### 5.1.1 Regler

; KALMYK.RUL 21-March-03 (14 June 03)

; E = low front unrounded (as long &)  
; I = high central unrounded (only in affixes)  
; O = mid front rounded (as long %)  
; y = high front rounded  
; G = velar n  
; N = palatalized n  
; L = palatalized l  
; Z = voiceless alveopalatal affricate  
; C = voiceless alveopalatal affricate  
; j = voiced alveopalatal affricate  
; h = voiced uvular stop/fricative  
; S = voiceless alveolar sibilant  
; Q = long vowel a  
; W = long vowel e  
; T = long vowel i  
; H = long vowel y  
; B = long vowel o  
; M = long vowel u  
; X = palatalized t  
; R = palatalized d

; archiphonemes:  
; D = alveolar stop  
; A = low unrounded  
; ? = high unrounded  
; q = palatal affricate  
; U = high rounded

ALPHABET

p t k b d g m n G N f s S z Z c C j l L r w h D q + a E e i I o u O y A ?

U v Q & W T B M % H R X

NULL 0

ANY @

BOUNDARY #

SUBSET K p t k b d g m n N G f s S z Z c C j l L r R X w h D q v

SUBSET V a E e i I o u y O U A ? Q W T H B M & %

SUBSET Vshrt a E e i o u y O

SUBSET Vlng Q W T H B M & %

SUBSET Vbck a u o I Q B M

SUBSET Vfrnt E i e y O W T H & %

RULE "defaults" 1 31

p t k b d g m n G N f s S z Z c C j l L r w h h D D q q + a @  
p t k b d g m n G N f s S z Z c C j l L r w h 0 d t C Z 0 a @  
1: 1

RULE "defaults" 1 31

```

E e i I o u O y A A ? ? U U v Q Q & & W W T T B B M M % % H @
E e i I o u O y a E i I u y 0 a Q E & e W i T o B u M O % y @
1: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

RULE "defaults" 1 2

```

H @
H @
1: 1 1

```

RULE " Vlng:Vshrt <=> (K:@) \_\_\_ [ (K)# | (K) +:0 K# ]" 9 7

	Vlng	Vlng	K	K	#	+	@
	Vshrt	@	@	K	#	0	@
1:	6	2	1	1	1	1	1
2:	6	2	1	3	0	4	1
3:	6	2	1	1	0	4	1
4:	6	2	1	5	1	1	1
5:	6	2	1	1	0	1	1
6:	0	0	0	7	1	8	0
7:	0	0	0	0	1	8	0
8:	0	0	0	9	0	0	0
9:	0	0	0	0	1	0	0

RULE " h:0 <=> K:@ +:0 \_\_\_ @" 5 6

	h	h	K	+	@	@
	0	@	@	0	@	@
1:	0	2	2	1	1	1
2:	0	2	2	3	1	1
3:	5	4	2	1	1	1
4:	0	0	0	0	0	1
5:	0	2	2	3	1	0

RULE " D:t <=> [ s | r | g | w ] +:0 \_\_\_" 3 8

	D	D	s	+	r	g	w	@
	t	@	s	0	r	g	w	@
1:	0	1	2	1	2	2	2	1
2:	0	1	2	3	2	2	2	1
3:	1	0	2	1	2	2	2	1

RULE " q:C <=> [ s | r | g | w ] +:0 \_\_\_" 3 8

	q	q	s	+	r	g	w	@
	C	@	s	0	r	g	w	@
1:	0	1	2	1	2	2	2	1
2:	0	1	2	3	2	2	2	1
3:	1	0	2	1	2	2	2	1

RULE " A:a <=> Vbck @\* + (@) \_\_\_" 4 6

