Uralic and Siberian Lexicology and Lexicography

Edited by
Sándor Szeverényi and Bayarma Khabtagaeva

Studia uralo-altaica 51

Uralic and Siberian Lexicology and Lexicography

Edited by Sándor Szeverényi and Bayarma Khabtagaeva

Studia uralo-altaica 51

Uralic and Siberian Lexicology and Lexicography

Proceedings of the 4th Mikola Conference
14-15, November 2014

Edited by Sándor Szeverényi and Bayarma Khabtagaeva

Szeged, 2017

University of Szeged
Department of Altaic Studies
Department of Finno-Ugrian Philology

ISBN 978-963-306-548-8
ISBN 978-963-306-549-5 (pdf)

ISSN 0133-4239

Content

Content 5

Foreword 8

Typology of the Ket finite verb 9
Edward Vajda

Sayan Turkic reindeer terminology 39
Elisabetta Ragagnin

On the Mongolian verb of motion *yav-* ‘to go, to travel, to leave’ 51
Veronika Zikmundova

Tungusic loanwords in Yeniseian languages 75
Bayarma Khabtagaeva

Lexical review of disease names in the Udmurt language 89
Rebeka Kubitsch

The Nganasan lexicon from a diachronic onomasiological point of view: The case of metonymy 107
Sándor Szeverényi

The first workday or the Moon’s day? Germanic and Slavic traditions in naming the days of the week in the Finnic languages 123
Kasperi Hasala

The possessive plural marker in the Burgenland dialect of Hungarian in Austria 131
Oshima Hajime

Russian impact on northern Khanty conditional sentences 149
Mária Sipos

A language without ‘get’? 173
Katalin Sipőcz

The grammaticalization of Northern Mansi *mā* ‘earth, world, land, place’ 187
Bernadett Bíró

Compounding in Aral–Caspian Kipchak languages 203
Bence Grezsa

Foreword

The 4th Mikola Conference was organized at the Finno-Ugric Department of the University of Szeged on November 13–14, 2014, in commemoration of the memory of Tibor Mikola, chair of the department for a quarter of a century and remembered primarily as a scholar of Samoyedic languages.

Despite compiling an Enets dictionary and creating the basis for the Nganasan morphological dictionary, Tibor Mikola is not remembered as a lexicologist or lexicographer, although, as his works demonstrate well, he was concerned with such issues, from both descriptive and historical perspectives.

"Uralic" and "Siberian" in the name of the conference and the title of the present volume refer to the fact their focus goes beyond Uralic language to include other indigenous Siberian languages as well. Such a broader focus of the volume symbolically signals that Tibor Mikola always paid attention to research into other Siberian languages and also regarded issues of language contact and typology to be of great importance in his work.

The present volume includes the written version of a selection of the papers presented at the conference.

The editors

Typology of the Ket finite verb

Edward Vajda

Western Washington University

1. Introduction

Finite verbs in Yeniseian are known for their formal elaboration, which strikes an obvious contrast with the comparatively straightforward suffixal agglutinating morphologies of other Inner Asian languages. The best-described Yeniseian language is Ket, the family's sole surviving member, spoken in three closely related dialects by a few dozen elderly people in central Siberia near the Yenisei River in the Turukhansk District of Russia's Krasnoyarsk Krai. New publications on the Ket finite verb over the past two decades make it possible to describe this unusual morphological object succinctly yet with precision, and even offer diachronic explanations for its most unusual features. The article consists of two main sections and a conclusion. The first section describes the verb's formal architecture, explaining what information must be included in the lexical entry of a Ket finite verb. The second explains how inflectional and lexical semantic categories are expressed, and offers a diachronic perspective on how Yeniseian verb structure developed. The conclusion summarizes how form and function in the Modern Ket verb are often mismatched, a situation arising in large part from the areal position of Yeniseian as an isolated microfamily surrounded by languages of a radically different morphological type. Along the way, credit is given to scholars whose research has shed important light on Ket verb structure. The present author also points out his own earlier missteps in analyzing this challenging morphological system. The article is dedicated to the outstanding typologist Bernard Comrie. As director of the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany) from 1998 to 2015, he invited Siberianists from across the globe to work collaboratively, greatly advancing the synchronic description of minority Eurasian languages as well as illuminating their genealogical and areal relationships.

2. The formal architecture of Ket verb morphology

The most striking feature of the Ket finite verb is its rigid position class structure, which involves a complex interdigitation of lexical and grammatical morphemes quite unlike the neighboring suffixal agglutinating languages. The authors of the first two fundamental descriptions of the Ket verb (Dul'zon 1968; Krejnovich 1968) essentially treated every stem as irregular, which was fortuitous since their descriptions were much richer in conjugated example forms than might otherwise have been the case. In actual fact, many regular patterns underlie the formal surface diversity, including three core principles of structure. The first is that virtually every conjugated form adheres to one and the same position-class template, though historical investigation shows that metathesis has altered this linear arrangement in significant ways. Second, the template's lexical morphemes are discontinuous, being separated by grammatical affixes in most forms. Third, the Ket verb stem is a position-class formula that specifies not only the lexical morphemes but also the locations of subject and object agreement markers and the choice of tense-mood morpheme shapes – neither feature being predictable by any overarching grammatical principle.

Let us begin with the position-class template that underlies every Ket verb stem and regulates its conjugated forms:

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
sbj person agr or detrans marker	incor-porated noun, modifier, or action nominal	obj or sbj agr	thematic consonant(s) (*k, t, d, q, h, n, ŋ, kd, kt, nt, ŋt*)	3 anim agr or conj (*s/i ~ a/o*)	3 inan agr *b* or thematic *b*	tense-mood (*n~l*)	1,2 sbj or obj agr or result affix	BASE (verb root or lexical aspect marker)	anim pl sbj agr

Table 1.
Ket finite verb template (sbj = subject, obj = object, agr = agreement, anim = animate-class, inan = inanimate class, detrans = detransitive, pl = plural, result = resultative, conj = conjugation)

The position-class approach to Ket verb morphology was pioneered independently by Butorin (1995), Reshetnikov and Starostin (1995), and Werner (1997), while Vajda (2001) developed the template model shown above. The examples in (1) contain diverse morphological combinations, but each verb form can easily be divided according to the ten position-classes listed in Table 1:

(1) Ket verbs with diverse position class combinations

a. *d*^{8}*-əla*^{7}*-bo*^{6}*-k*^{5}*-s*^{4}*-aq*^{0}*-in*$^{-1}$
3SBJ^{8}-out^{7}-1SG.OBJ^{6}-with^{5}-PRES^{4}-go^{0}-ANIM.PL.SBJ^{-1}
'They lead me out.'

b. *k*^{8}*-eda*^{7}*-q*^{5}*-o*^{4}*-l*^{2}*-di*^{1}*-da*^{0}
2SBJ^{8}-send^{7}-INCEPT^{5}-PAST$^{4/2}$-1SG.OBJ^{1}-ITER.TRANS^{0}
'You (SG) used to send me.'

c. *tip*^{7}*-di*^{1}*-bed*^{0}
dog^{7}-1SG.SBJ^{1}-have^{0}
'I have a dog.'

d. *ba*^{6}*-k*^{5}*-in*^{2}*-saal*^{0}
1SG.SBJ^{1}-TC^{5}-PAST^{2}-spend.night^{0}
'I spent the night.'

Several position classes display surprising alternations in function. The semantic contrasts between subject vs. object (in positions 6, 4, 3 and 1), verb root vs. modifier (position 7), and verb root vs. aspect or transitivity marker (base position) – are in fact strictly constrained and also afford pathways toward understanding the verb's historical development. In other words, there is astonishingly order submerged within the apparent chaos.

The second core principle of Ket finite verb structure is that the basic meaning of the stem is expressed by lexical morphemes in the non-adjacent positions P7, P5, and P0. Verb stems can be categorized according to which of these three positions are actually filled. Position P0 is called the 'base' because it represents the original locus of the verb root in the template. The oldest and structurally most basic stems lack morphemes in P7 and P5, so that their conjugated forms contain only the base and grammatical affixes:

(2) Simple, root-final Ket verb stems (lacking P7 and P5)

a. *k*^{8}*-il*^{2}***-aq*^{0}***-in*$^{-1}$
2SBJ^{8}-PAST^{2}-go^{0}-ANIM.PL.SBJ^{-1}
'You (PL) went (and returned).'

b. *du*^{8}*-n*^{2}***-qo*^{0}**
3MASC.SBJ^{8}-PAST^{2}-**die**^{0}**
'He died.'

c. *du*^{8}*-di*^{2}***-ted***^{0}
3MASC.SBJ^{8}-1SG.OBJ^{1}-**beat**^{0}
'He beats me.'

d. *da*^{8}*=b*^{3}*-il*^{2}***-a***^{0}
3FEM.SG.SBJ^{8}=3INAN.OBJ^{3}-PAST^{2}-**eat**^{0}
'She ate it.'

e. *du*^{8}*-b*^{3}*-in*^{2}***-bəkʲ***^{0}*-n*$^{-1}$
3SBJ^{8}-3INAN.OBJ^{3}-PAST^{2}-**stretch**^{0}-ANIM.PL.SBJ^{-1}
'They stretched it.'

Although every stem fills the base position, on rare occasions the base morpheme itself elides phonologically in certain forms, as in the Southern Ket pronunciation of (2c) *dabīˑl* 'she ate it' (cf. Central Ket *dabīˑla* 'she ate it'). Verbs that contain a base but no P7 or P5 are strongly prefixing, their only suffix being the animate-class subject agreement marker in position P-1. Why a clitic boundary (=) separates the P8 subject marker from the rest of the morphological verb in forms like (2d) but not in others is explained in section 2.2 and can be ignored for now.

Many verb stems also contain a thematic consonant in position P5. These morphemes are usually difficult to etymologize, but in a small number of basic verbs they add a definable meaning to the verb stem.

(3) Root-final Ket verbs with P5 thematic consonant

a. *d*^{8}*=****n***^{5}*-a*^{4}*-b*^{3}*-do*^{0}
3SBJ^{8}=**round**^{5}-PRES^{4}-3INAN.OBJ^{3}-chop^{0}
'He fashions it by hewing (a round shape).'

b. *d*^{8}*=****d***^{5}*-a*^{4}*-b*^{3}*-do*^{0}
3SBJ^{8}=**long**^{5}-PRES^{4}-3INAN.OBJ^{3}-chop^{0}
'He fashions it by hewing (a long object, such as a canoe).'

c. *d*^{8}*=****t***^{5}*-a*^{4}*-b*^{3}*-do*^{0}
3SBJ^{8}=**surface**^{5}-PRES^{4}-3INAN.OBJ^{3}-chop^{0}
'He cuts it (rough-hews an object's surface).'

d. *d*^{8}*=****k***^{5}*-a*^{4}*-b*^{3}*-do*^{0}
3SBJ^{8}=**away**^{5}-PRES^{4}-3INAN.OBJ^{3}-chop^{0}
'He clears it (creates a trail by cutting away underbrush).'

Unlike these examples with the base *-do* ‘cut’, the P5 affix in most stems is semantically opaque and must be glossed TC for ‘thematic consonant’. A few stems contain more than one thematic consonant, showing that the P5 slot contains three linearly distinct classes of morphemes. These include, in order from left to right: 1) postpositions such as *k* ‘with’ that require a preceding object marker; 2) spatial prefixes like *d* ‘long, along’, *h* ‘flat surface, area’, or *n* ‘round, around’; and 3) the pluractional marker *t*. Example (4a) contains a single thematic consonant, while (4b) contains two.

(4) Rare example of stem with concatenated P5 thematic consonants

a. *d⁸=**t⁵**-aj⁴-ka⁰*
3SBJ⁸=**PLURACTIONAL⁵**-PRES⁴-travel⁰
‘He walks around (in various directions).’

b. *d⁸=bo⁶-**k/t⁵**-aj⁴-ka⁰*
3SBJ⁸=1SG.OBJ⁶-with/**PLURACTIONAL⁵**-PRES⁴-travel⁰
‘He leads me around (in various directions).’

Because concatentations of thematic consonants are completely unproductive, they occupy a single position class (P5) in the modern template.

Let us now turn to the last of the three core lexical position classes, the P7 incorporate slot. This position is usually unfilled in basic vocabulary. However, all productive stem patterns fill both P7 and P0, and most contain a P5 thematic consonant, as well. Stems with P7 can be called ‘compound stems’. In compound stems with a semantically salient base, the P7 slot incorporates a noun (5a), adjective (5b), directional (5c), or adverb (5d) modifying the P0 base verb root.

(5) Ket incorporation in compound stems

a. *d⁸=**suul⁷**-il²-bed⁰-n⁻¹*
3SBJ⁸=**snowsled⁷**-PAST²-make⁰-ANIM.PL.SBJ⁻¹
‘They made a snowsled.’

b. *d⁸-**ugde⁷**-t⁵-a⁴-b³-sin⁰*
3SBJ⁸-**long⁷**-TC⁵-PRES⁴-3INAN.OBJ³-change⁰
‘He elongates it.’

c. *d⁸-**aka⁷**-u⁶-k⁵-s⁴-aq⁰-in⁻¹*
3SBJ⁸-**river.to.forest⁷**-3INAN.OBJ⁶-with⁵-PRES⁴-go⁰-ANIM.PL.SBJ⁻¹
‘They carry it from the river up into the forest.’

d. ***hel*^{7}-*t*^{5}-*a*^{4}-*b*^{3}-*qut*^{0}**
out.of.place^{7}-TC^{5}-PRES^{4}-3INAN.SBJ^{3}-assume.position^{0}
'It slips out of place.' ~ 'It gets dislocated.'

In a majority of the three dozen or so productive Ket verb stem patterns, however, P7 contain what is called an 'action nominal' rather than a true incorporate. The P0 base in such stems has eroded semantically, and the action nominal expresses the stem's basic lexical meaning. The examples in (6) contain the action nominal *bakdeŋ* 'pull':

(6) Ket action nominal-based verb stems (suffixing use of the template)

a. ***bakdeŋ*^{7}-*ba*^{6}-*k*^{5}-*a*^{4}-*qan*^{0}**
pull^{7}-1SG.SBJ^{6}-TC^{5}-PRES^{4}-**MOM.INCEPT**^{0}**
'I start pulling (once).'

b. *d*^{8}=***bakdeŋ*^{7}-*ku*^{6}-*k*^{5}-*a*^{4}-*bed*^{0}-*n*$^{-1}$**
1SBJ^{8}=**pull**^{7}-2SG.OBJ^{6}-TC^{5}-PRES^{4}-**ITER.TRANS**^{0}-ANIM.PL.SBJ^{-1}
'We keep (on) pulling you (SG).'

c. *d*^{8}=***bakdeŋ*^{7}-*q*^{5}-*in*^{2}-*ku*^{1}-*t*^{0}-*n*$^{-1}$**
1SBJ^{8}=**pull**^{7}-INCEPT^{5}-PRES^{4}-2SG.OBJ^{1}-**MOM.TRANS**^{0}-ANIM.PL.SBJ^{-1}
'We started pulling you (SG) once.'

d. *d*^{8}=***bakdeŋ*^{7}-*q*^{5}-*a*^{4}-*ku*^{1}-*da*^{0}**
1SBJ^{8}=**pull**^{7}-INCEPT^{5}-PRES^{4}-2SG.OBJ^{1}-**ITER.TRANS**^{0}**
'We repeatedly (start to) pull you (SG).'

Used as a separate word outside the finite verb, an action nominal such as *bakdeŋ* would express the following meanings, depending on context: 1) infinitival 'to pull', 2) gerundive '(the act of) pulling', and 3) participial '(someone who is) pulling' or '(someone or something that was or is being) pulled'. When incorporated into the P7 slot of a verb stem, the action nominal serves as the verb's semantic head. The P0 base in such stems has become grammaticalized as a suffix specifying transitivity (TRANS vs. INTRANS), start of action (INCEPT), or single complete action (MOM) in contrast to multiple or persistent action (ITER).