```

A A Vbck @ + @

```

	a	@	Vbck	@	+	@
1:	0	1	2	1	1	1
2:	2	2	2	2	3	1
3:	4	4	4	4	1	1
4:	1	0	2	1	1	1

RULE " A:E <=> Vfrnt @\* + (@) \_\_\_" 4 6

	A	A	Vfrnt	@	+	@
	E	@	Vfrnt	@	+	@
1:	0	1	2	1	1	1
2:	2	2	2	2	3	1
3:	4	4	4	4	4	1
4:	1	0	2	1	1	1

RULE " U:u <=> Vbck @\* + (@) \_\_\_" 4 6

	U	U	Vbck	@	+	@
	u	@	Vbck	@	+	@
1:	0	1	2	1	1	1
2:	2	2	2	2	3	1
3:	4	4	4	4	4	1
4:	1	0	2	1	1	1

RULE " U:y <=> Vfrnt @\* + (@) \_\_\_" 4 6

	U	U	Vfrnt	@	+	@
	y	@	Vfrnt	@	+	@
1:	0	1	2	1	1	1
2:	2	2	2	2	3	1
3:	4	4	4	4	4	1
4:	1	0	2	1	1	1

RULE " n:0 <=> n +:0 (@:0) \_\_\_" 4 6

	n	n	n	+	@	@
	0	@	n	0	0	@
1:	0	2	2	1	1	1
2:	0	2	2	3	1	1
3:	4	0	0	4	4	1
4:	1	0	0	1	1	1

END

## 5.1.2 Lexikon

Underlexikonen är uppbyggda av ett antal "fältkoder" (fieldcodes), vilka beskriver och bestämmer de listade formerna. Dessa är:

\lf (lexikal form), \lx (lexikon till vilket formen hör), \alt (uttryck för "fortsättningsregeln", vilken bestämmer möjliga fortsatta underlexikon), \gl (lexikalisk information).

```

;entries for          \lx NOUN3          \gl +Abl
KALMYK.LEX          \alt Noun3
                   \gl N(raft)          \lf +nA
;LEXICON INITIAL    \lx CN2
\lf 0                ;LEXICON CN1      \alt Poss
\lx INITIAL          \lf +Dv          \gl +Gen
\alt Begin           \lx CN1
\gl [                \alt Poss          ;LEXICON CN3
                   \gl +Dat          \lf +Dv
;LEXICON NOUN1      \lx CN3
\lf Qwv              \lf +hin          \alt Poss
\lx NOUN1            \lx CN1          \gl +Dat
\alt Noun1           \alt Poss
\gl N(grandfather)  \gl +Gen          \lf +as
                   \lx CN3
\lf Qhv              \lf +hAs          \alt Poss
\lx NOUN1            \lx CN1          \gl +Abl
\alt Noun1           \alt Poss
\gl N(cup)           \gl +Abl          \lf +mud
                   \lx CN3
\lf &                \lf +mUd          \alt Noun1
\lx NOUN1            \lx CN1          \gl +P11
\alt Noun1           \alt Noun1
\gl N(sound)        \gl +P11          \lf +nvr
                   \lx CN3
\lf zBg              \lf +nvr          \alt Noun1
\lx NOUN1            \lx CN1          \gl +P12
\alt Noun1           \alt Noun1
\gl N(joke)         \gl +P12          \lf +s
                   \lx CN3
\lf kWr              \lf +s           \alt Noun1
\lx NOUN1            \lx CN1          \gl +P13
\alt Noun1           \alt Noun1
\gl N(bay horse)    \gl +P13          \lf +In
                   \lx CN3
;LEXICON NOUN2      \alt Poss
\lf d&                \lf +d           \gl +Gen
\lx NOUN2            \lx CN2
\alt Noun2           \alt Noun1
\gl N(war)          \gl +P14
                   \lf +ndv
\lf dM              \lf +ndv          \lf +N
\lx NOUN2            \lx CN2          \lx POSS
\alt Noun2           \alt Poss        \alt End
\gl N(song)         \gl +Dat          \gl +Poss3
                   \lf +m
;LEXICON NOUN3      \lf +nAs          \lx POSS
\lf sal              \lx CN2          \alt End
                   \alt Poss          \gl +Poss1

```