Krejnovich (1968) first recognized the difference between strongly prefixing stems with their semantic head in the original root position (the P0 base), and suffixing stems with a salient lexical morpheme near the verb's beginning (in the P7 incorporate position). Vajda (2009) argued that areal influence from the surrounding

suffixing languages led to "pseudo-incorporation" of action nominals in P7 and semantic bleaching of the P0 base

a. Verb root in base position (P0) and original prefixing configuration of position classes:

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
sbj pers. agr	incor-porated noun or modifier	obj or sbj agr	thematic consonant(s) (*k, t, d,* etc.)	3 anim agr or conj (*s/i ~ a/o*)	3 inan agr or themat. *b*	tense-mood (*n ~ l*)	1,2 sbj or obj agr	BASE (verb root)	anim pl sbj agr

b. Action nominal in P7 and innovative suffixing configuration:

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
sbj pers. agr	action nominal as semantic head	obj or sbj agr	thematic consonant (s) (*k, t, d,* etc.)	3 anim agr or conj (*s/i ~ a/o*)	3 inan agr or themat. *b*	tense-mood (*n ~ l*)	1,2 sbj or obj agr	BASE (aspect or transiti-vity suffix)	anim pl sbj agr

Table 2. Contrast between prefixing (a) and suffixing (b) verb models

The typological shift from prefixing to suffixing is also implicated in the partial change of P8 subject person agreement morphemes from prefix to special clitic. This is a convenient place to describe the phonological behavior of this position class, before concluding the discussion of lexical morpheme categories. Reshetnikov and Starostin (1995) were the first to identify the P8 subject markers as special clitics. Because they often elide, previous researchers were unaware of their structural presence in the morphological verb, making it impossible to give an accurate typological assessment of the Ket agreement system.

Vajda (2001) first identified the complex conditions under which P8 allomorphs are realized. They appear as syllabic prefixes only before certain short position-class strings:

P5-a^4-P0	P3-(n^2)-P0	P0
***di*^{8}-*k*^{5}-*a*^{4}-*daq*^{0}** 'I live'	***di*^{8}-*b*^{3}-*ted*^{0}** 'I hit it'	***di*^{8}-*doq*^{0}** 'I fly'
***ku*^{8}-*k*^{5}-*a*^{4}-*daq*^{0}** 'you (SG) live'	***ku*^{8}-*b*^{3}-*ted*^{0}** 'you (SG) hit it'	***ku*^{8}-*doq*^{0}** 'you (SG) fly'
***du*^{8}-*k*^{5}-*a*^{4}-*daq*^{0}** 'he lives'	***du*^{8}-*b*^{3}-*ted*^{0}** 'he hits it'	***du*^{8}-*doq*^{0}** 'he flies'
***də*^{8}-*k*^{5}-*a*^{4}-*daq*^{0}** 'she lives'	***də*^{8}-*b*^{3}-*ted*^{0}** 'she hits it'	***də*^{8}-*doq*^{0}** 'she flies'

Table 3. Environments that preserve the full prefixal forms of P8 markers

Before other position-class strings, feminine/inanimate *də*^{8}- changes its vocalism from [də] to [da]. Unlike *də*^{8}-, which is always a prefix, *da*^{8}= is a clitic that can attach either to the verb or to a preceding word, as it normally does in fast speech, especially if the preceding word ends in a vowel. In examples with morpheme breakdown, the clitic boundary [=] after *da*^{8} indicates its ability to use either the verb or a preceding word as its host:

proclitic on verb	enclitic on preceding word
***da*^{8}=*o*^{4}-*l*^{2}-*daq*^{0}** ‘she lived’	*bu*=***da*^{8}** *o*^{4}-*l*^{2}-*daq*^{0} ‘she lived’ (*bu* ‘she’)
***da*^{8}=*b*^{3}-*il*^{2}-*bed*^{0}** ‘she made it’	*bu*=***da*^{8}** *b*^{3}-*il*^{2}-*bed*^{0} ‘she made it’
***da*^{8}=*in*^{2}-*doq*^{0}** ‘she flew’	*bu*=***da*^{8}** *in*^{2}-*doq*^{0} ‘she flew’

Table 4. Allomorphs of P8 feminine- and inanimate-class markers

The shape taken by the other three P8 markers in longer verb forms depends crucially on whether they precede a vowel or consonant. Before a vowel they assume the non-syllabic prefix forms *d*- (1st person, 3rd person masculine-class singular or animate-class plural) and *k*- (2nd person), which cannot encliticize to a preceding word. In this way they differ from feminine- or inanimate class *da*^{8}=, which exhibits clitic behavior in longer strings even before a vowel (7b).

(7) Allomorphs of P8 markers in longer verbs strings before a following vowel

a. Prefix form only before a vowel-initial verb string

quska ***d*^{8}**-*o*^{4}-*l*^{2}-*daq*^{0}	‘I lived in a tent.’
quska ***k*^{8}**-*o*^{4}-*l*^{2}-*daq*^{0}	‘You (SG) lived in a tent.’
quska ***d*^{8}**-*o*^{4}-*l*^{2}-*daq*^{0}	‘He lived in a tent.’

b. Prefix/clitic alternation in 3rd person feminine/inanimate markers

quska ***da*^{8}=***o*^{4}-*l*^{2}-*daq*^{0} ‘She lived in a tent.’
(*quska* # ***da*=***oldaq* ~ *quska*=***ra*** # *oldaq*)

Before a consonant, the shortened allomorphs *d* (1st person), *k* (2nd person), *d* (3rd person masculine-class singular or animate-class plural) normally elide unless they can encliticize to a preceding vowel-final word:

(8) Allomorphs of P8 markers in longer verbs strings before a consonant

d^{8}=*nan*^{7}-*s*^{4}-*ibed*^{0}	‘I make bread’ (pronounced *nansibed* or *=d # nansibed*)
k^{8}=*nan*^{7}-*s*^{4}-*ibed*^{0}	‘you (SG) make bread’ (*nansibed* or *=k # nansibed*)
d^{8}=*nan*^{7}-*s*^{4}-*ibed*^{0}	‘he makes bread’ (*nansibed* or *=d # nansibed*)
da^{8}=*nan*^{7}-*s*^{4}-*ibed*^{0}	‘she makes bread’ (*da=nansibed* or *=da # nansibed*)

The unusual allomorphy of P8 markers follows the Modern Ket preference for verbs with a lexical root in the first syllable, whether this root be an action nominal or a true incorporate. This tendency arose through influence from the surrounding suffixing languages and is not seen in the closely related Yugh language, where most P8 morphemes remained prefixes. In Ket, cliticization of P8 agreement markers fails to occur only in short strings like *du*^{8}*-b*^{3}*-ted*^{0} 'he hits it', where the P8 syllable was retained to prevent the P0 base from occupying the verb's initial syllable, an ancient restriction apparently inherited from Proto-Yeniseian.

Returning to our discussion of stem elements, recall that the core lexical material is normally found in positions P7-P5-P0. The remaining seven position classes usually contain subject/object agreement or tense-mood affixes. However, three of the agreements slots occasionally contain lexical affixes that remain unchanged in all conjugated forms of the given stem. In slot P8 the inanimate-class agreement marker *da=*^{8} was occasionally reanalyzed as a detransitivizer on the basis of transimpersonal constructions such as 'it reddens me' → 'I blush' (Nefedov, Malchukov and Vajda 2011):

(9) Verb form with detransitivizing marker *da=*^{8}

***da*^{8}*=sulej*^{7}*-bo*^{6}*-k*^{5}*-s*^{4}*-a*^{0}
DETRANS^{8}**=**red^{7}-1SG.SBJ^{6}-TC^{5}-PRES^{4}-process.occurs^{0}
'I blush.'

Another lexical morpheme occurs in position P1, which normally expresses subject or object agreement. In some intransitives, P1 instead contains the resultative prefix *a-* ~ *ja* ~ *aja*. This ancient element, which was part of the stem's lexical aspect system, once appeared in tandem with an intransitive or resultative suffix *-ej* or *-ŋ*, creating a sort of circumfix around the verb root. The suffix occupied a separate suffixal slot in Proto-Yeniseian, but in Modern Ket it interacts so irregularly with the verb root that Vajda (2001, 2004) treated it as part of the P0 base. The intransitive/resultative suffix remains discernable as a separate position class in Kott (Vajda, in press 2) and was also treated as a separate slot in Butorin's Ket template (Butorin 1995).

(10) Verb form with resultative *a-*^{1}

a^{4}*-b*^{3}***-a*^{1}***-bed*^{0}
PRES^{4}-3INAN.SBJ^{3}-**RESULT**^{1}-wipe^{0}
'It is wiped.'

Finally, some stems have a lexical affix in P3, a position that normally contains the inanimate-class agreement marker *b*[3]. Thematic *b*[3] resulted from metathesis of a consonant formerly occupying either P5 or P0. Some instances of lexical *b*[3] arose when the P5 area prefix *h* (from earlier **p*) co-occurred with another P5 thematic consonant and jumped forward across the P4 tense/mood marker into position P3. Because the metathesis also voiced **p* to *b* (why this happened is not clear), it now resembles the homonymous inanimate-class agreement marker *b*[3], which occupies the same position much more frequently:

(11) Example of metathesis of P5 thematic **p* into P3

a. *d*[8]*=qoqə*[7]*-ba*[6]***-h*[5]***-a*[4]*-ted*[0]
3SBJ[8]=fist[7]-1SG.OBJ[6]-area[5]-PRES[4]-hit.endwise[0]
'He punches me (once).'

b. *d*[8]*=qoqən*[7]*-ba*[6]***-t*[5]***-a*[4]***-b*[3]*-ted*[0] (< **d*[8]*=qoqn*[7]*-ba*[6]***-p/t*[5]*-a*[4]*-ted*[0])
3SBJ[8]=fists[7]-1SG.OBJ[6]-**PLURACTIONAL**[5]-PRES[4]-**area**[3]-hit.endwise[0]
'He punches me (repeatedly).'

The second type of thematic *b* in position P3 resulted from a different metathesis. Where the initial *b* of a P0 base derived historically from **w*, it metathesized into position P3 in past-tense forms. In Modern Ket, this element must be regarded as lexical, though it has no independent function apart from the verb root, the phonological remainder of which still occupies position P0:

(12) Leftward metathesis of base anlaut *b* < **w* into P3

a. *da*[8]*=ba*[6]***-t*[5]*-a*[4]***-bet*[0]
DETRANS[8]=1SG.SBJ[6]-TC[5]-PRES[4]-**feel**[0]
'I understand.'

b. *da*[8]*=ba*[6]*-t*[5]*-o*[4]***-b*[3]*-n*[2]***-et*[0] (pronounced *bato**m**n**et*** < **d*[8]*=ba*[6]*-t*[5]*-o*[4]*-n*[2]***-wet*[0])
DETRANS[8]=1SG.SBJ[6]-TC[5]-PAST[4]-**TC**[3]-PAST[2]-**feel**[0]
'I understood.'

Vajda (2004: 66–68) originally misinterpreted the various metathesized *b*-elements in position P3 as morphemes expressing lexical categories such as 'applicative', 'involuntary causative' or 'intensive', seeing them as etymologically derived from inanimate-class *b*[3]. Even native speakers of Ket occasionally reanalyze thematic *b*[3] as inanimate-class agreement, replacing it by analogy with other object markers; Krejnovich (1968: 91) gives an example of such non-canonical forms,

which exhibit multi-site object marking, since the P6 object markers are also present. Section 2 below will discuss other instances of reanalysis that actually did lead to permanent shifts in conjugation class.

Table 5 summarizes the three primary lexical position classes (shaded) and lexical morpheme categories (bold print) found in the Ket verb template.

P8	**P7**	P6	**P5**	P4	P3	P2	P1	**P0**	P-1
sbj agr or **detrans *da=***	**incor-porated noun, modifier, or action nominal**	obj or sbj agr	**thema-tic conso-nant (*k, t, d, q, h, n, ŋ, kd, kt, nt, ŋt*)**	3 anim agr or conj (s/*i* ~ *a*/*o*)	3 inan agr or **themat. *b* (< P5 **p* or P0 anlaut **w*)**	aspect (n ~ l)	1,2 agr or **result affix**	**BASE (verb root or lexical aspect marker**	anim pl sbj agr

Table 5.
Summary of lexical position classes and morphemes in Modern Ket finite verbs

As mentioned earlier, every verb has a 'formulaic stem'. There are the three core lexical positions P7-P5-P0 and any lexical affixes occupying positions P8, P3 and P1 in place of the agreement morphology usually found in these slots. Two additional pieces of information must also be included in the verb's lexical entry. First, tense-mood inflection involves an unpredictable choice of affix shapes in positions P4 and P2. The positions themselves are identical for most stems, but the combination of morphemes – *s*/*i* or *a*/*o* in P4 with *l* or *n* in P2 – is lexically idiosyncratic. Second, the positions used to express subject and object agreement cannot be predicted based on any overall semantic or formal pattern; therefore, they must be listed as part of the lexical entry, as well. This can be done by placing capital letters S and O in the appropriate position(s) in the stem formula or by placing a designation for agreement class such as *vt1* (transitive class I) or *v5* (intransitive class V) after the stem. Superscript numerals can specify position class, but this becomes unnecessary after one becomes more familiar with the verb structure. Here are three ways Ket verb stems might be listed in a dictionary:

(13) Different formalisms for listing the Ket stem meaning 'S leads O out':

a. Minimal formula: *əla*^{7}-*k*^{5}-*aq*^{0} (*s*/*il* tense/mood class, transitive agreement II)

b. Extended formula (Kotorova and Nefedov 2015): *əla*^{7}-*k*^{5}-[*s*^{4}]-[*l*^{2}]-*aq*^{0} (*vt2*)

c. Maximal formula (this article): S^{8}-*əla*^{7}-O^{6}-*k*^{5}-(*s*^{4}/*l*^{2})-*aq*^{0}-$S_{PL}$$^{-1}$

The minimal formula in (13a) contains only the three core lexical morphemes classes, with information about tense-mood and agreement class provided afterward. The extended formula used in Kotorova and Nefedov's (2015) *Comprehensive Ket Dictionary* includes all lexically fixed morpheme shapes, including detransitive P8, thematic P3, and resultative P1, as well as the P4 and P2 tense-mood shapes, which appear in square brackets if they occur only in some forms; information about agreement class is given outside the formula, and the reader must become intimately familiar with the positions occupied by subject and object markers in each of the five productive intransitive agreement types and three transitive agreement types. The maximal approach lists agreement positions and tense-mood affix shapes in the stem itself. This approach is used in Vajda & Werner's (in preparation) *Etymological Dictionary of the Yeniseian Languages*, though without superscript numbers. Because the P7 incorporate is syllabic and the P5 thematic consonant is not, the position of agreement markers is discernable even without numbering. The present article marks position class explicitly to help explain basic Ket verb structure as clearly as possible.[1]

Each of these formalisms has its merits. The extended formula (13b) is well suited for a synchronic dictionary, while the maximal formula (13c) is useful in an etymological dictionary where finite verb formulas appear under entries headed by a lexical root. The architecture of stem formulas will become clearer after tense-mood marking and subject/object agreement are discussed more fully in the next section.

2. The templatic expression of inflectional categories and lexical semantic patterns

Despite its formal exuberance, the Ket verb contains only two inflectional subsystems: tense-mood marking and subject/object agreement. Let us look at each in turn and identify the idiosyncrasies associated with them that must be listed in the verb's lexical entry, before moving on to lexical semantics in 2.3.