```

\lf +hAn                ;LEXICON INITIAL                \lf +qv
\lx POSS                \lf 0                            \lx CONJ1
\alt End                \lx INITIAL                    \alt End
\gl +Poss2             \alt Begin                      \gl +pret

\lf +w                  \lx VERB                       \lf +n
\lx POSS                \lf b&                          \lx CONJ1
\alt End                \lx VERB                       \alt End
\gl +Possa             \alt Verb                       \gl +Conv

\lf +C                  \gl V(be)

\lx POSS                \lf sur                          ;LEXICON CONJ2
\alt End                \lx VERB                       \lf +svn
\gl +Possb             \alt Verb                       \lx CONJ2
\lf +wvdn              \gl V(ask)                      \alt Noun2
\lx POSS                \lf bErv                       \gl +pret2
\alt End                \lx VERB                       \lf +nA
\gl +Possc             \alt Verb                       \lx CONJ2
\lf bErv               \gl V(hold)                    \alt Poss
\lx VERB               \gl +pres

;LEXICON End
\lf 0                   ;LEXICON CONJ1
\lx End                 \lf +0
\alt #                  \lx CONJ1
\gl ]                   \alt End
                        \gl +Imp1

                        \lf +g
                        \lx CONJ1
                        \alt End
                        \gl +Imp2