2.1. Tense-mood marking

Stems regularly inflect to distinguish three synthetic tense-mood forms: past indicative, non-past indicative (interpretable as either present or future tense) and

[1] I thank Andrey Nefedov for suggestions on the wording of this section, and for helping proofread the article.

imperative (limited to stems expressing intentional actions). Other grammatical tense-mood-aspect categories can be expressed by pre-verbal particles (*as* 'future action', *ba* 'habitual past action', *sim* 'conditional', etc.). Formally, tense-mood inflection is based on two interacting sets of affixes in position P4 (conjugation markers of unknown original meaning) and P2 (originally aspect markers). Table 6 highlights the template's tense-mood position classes:

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
sbj pers. agr or detrans marker	incor-porate	obj or sbj agr	the-matic cons (*k, t, d,* etc.)	3 anim agr or **conj *s/i ~ a/o,*** (rarely ***q***)	3 inan agr or themat. *b*	**tense-mood *n ~ l,*** (rarely ***j***)	1,2sbj or obj agr or result affix	BASE	anim pl sbj agr

Table 6. Tense-mood-aspect related positions in the Ket verb template

The template's historical development cannot be understood without taking metathesis into account (Vajda 2013). We have already seen how metathesis led to the rise of thematic *b* in position P3. The "authentic" inanimate-class agreement marker b^3 itself once occupied P4, forming a single position class with the 3rd person animate-class agreement markers. However, it metathesized ahead of non-past *a* and past-tense *o* due to the same phonological rule that sporadically switched labial and non-labial segments elsewhere (cf. non-metathesized Southern Ket *qo**qp**un* 'cuckoo bird' and metathesized Central and Northern Ket *qo**pq**un* 'cuckoo bird'). Labial metathesis did not occur in Kott, where the template retains its original ordering of 3rd person agreement followed by tense-mood-aspect-transitivity affixes:

(14) Kott verbs with inanimate-class agreement marker in original position P4

a. ***b*^{4}-*a*^{3}-*pi*^{0}**
3INAN.SBJ^4-PRES^3-grow^0
'It grows.'

b. ***m*^{4}-*a*^{3}-*n*^{2}-*a*^{1}-*pi*^{0}**
3INAN.SBJ^4-PAST^3-PERF^2-RESULT^1-grow^0
'It grew up.'

(15) Ket verbs with inanimate-class *b* metathesized forward into position P3

a. ***a*^{4}-*b*^{3}-*a*^{1}-*tij*^{0}**
PRES^4-**3INAN.SBJ^3**-INTRANS^1-grow^0
'It grows.'

b. *o*^{4}-***b***^{3}-*l*^{2}-*a*^{1}-*tij*^{0}
PAST^{4}-**3INAN.SBJ^{3}**-IMPERF^{2}-INTRANS^{1}-grows^{0}
'It was growing.'

More is said about metathesis in section **2.2**, as it has complicated the Ket agreement system in other ways too.

The locus of grammatical tense-mood marking in Modern Ket is thus the linearly separate position classes P4 and P2. The P4 slot usually contains one of two conjugation marker forms: *a*^{4} or *s*^{4}. Conjugation marker *a*^{4} regularly labializes to *o*^{4} in past-tense forms, but remains *a*^{4} in non-past and imperative forms. Conjugation marker *s*^{4} never appears in past indicative or imperative forms, being replaced by *i* or *Ø* whenever a P2 consonant appears[2]. This often occurs in non-past indicative forms, as well, since the phonetic realization of *s*^{4} appears to be subject to complicated rules (Vajda 2004: 74–76) that are not yet fully understood. There is no discernable functional difference between *a/o* and *s/i/Ø* verbs. The P4 allomorph *s* ~ *i* ~ *Ø* ~ *a/o* characteristic of a given stem developed through complex morphophonemic interactions with the surrounding affixes, processes that are no longer obvious. Also, when an animate-class agreement marker occupies P4, it normally cancels out or merges with the conjugation marker vowel in the same position.[3]

The P2 slot contains consonants that originally marked aspect. Most past-indicative and imperative forms have either *n*^{2} or *l*^{2}, which interact with *a/o*^{4} or *i/Ø*^{4} to form the inflectional tense-mood system of Modern Ket. Before most vowel-initial bases, imperative forms add *d*. The resulting combinations of *nd* and *ld* that appear before vowel-initial bases are counted as part of position class P2.

(16) Examples of imperative forms with *d*^{2}

a. *a*^{4}-*nd*^{2}-*i*^{0}
IMP$^{4/2}$-sharpen^{0}
'Sharpen it!'

[2] It is not always clear when *i* should be treated as a tense-mood marker occupying position P4 and when it is epenthetic. The present article treats it as *i*^{4} under word stress or when it is linearly separated from the P2 consonant. In other cases, it is included in P2 as *il*^{2} or *in*^{2}.

[3] Rarely, the agreement marker and conjugation marker preserve the original Proto-Yeniseian (and presumably pre-Ket) ordering: *d*^{8}-*ul*^{7}-*d*^{5}-***aŋ*^{4}-*s*^{3}**-*kəŋ*^{0} 'he washes them' (< 3MASC.SBJ^{8}-water^{7}-TC^{5}-**3ANIM.PL.SBJ^{4}-PRES^{3}**-wash^{0}). See Vajda (2001: 434-435) for more examples, though the phenomenon was not understood at the time.

b. *i*^{4}-*ld*^{2}-*il*^{0}
IMP$^{4/2}$-sing^{0}
'Sing!'

Imperative *d* is vestigial in Modern Ket, appearing only before bases that were vowel-initial in Proto-Yeniseian. The distribution of *l* and *n* across past-indicative and imperative forms is normally parallel. If a stem's past-indicative forms have *n*, then the imperative also has *n* (or *nd*). If the past-indicative has *l*, the imperative also has *l* (or *ld*). Non-past indicative forms, by contrast, have no P2 marker at all (except for some forms of the irregular stem 'know', which has *l*^{2} in past and non-past forms alike: *it*^{7}-*a*^{4}-*l*^{2}-*am*^{0} 'he knows / he knew'). The choice between *n*^{2} and *l*^{2} partly follows lexical aspect. P2 *n* appears in many stems denoting single complete actions or changes of state, a vestige of its original perfective aspect marking function. P2 *l* appears in the past tense of all stems that specifically denote multiple actions or static situations, a vestige of its former function of marking imperfective aspect. Verbs that express a state resulting from a prior action (called 'resultatives', 'stative-resultatives' or 'perfective-statives') likewise mark past tense with *l*^{2}. P2 *n* never appears in stems that specifically denote static states, ongoing activity, or repeated action. However, some stems with *l*^{2} express single complete actions or express either single complete action or repeated action, depending on context. It is thus not possible to say the P2 consonants mark grammatical aspect in Modern Ket. The choice of P2 consonant for past indicative and imperative is lexically fixed and must be listed in the stem formula. A tiny number of stems do show paradigmatic alternations between perfective *n*^{2} and imperfective *l*^{2} to signal single vs. multiple action, but this too is an idiosyncrasy of the stems in question and also must be listed as part of their lexical entry. A few common stem types have no P2 marker at all and signal tense solely by alternating non-past *a* with past-tense *o* in slot P4 (and sometimes in the P0 base vowel, as well, displaying a rare instance of Ket vowel harmony).

Most stems belong to one of five productive tense-mood classes, which must be specified in the verb's lexical entry:

(17) Tense-mood class ***s/in*** (= ***i/n***) (stems with ***s***^{4} and ***n***^{2} or only ***[i]n***^{2})

Example stem: *S*^{8}-*O*^{6}-*k*^{5}-*(s*^{4}*/n*^{2}*)-doq*^{0}-*S*$_{PL}$$^{-1}$
SBJ^{8}-OBJ^{6}-toward^{5}-(tense/mood$^{4/2}$)-jump^{0}-ANIM.PL.SBJ^{-1}
'S attacks O'

Non-past indicative: *d*^{8}-*a*^{6}-*k*^{5}-*s*^{4}-*doq*^{0} 'he attacks him'

Past indicative:	d^{8}-a^{6}-k^{5}-in^{2}-doq^{0}	‘he attacked him’
Imperative:	a^{6}-k^{5}-in^{2}-doq^{0}	‘attack him!’

(18) Tense-mood class ***s/il*** (= ***i/l***) (stems with ***s⁴*** and ***l²*** or only ***[i]l²***)

Example stem: S^{8}-nan^{7}-(s^{4}/l^{2})-bed^{0}-$S_{PL}{}^{-1}$
SBJ^{8}-bread^{7}-(tense/mood$^{4/2}$)-make^{0}-ANIM.PL.SBJ^{-1}
‘S makes bread’

Non-past indicative:	da^{8}=nan^{7}-$\boldsymbol{s^{4}}$-$ibed^{0}$	‘she makes bread’
Past indicative:	da^{8}=nan^{7}-$\boldsymbol{l^{2}}$-$ibed^{0}$	‘she maked bread’
Imperative:	nan^{7}-$\boldsymbol{il^{2}}$-get^{0}	‘bake bread!’

(19) Tense-mood class ***a/on*** (stems with ***a/o⁴*** and ***n²***)

Example stem: S^{8}-k^{5}-$(a/o^{4}n^{2})$-qut^{0}-$S_{PL}{}^{-1}$
SBJ^{8}-up^{5}-(tense/mood$^{4/2}$)-go^{0}-ANIM.PL.SBJ^{-1}
‘S ascends, goes up’

Non-past indicative:	$də^{8}$=k^{5}-$\boldsymbol{a^{4}}$-qut^{0}	‘she goes up’
Past indicative:	da^{8}=k^{5}-$\boldsymbol{o^{4}}$-$\boldsymbol{n^{2}}$-qut^{0}	‘she went up’
Imperative:	k^{5}-$\boldsymbol{a^{4}}$-$\boldsymbol{n^{2}}$-qut^{0}	‘go up!’

(20) Tense-mood class ***a/ol*** (stems with ***a/o⁴*** and ***l²***)

Example stem: S^{8}-O^{6}-k^{5}-$(a/o^{4}l^{2})$-do^{0}-$S_{PL}{}^{-1}$
SBJ^{8}-toward^{5}-(tense/mood$^{4/2}$)-look^{0}-ANIM.PL.SBJ^{-1}
‘S looks, stares at O’

Non-past indicative:	d^{8}-a^{6}-k^{5}-$\boldsymbol{a^{4}}$-do^{0}	‘he looks at him’
Past indicative:	d^{8}-a^{6}-k^{5}-$\boldsymbol{o^{4}}$-$\boldsymbol{l^{2}}$-do^{0}	‘he looked at him’
Imperative:	a^{6}-k^{5}-$\boldsymbol{a^{4}}$-$\boldsymbol{l^{2}}$-do^{0}	‘look at him!’

(21) Tense-mood class ***a/o*** (stems with ***a⁴*** becoming past-tense ***o⁴*** and no P2 indicated)

Example stem: $nanbed^{7}$-S^{6}-k^{5}-(a/o^{4})-$qan{\sim}qon^{0}$
bread.make^{7}-SBJ^{6}-TC^{5}-(tense/mood$^{4/2}$)-INCEPT^{0}
‘S starts making bread’

Non-past indicative:	$nanbed^{7}$-i^{6}-k^{5}-$\boldsymbol{a^{4}}$-qan^{0}	‘she starts making bread’
Past indicative:	$nanbed^{7}$-i^{6}-k^{5}-$\boldsymbol{o^{4}}$-qon^{0}	‘she started making bread
Imperative:	$nanbed^{7}$-ku^{6}-k^{5}-$\boldsymbol{a^{4}}$-qan^{0}	‘start making bread!’

A few stems contain j^2, which appears to be derived morphophonemically from n^2 and sometimes alternates with it (22a). And q^2 appears in the past tense of a few stems meaning 'kill' (22b)[4]:

(22) Irregular past-tense marker shapes

a. *d⁸-at⁷-b³-**ij²**-aq⁰ ~ d⁸-at⁷-b³-**in²**-daq⁰*
1SBJ⁸-down⁷-3INAN.OBJ³-**PAST²**-put⁰
'I poured it.'

b. *d⁸-a⁴-**q²**-ej⁰* 'I killed him'
1SBJ⁸-3MASC.SG.OBJ⁴-**PAST²**-kill⁰
'I killed him.'

The unpredictable appearance of j^2 or q^2 must be listed in the stem formula, just as the distribution of n^2- and l^2- is no longer fully predictable based on lexical aspect.

There are two key points to summarize regarding tense-mood marking. First, from the perspective of synthetic inflectional morphology, Ket has only three tense-mood forms – past indicative, non-past-indicative, and imperative – despite the template's formal complexity. Second, the position of tense-mood inflections in the template is predictable (P4 + P2), but the inflection shapes themselves are idiosyncratic, forming five productive tense-mood classes. These cannot properly be called 'conjugations' because they minimally interact with the agreement morphology.

2.2. Agreement marking

The situation with Ket agreement is the mirror opposite of the tense-mood system. Tense-mood inflections have lexically unpredictable forms that occupy predictable positions. Agreement morpheme shapes are predictable, but their position classes are not and must be specified in the verb's lexical entry.

Ket verb-internal subject and object markers distinguish agreement along the following three parameters. First, there is a distinction in person (1st, 2nd or 3rd) that

[4] Historically, past-tense *q* (< **qo*) was conjugation marker that alternated with **si* in position P3, which is why the P4 agreement markers *-a* and *-aŋ* do not labialize preceding it. However, because *q* fulfills the same past-tense function as l^2 and n^2 in Modern Ket, it is best treated as occupying P2.

encompasses all subjects and objects. Second, there is a distinction between singular and plural. Finally, third person distinguishes noun class. Person, number and class agreement interact in the following way: plural is distinguished for all animate-class subjects and objects. The noun class of third person subjects and objects interacts with number to create four subcategories: feminine singular animate, masculine singular animate, plural animate, and inanimate (normally without distinguishing singular from plural). Table 7 summarizes these functions across the position classes where they are expressed:

position class	P8 sbj person/class	P6 sbj or obj person/class/ number	P4 sbj or obj 3p anim-class	P3 sbj or obj 3p inan-class	P1 sbj or obj person/ number	P-1 anim-class plural suffix
1st sg.	*di- ~ d-*	*ba- ~ bo-*			*di-*	
1st pl.	*d=*	*dəŋ-*			*daŋ-*	*-n*
2nd sg.	*ku- ~ k-*	*ku-*			*ku-*	
2nd pl.	*k=*	*kəŋ-*			*kaŋ-*	*-n*
3rd inan-class	*də- ~ da=*	*Ø- ~ u- ~ bu-*		*b-*	*a-*	
3rd fem-class		*i- ~ u- ~ bu-*	*i- (d)id- ~ (d)it*			
3rd masc-class	*du- ~ d- d=*	*a- ~ o- ~ bu-*	*a- ~ o-*			
3rd anim-class pl.		*aŋ- ~ oŋ- ~ bu-*	*aŋ- ~ oŋ-*		*aŋ-*	*-n*

Table 7.

Position, form and agreement functions of Ket subject and object markers

Although six position classes out of ten are involved in agreement, Ket has no morphologically trivalent verbs. In terms of verb-internal agreement morphology, Ket verb stems are avalent, monovalent or bivalent (Nefedov, Malchukov and Vajda 2010 and 2011).

(23) Examples of avalent (a), monovalent (b), and bivalent (c) Ket verb stems

a. *sil*^{7}*-d*^{5}*-(a/o*^{4}*)-b*^{3}*-qan*^{0} ‘summer begins’
sil^{7}*-d*^{5}*-o*^{4}*-b*^{3}*-qan*^{0} (pronounced in Southern Ket as *sildovon*)
summer^{7}-TC^{5}-PAST^{4}-TC^{3}-INCEPT^{0}
‘Summer began.’

b. ***S*^{8}**=*sil*^{7}*-h*^{5}*-(a/o*^{4}*n*^{2}*)-aq*^{0}*-S*$_{PL}$$^{-1}$ ‘S spends (one) summer’
***d*^{8}**=*sil*^{7}*-h*^{5}*-o*^{4}*-n*^{2}*-aq*^{0} (pronounced in Southern Ket as *silunaq*)
1SBJ^{8}-summer^{7}-TC^{5}-PAST$^{4/2}$-spend^{0}
‘I spent a summer.’

c. ***S*^{8}**=*silaqŋ*^{7}*-q*^{5}*-(a/o*^{4}*l*^{2}*)-**O*$^{4/3/1}$*-da*^{0}*-S*$_{PL}$$^{-1}$ ‘S makes / lets O spend the summers’
***d*^{8}**=*silaqŋ*^{7}*-q*^{5}*-**o*^{4}*-l*^{2}*-da*^{0}*-n*$^{-1}$
1SBJ^{8}=spend.summer^{7}-TC^{5}-**3MASC.SG.OBJ^{4}**-PAST^{2}-TRANS.MOM^{0}-SANIM.PL^{-1}
‘They used to make him spend the summer (somewhere).’