;entries for
KALMYK.LEX

```

## 5.2 Ordlistor (sammanlagt 180 ord) [8, ,15, 17]

### 5.2.1 Enstaviga ord med lexikaliskt lång vokal (82 st.)

	ord	Todo bičig	LF för PC-Kimmo	betydelse
1	ä	ai	EE	ljud
2	ba	bā	baa	böter
3	bä	bai	bEE	mål, måltavla
4	bi	biy	bii	dans
5	bö	böö	bOO	shaman
6	bu	bou	buu	gevär; amulett
7	gyn	göün	gyyn	sto, märr
8	gyr	güyir	gyyr	mal (insekten)
9	hä	yai	hEE	olycka, sorg
10	ho	yō	hoo	rak, rät; rättfram, ärlig
11	hu	you	huu	melon; ram
12	dän	dayin	dEEen	krig
13	dun	doun	duun	sång; röst; ljud
14	dy	döü	dyy	småsyskon
15	zar	zār	zaar	mysk
16	zä	zai	zEE	plats
17	ze	zē	zee	dotterbarn; syskonbarn
18	zeg	zēq	zeeg	mönster, ornament
19	zi		zii	seldon
20	zo	zō	zoo	ryggslut; ås, bergrygg
21	zog	zōq	zoog	mål(tid); skämt, ironi
22	zun	zoun	zuun	hundra
23	zyn	zöün	zyyn	öster; vänster
24	zyn	zöün	zyyn	nål
25	ig	iyiq	iig	garn, garnnystan
26	jyr	jöür	jyyr	översvämning
27	ka	ka	kaa	
28	ke	kē	kee	vacker, uppklädd
29	kö	köö	kOO	sot
30	ker	kēre?	keer	skimmel
31	ki	kei	kii	luft; andning
32	kör	köör	kOOor	skrytsam
33	kyn	kümün	kyyn	människa
34	kyr	köür	kyyr	samtal
35	la	lā	laa	ljus (levande); träsk, sumpmark
36	lir	liyir	liir	akacia
37	lu	lou	luu	drake; åska
38	lus	lousa?	luus	mula, mulåsna
39	mä	mai	mEE	sill
40	mu	mou	muu	dålig
41	mis	miyis	miis	katt
42	na	nā	naa	på denna sidan
43	nä	nai	nEE	hopp, tro

44	när	nayir	nEEr	fest
45	ni	nei	nii	endräktig, samfäld
46	nör	noyir	nOOr	sömn; bukspottkörteln
47	nur	nour	nuur	sjö
48	nyl	nöül/ nügel	nyyl	synd, last
49	nyr	nour/ niyur	nyyr	ansikte
50	ö	oi	OO	skog, dunge
51	ö	öö	OO	ojämnhet; missnöje, förnärmelse
52	sän	sayin	sEEEn	god, bra
53	sö	söö	sOO	natt
54	söm	sööm	sOOm	avståndet mellan tummen och pekfingret
55	sy		syy	armhåla
56	syl	söül	syyl	svans; slut
57	tar	tāra	taar	enighet, samtycke
58	teg	tēq	teeg	stäpp; tvärbjälke, dörrklinka
59	to	tō	too	tal (antal)
60	tö	töö	tOO	avståndet mellan tummen och pekfingret
61	u		uu	vid, rymlig, öppen
62	ul	?	uul	fnöske
63	ur	our	uur	ånga; ilska, vrede
64	yr	öür	yyr	rede; vän
65	xa	xā	xaa	framben på djur (med revben)
66	xo	xō		ljus (färg)
67	xol	xōlo?	xool	hals; mat
68	xön	xoni	xOOEn	får
69	xur	xour	xuur	fiol
70	xu		xuu	alldeles, helt och hållet
71	xy	xui	xyy	virvelvind
72	ca	cā	caa	efter, hinsides; ren (djuret)
73	cä	cai	cEE	té
74	cer	cēr	ceer	förbud
75	cö	cöö	cOO	gjutjärn
76	cön	cöön	cOOEn	få, lite
77	chi		Cii	körbär
78	chig	čiyiq	Ciig	fukt, väta
79	shi		Sii	föreställning
80	shud	šoud	Suud	rätt, rättframt, rätt och slätt
81	shyr		Syyr	säll, filter
82	jun	joun/ jayun	juun	vad (inter.pron.)

### 5.2.2 Fördelning på olika typer av stavelsestrukturer (till tabell 1)

V	VC	CV	CVC (-CVn)	CVn
4	4	37	27	10

### 5.2.3 Tvåstaviga ord (utan andrastam) med lång vokal (63 st.)

Angivna ord är alla enstammiga. Detta emedan alla enstaviga tvåstammiga ord hör till (5.2.1) och återfinns där. I klassisk oiratiska slutar alla former på vokal.

	<b>ord</b>	<b>LF för PC-Kimmo</b>	<b>betydelse</b>
1	aaw	aawv	farfar/morfar
2	aah	aahv	kopp, skål
3	baar	baarv	taktegel
4	bääc	bEEcv	livsföring
5	biiw	biiwv	klarinet