Nor do agreement positions predictably distinguish semantic roles, in contrast to interpretations by Belimov (1991), Reshetnikov and Starostin (1995), and Butorin (1995). Modern Ket has five productive intransitive classes, each with a different position-class expression of subject agreement. These patterns were called ‘actant conjugations’ in Vajda (2001, 2003, 2004), who gave them names such as ‘absolutive conjugation’ or ‘inactive conjugation’, which created the misleading impression that agreement class membership was based on functional differences in valency. Here they are called simply ‘agreement classes’, as they resemble Indo-European conjugation classes, except that Ket involves differences in position class whereas Indo-European agreement suffixes occupy the same position in every stem. Some Ket agreement classes involve multi-site subject marking, while others mark the subject only once.

(24) The five productive Ket intransitive agreement classes

i. Intransitive class I (*v1*): (animate-class subject person^{8} and plural^{-1})

a. *d*^{8}*-es*^{7}*-o*^{4}*-l*^{2}*-ij*^{0}*-n*$^{-1}$
3SBJ^{8}-into.open.space^{7}-PAST$^{4/2}$-call^{0}-ANIM.PL.SBJ^{-1}
‘They shouted.’

b. *du*^{8}*-n*^{2}*-qo*^{0}*-n*$^{-1}$
3SBJ^{8}-PAST^{2}-die^{0}-ANIM.PL.SBJ^{-1}
‘They died.’

c. *d[8]-ikda[7]-k[5]-l[2]-aq[0]-n[-1]*
3SBJ[8]-forest.to.river[7]-TC[5]-PAST[2]-go[0]-ANIM.PL.SBJ[-1]
'They made a round trip to the riverbank (lasting several days).'

d. *d[8]=kət[7]-h[5]-o[4]-n[2]-aq[0]-n[-1]*
3SBJ[8]=winter[7]-TC[5]-PAST[4/2]-pass[0]-ANIM.PL.SBJ[-1]
'They spent the winter.'

ii. Intransitive class II (*v2*): (any subject person, class, number marked in P6)

a. *ba[6]-k[5]-s[4]-sal[0]*
1SG.SBJ[6]-TC[5]-PRES[4]-spend.night[0]
'I spent the night.'

b. *ɨntip[7]-bo[6]-l[2]-a[1]-bed[0]*
puppy[7]-1SG.SBJ[6]-PAST[2]-RESULT[1]-possess[0]
'I had a puppy.'

c. *ɨs[7]-aŋ[6]-k[5]-o[4]-qan[0]*
row[7]-3ANIM.PL.SBJ[6]-TC[5]-PAST[4]-INCEPT[0]
'They started rowing.'

iii. Intransitive class III (*v3*): (multi-site subject marking in P8, P6, and P-1)

a. *d[8]-ikda[7]-bu[6]-t[5]-l[2]-aq[0]-n[-1]*
3SBJ[8]-to.river[7]-3SBJ[6]-TC[5]-PAST[2]-go[0]-ANIM.PL.SBJ[-1]
'They made a quick round trip to the riverbank.'

b. *d[8]=bu[6]-t[5]-o[4]-l[2]-ok[0]-n[-1]*
3SBJ[8]=3SBJ[6]-TC[5]-PAST[4/2]-jump[0]-ANIM.PL.SBJ[-1]
'They shuddered.'

c. *d[8]=bu[6]-ŋ[5]-l[2]-uqo[0]-n[-1]*
3SBJ[8]=3SBJ[6]-TC[5]-PAST[2]-look[0]-ANIM.PL.SBJ[-1]
'They looked (searched for something).'

iv. Intransitive class IV (*v4*): (multi-site marking of anim.-class subjects in P8 + P1)
(multi-site marking of inan.-class subjects in P3 + P1)

a. *d[8]-ikda[7]-o[4]-n[2]-aŋ[1]-daq[0]*
3SBJ[8]-to.river[7]-PAST[2]-3ANIM.PL.SBJ[1]-go[0]
'They went to the river (and stayed there).'

b. *d*[8]*=d*[5]*-o*[4]*-l*[2]*-aŋ*[1]*-dun*[0]
3SBJ[8]=TC[5]-PAST[2]-3ANIM.PL.SBJ[1]-shout[0]
'They gave a shout.'

c. *d*[8]*-o*[4]*-l*[2]*-aŋ*[1]*-tij*[0]
3SBJ[8]-PAST[4/2]-3ANIM.PL.SBJ[1]-grow[0]
'They grew up.'

v. Intransitive class V (*v5*): (3animate subject[4], 1st or 2nd person subject[1])
(3inanimate subject[3])

a. *i*[7]*-di*[1]*-bed*[0]
day[7]-1SG.SBJ[1]-spend[0]
'I spend the day.'

b. *tip*[7]*-il*[2]*-di*[1]*-bed*[0]
dog[7]-PAST[2]-1PL.SBJ[1]-possess[0]
'I had a dog.'

c. *us*[7]*-oŋo*[4]*-n*[2]*-den*[0]
sleep[7]-3ANIM.PL.SBJ[4]-PAST[2]-event.occurs[0]
'They fell asleep.'

Stems that use intransitive class V (*v5*) to mark animate-class subjects are logically precluded from having inanimate-class subjects. Inanimate-class marking in this class occurs in stems, that mark their animate-class subjects according to the pattern of intransitive class I (*v1*).

(25) Agreement class dichotomy between animate-class and inanimate-class subjects

i. Intransitive class I (*v1*): (animate-class subject person[8] and plural[-1])

d[8]*-il*[2]*-loqŋ*[0]*-n*[-1]
3SBJ[8]-PAST[2]-shake[0]-ANIM.PL.SBJ[-1]
'They (animate-class) shook.'

ii. Intransitive class V (*v1*): (inanimate-class subject[3])

b[3]*-il*[2]*-loqŋ*[0]
3INAN.SBJ[3]-PAST[2]-shake[0]-ANIM.PL.SBJ[-1]
'They (inanimate-class things) shook.' or 'It shook.'

Intransitive class V stems with inanimate-class subjects could therefore instead be interpreted as belonging to inanimate-class I. This, in fact, would corresponds to their historical development and would also parallel the class-based dichotomy found in intransitive class IV stems. (Stems belonging to intransitive classes II or III have no positional dichotomy between their animate-class and inanimate-class subjects.) Under such an interpretation, intransitive class V would contain only animate-class verbs expressing static conditions or changes of state that are logically limited to animate-class participants (such as 'wake up', 'own', 'get sleepy', etc.). In any case, stems with such meanings are not restricted to intransitive class I and can be found in the other four intransitive agreement classes as well, so the semantic correlation here is only partial.

Transitive stems are divided into three productive agreement classes (see Vajda 2001, 2004, 2009; Georg 2007).