6	biir	biirv	pensel
7	biič	biiCv	dansör
8	böör	bOOrv	njure
9	buur	buurv	avelstjur (kamel)
10	byyl	byylv	tandkött
11	byyr	byyrv	läger, boplat
12	däär	dEErv	skavsår (på hästrygg)
13	deer	deerv	på, ovanpå
14	deed	deedv	högsta, övre, över-, hög-
15	deež	deeZv	bästa delen (av mat)
16	zöör	zOOrv	rikedom
17	keel	keelv	foster (hos djur)
18	keer	keerv	öppet fält, stäpp
19	kiid	kiidv	kloster
20	koof	koofv	kaffe
21	nuuc	nuucv	hemlighet; hemlig
22	ööd	OODv	på, upp, mot
23	öön	OOnv	jubileum, årsdag
24	öör	OOrv	röta
25	öör	OOrv	vid, i närheten av
26	saak	saakv	just han/hon, densamme
27	saal'	saaLv	(mjölk av en) mjölkning
28	saam	saamv	mjölkning
29	siik	siikv	örhänge
30	syyw	syywv	famn, fång
31	suuž	syyZv	lår (kroppsdel)
32	syyr	syyrv	grund, fundament, bas
33	taal	taalv	vilja
34	tääž/tääš	tEEZv/tEESv	furste
35	tiiz	tiizv	stämpel, sigill
36	tiim	tiimv	sådan; så, sålunda
37	toon	toonv	rököppning i jurta
38	toor	toorv	persika
39	tooč	tooCv	bokförare



40	tööl	tOOLv	stjärna, vit fläck (i pannan på djur)
41	töön	tOONv	bränning (i folkmedicin)
42	tuuž	tuuZv	historia; historisk
43	tuul'	tuuLv	saga
44	tyyk	tyykv	berättelse
45	uud	uudv	betsel
46	uul	uulv	berg
47	uul'	uuLv	uggla
48	uut	uutv	stor skinsäck
49	uuc	uucv	hästlånd
50	uuš	uuSv	matvaror
51	yyc	yycv	beredning av kött inför vintern
52	xääw	xEEwv	åra
53	xäär	xEErv	medlidande
54	xääč	xEEcv	sax
55	xöön	xOONv	efter
56	xööt	xOOTv	nordlig; nästa
57	xööč	xOOCv	fåraherde
58	xuur	xuur	trad. kalm./mong. stränginstrument
59	xuux	xuuxv	skalp
60	cuuw	cuuwv	braxen
61	caaž	caaZv	lag, lagsamling
62	čeež	CeeZv	bröst
63	eem	eem	axel

### 5.2.4 Enstaviga verb med lång vokal (35 st.)

Som nämndes ovan (1.2; 3.1) hamnar de enstaviga verbstammarna med lång vokal i en kategori för sig. Detta beror på de ”naturliga” morfologiska skillnaderna mellan nomina och verb – nominalböjning och verbalböjning innebär helt enkelt olika regler.

De enkeltecknade formerna för orden i nedanstående lista är: imperativ1 (för 2p.), imperativ2 (för 3p.), ”momentalgerundium”.

Som uppslagsform står verben i fut.part., dvs. så som de normalt listas i kalmuckiska ordböcker.

	<b>verb</b>	<b>LF för PC-Kimmo</b>	<b>betydelse</b>
1	äax	EE	vara rädd, frukta
2	bäax	bEE	vara
3	boox	boo	binda/linda om, förbinda
4	buux	buu	gå ned; landa
5	geex	gee	tappa bort
6	giix	gii	sluta (regna)
7	gyyx	gyy	springa; rinna
8	huux	huu	be, tigga
9	daax	daa	förmå (lyfta något); tåla
10	žiix	zhii	sträcka ut
11	zööx	zOO	bära; transportera
12	zuux	zuu	bita
13	zyyx	zyy	knäppa (fast)
14	jyyx	jyy	ösa (upp ur)
15	kööx	kOO	jaga iväg; svälla
16	naax	naa	klistra (fast)
17	niiix	nii	snyta sig
18	nuux	nuu	gömma, dölja
19	nyyx	nyy	flytta (om nomader)
20	ööx	OO	kassera
21	saax	saa	mjölka
22	suux	suu	sitta
23	taax	taa	gissa; förmoda
24	täax	tEE	offra, frambringa offer
25	teex	tee	transportera
26	toox	too	visa uppmärksamhet; bjuda
27	tuux	tuu	driva (boskap)
28	tyyx	tyy	samla, plocka
29	uux	uu	dricka
30	xaax	xaa	stänga
31	xäax	xEE	leta, söka
32	cäax	cEE	ljusna; blekna; gråna (om hår)

33	šaaḫ	Saa	sticka/slå (in/ner i)
34	šyyx	Syy	sil, filtrera
35	eex	ee	värma

### 5.3 Transkribering

Kalmuckiska alfabetet	Transkribering för PC-Kimmo	I uppsatsen använd transkribering (vilken används för representation av såväl fonem, som grafem och morfem)	Ungefärligt motsvarande IPA-tecken
а (аа)	а – Q	а – аа	а
ә (әә)	Е – &	ä – ää	æ
б	б	б	б
в	в	в	в
г	г	г	г
һ	һ	һ	ɢ ʁ
д	д	д	д
дь, д(+<ю, я>)	Р	д'	d <sup>j</sup>
е (ее)	е – W	е – ее	ɛ
ж	З	ž	ʤ
з	з	з	з
и (ии)	и – Т	И – ии	и
й	й	й	й
к	к	к	k <sup>h</sup>
л	л	л	l (t)
ль, л(+<ю, я>)	Л	л'	ʎ
м	м	м	м
н	н	н	н
нь, н(+<ю, я>)	Н	н'	ɲ
ң	Г	ɳ	ŋ N
о (оо)	о – В	о – оо	о
ө (өө)	О - %	ö – öö	œ
п	п	п	p <sup>h</sup>
р	р	р	р
с	с	с	с
т	т	т	t <sup>h</sup>
ть, т(+<ю, я>)	Х	т'	t <sup>j</sup>
у (уу)	у – М	у – уу	у
ү (үү)	ү – Н	ү – үү	ʏ
ф	ф	ф	ф
х	х	х	χ x
ц	с	с	ʦ <sup>h</sup>

ч	C	č	$\overline{tj^h}$
ш	S	š	ʃ
ы	I	ï	i
ю	ju	ju, 'u	
я	ja	ja, 'a	
э (#ээ)	e – ee	e – ee	
-	v	v	ə ʒ ɪ

## Bibliografi.

1. **Antworth, Evan L.** (1990) *PC-KIMMO: a two-level processor for morphological analysis*, Dallas.
2. **Andra, B.** (Badmaev, A.) (1998) *Ojuni zul – lärobok i Todo bičiq*, Elista. [Б. Андра. *Оюни зул – Тод үзгин Цаһан тоһан даслур*, Элст]
3. **Bitkeev, P.** (1995) *De mongoliska folkens språk och skriftsystem*, Moskva-Elista. [Биткеев, П.Ц. *Языки и письменные системы монгольских народов*, Москва-Элиста]
4. **Bobrovnikov, A.** (1849) *Mongolisk-kalmuckisk grammatik*, Kazan'. [Бобровников, А.А. *Грамматика монгольско-калмыцкого языка*, Казань]
5. **Chomsky, Noam/Morris Halle.** (1968) *The sound pattern of English*, New York.
6. **Darvaev, P.** (1995) *Lärobok i kalmuckiska*, Elista. [Дарваев, П.А. *Самоучитель калмыцкого языка*, Элиста]
7. **Jaxontova, N.** (1996) *Det oiratiska skriftspråket på 1600-talet*, Moskva. [Яхонтова, Н.С. *Ойратский литературный язык XVII века*, Москва]
8. **Kalmuckisk-mongolisk-rysk ordbok**, Elista (1986). [Хальмг-моңыл-орс толь, Элст]
9. **Kara, György.** (1997) "Early Kalmyk primers and other schoolbooks – samples from Textbooks 1925-1930", *The Mongolia Society Special Papers 13*, Bloomington.
10. **Karttunen, Lauri.** (1983) KIMMO: a general morphological processor. *Texas Linguistic Forum* 22:163-186.
11. **Koskeniemi, Kimmo.** (1983) *Two-level morphology: a general computational model for word-form recognition and production*, Helsinki.
12. **Pavlov, D.** (2000) *Frågeställningar rörande kalmuckiska standardspråkets historia och byggnad*, Elista. [Павлов, Д.А. *Вопросы истории и строя калмыцкого литературного языка*, Элиста]
13. **Poppe, N.N.** (1960) *Buriat Grammar*, Bloomington.
14. **P'urbееv, G.** (1977) *Kalmuckisk grammatik*, Elista. [Пюрбеев, Г.Ц. *Грамматика калмыцкого языка*, Элиста]
15. **Rysk-kalmuckisk ordbok**, Moskva (1964). [Орс-хальмг толь, Москва]

16. **Sodmon, N.** (1999) *Solorγo*, Elista.
17. **Todaeva, B.** (2001) *Ordbok över xinjiang-oiraternas språk*, Elista. [Т. Буляш. *Шинжэнэ өөрд келнэ толь*, Элст]
18. **Jamca, T.** (1999) *Skrivregler för Todo bičig*, Hohhot. [Тү. Jamca. *Todo үзүгийн дүрүм*, Көкө-хото]