(26) The three productive Ket transitive agreement classes

i. Transitive class I (*vt1*): (subject person^{8}, subject plural^{-1})
(3animate object^{4}, 3inanimate object^{3}, 1p or 2p object^{1})

a. *d^{8}=bakdeŋ^{7}-q^{5}-a^{4}-di^{1}-da^{0}-n^{-1}*
3SBJ^{8}=pull^{7}-INCEPT^{5}-PRES^{4}-1SG.OBJ^{1}-ITER.TRANS^{0}-ANIM.PL.SBJ^{-1}
'They frequently pull me.' or 'They frequently start pulling me.'

b. *d^{8}-in^{2}-di^{1}-tek^{0}-n^{-1}*
3SBJ^{8}-PAST^{2}-1SG.OBJ^{1}-hit.endwise^{0}-ANIM.PL.SBJ^{-1}
'They beat me (past tense, single event).'

ii. Transitive class II (*vt2*): (subject person^{8}, subject plural^{-1})
(person, class and number of any object marked in P6)

a. *d^{8}=bakdeŋ^{7}-bo^{6}-k^{5}-a^{4}-bed^{0}-n^{-1}*
3SBJ^{8}=pull^{7}-1SG.OBJ^{6}-TC^{5}-PRES^{4}-ITER.TRANS^{0}-ANIM.PL.SBJ^{-1}
'They frequently pull me.' or 'They are pulling me (process).'

b. *d^{8}=don^{7}-ba^{6}-h^{5}-il^{2}-tek^{0}-n^{-1}*
3SBJ^{8}=knife^{7}-1SG.OBJ^{6}-area^{5}-PAST^{2}-hit.endwise^{0}-ANIM.PL.SBJ^{-1}
'They stabbed me (once).'

iii. Transitive class III (*vt3*): (multi-site subject marking in P8, P6, and P-1)
(3animate object^{4}, 3inanimate object^{3}, 1p or 2p object^{1})

a. *d^{8}-aka^{7}-bu^{6}-k^{5}-di^{1}-qos^{0}-n^{-1}*
3SBJ^{8}-river.to.forest^{7}-3SBJ^{6}-TC^{5}-1SG.OBJ^{1}-take^{0}-ANIM.PL.SBJ^{-1}
'They take me into the forest.'

b. *d^{8}=bu^{6}-t^{5}-a^{4}-b^{3}-daq^{0}-n^{-1}*
3SBJ^{8}=3SBJ^{6}-TC^{5}-PRES^{4}-3INAN.OBJ^{3}-put^{0}-ANIM.PL.SBJ^{-1}
'They pull it out.'

A handful of stems have unproductive agreement patterns (Vajda 2004: 70). The origins of Ket agreement classes – both productive and unproductive – are complex, involving factors such as reanalysis and metathesis, and generally having nothing directly to do with valency (Vajda 2013). Vajda (in press 2) implicates the partial phonological erosion of the original P1 subject markers and their replacement by P8 subject markers, leading to multi-site subject marking in stems where the P1 markers happened to remain. Another factor was metathesis between agreement markers and the stem's P5 thematic consonant. In some stems, object markers that originally occupied position P6 metathesized to the right of the P5 marker, assuming positions P4, P3 or P1, depending on person and class. In other stems, subject markers that originally occupied P4, P3 or P1 metathesized into position P6. Because metathesis involved a phonological trigger (adjacent segments differing in +/- labial), the resulting agreement classes do not pair up with distinctions in semantic or syntactic valency. A final complication involved sporadic reanalysis of morphemes across position classes. For example, in verbs meaning 'gulp down', the incorporate *buŋ^{7}* 'lump' was reanalyzed as the subject marker *bu^{6}* followed by thematic consonant *ŋ^{5}*.

(27) Reanalysis of incorporate causing shift in agreement class from *vt1* (a) to *vt3* (b)

a. Original structure
d^{8}=buŋ^{7}**-t^{5}-b^{3}-il^{2}-dop^{0}*
*3SBJ^{8}=lump^{7}-down^{5}-3INAN.OBJ^{3}-PAST^{2}-injest^{0}
'He (anim.-class bird) gulped it down.'

b. Reanalyzed structure
*d^{8}=**bu^{6}-ŋ**/t^{5}-b^{3}-il^{2}-dop^{0}*
3SBJ^{8}=**3SBJ^{6}-TC**/down^{5}-3INAN.OBJ^{3}-PAST^{2}-injest^{0}
'He (anim.-class bird) gulped it down.'

c. Replacement by other agreement markers
*d^{8}=**ba^{6}-ŋ**/t^{5}-b^{3}-il^{2}-dop^{0}*
1SBJ^{8}=**1SG.SBJ^{6}-TC**/down^{5}-3INAN.OBJ^{3}-PAST^{2}-injest^{0}
'I gulped it down.'

Only some cases of multi-site subject marking arose in this way, though all instances of thematic *ŋ*^{5} appear to derive from this kind of reanalysis, in contrast to the interpretation in Vajda (2003: 75), where a futile attempt was made to assign them semantic etymologies.

To summarize, Modern Ket agreement classes, however they are called, must be specified in the stem formula or notated in parentheses afterward. Not doing so would omit a core element of the verb's lexical entry. A second key point is that the Ket verb is strongly head-marking. The subject and object are somehow distinguished from one another inside the verb form, while the verb-external subject and object noun phrases themselves are never morphologically marked.

2.3. Lexical typology

As argued by Werner (1997 and elsewhere), Ket is basically a nominative-accusative language, with certain active typological traits, such as the positional contrast between animate- and inanimate-class marking observed in many intransitive stems. Otherwise, the various strategies for marking the subject or object lack a clear semantic basis. Syntactic valency is an obligatory lexical category. Also, Ket has no labile verbs. Avalent, monovalent, and bivalent stems with logically related meanings always differ in terms of the lexical morpheme configuration in P7-P5-P0, and not simply in the presence or absence of the agreement markers themselves. There are rather few derivational techniques for raising or lowering valency (Vajda 2015) and all of them are highly constrained lexically. Noun incorporation is restricted to a handful of bases. The base *bed*^{0} 'make' freely incorporates its object, but most bases cannot incorporate nouns at all or incorporate only a few specific object nouns.

(28) Verb phrase without (a) and with (b) object incorporation

a. *bu asl d*^{8}*=b*^{3}*-il*^{2}*-bed*^{0}
he ski 3SBJ^{8}=3INAN.OBJ^{3}-PAST^{2}-make^{0}
'He made a ski.'

b. *d*^{8}*-asl*^{7}*-il*^{2}*-bed*^{0}
3SBJ^{8}-ski^{7}-PAST^{2}-make^{0}
'He made a ski.'

The bases *ted*^{0} 'hit endwise' and *kit*^{0} 'rub' can incorporate their instrument, but not their direct object (Vajda, in press 1). No other bases permit transitivity raising via instrument incorporation.

(29) Verb phrase without (a) and with (b) instrument incorporation

a. *ke'd* *tīb* *sal-as* *d^{8}-a^{6}-t^{5}-a^{4}-kit^{0}*
person dog tobacco-INSTR 3SBJ^{8}-3MASC.SG.OBJ^{6}-TC^{5}-PRES^{4}-rub^{0}
'The man rubs the dog with tobacco ~ rubs tobacco on the dog.'

b. *ke'd* *tīb* *d^{8}=sal^{7}-a^{6}-t^{5}-a^{4}-kit^{0}*
person dog 3SBJ^{8}=tobacco^{7}-3MASC.SG.OBJ^{6}-TC^{5}-PRES^{4}-rub^{0}
'The man "tobaccoes" the dog (to ward off fleas).'

A subset of intransitive motion verbs can become transitive by adding an object marker followed by the thematic consonant *k^{5}* 'with'. No other stems generate this type of comitative applicative.

(30) Simple intransitive (a) and comitative applicative (b) motion verb form

a. *de'ŋ* *d^{8}-ik^{7}-in^{2}-bes^{0}-n^{-1}*
people 3SBJ^{8}-into.open.space^{7}-PAST^{2}-pass^{0}-ANIM.PL.SBJ^{-1}
'The people arrived.'

b. *de'ŋ* *d^{8}-ik^{7}-bo^{6}-k^{5}-in^{2}-bes^{0}-n^{-1}*
people 3SBJ^{8}-into.open.space^{7}-1SG.OBJ^{6}-with^{5}-PAST^{2}-pass^{0}-ANIM.PL.SBJ^{-1}
'The people brought me.' ~ 'The people arrived with me.'

Aspect-related categories play a significant role in verb morphology. Single vs. multiple action is distinguished in many stems. Inceptivity is another prominent lexical category. Finally, the distinction between action vs. resultant state is also regularly marked by the stem morphology. The examples in (31) show several highly productive stem creation models used to express contrasts in lexical aspect:

(31) Lexically related verbs containing the action nominal *toqojiŋ* 'dry off'

a. *da^{8}=toqojiŋ^{7}-q^{5}-i^{4}-b^{3}-t^{0}*
3FEM.SG.SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-3INAM.OBJ^{3}-MOM.TRANS^{0}
'She begins drying it off (once).' ~ 'She dries it off (once)'

b. *da^{8}=toqojiŋ^{7}-q^{5}-s^{4}-a^{1}-tn^{0}*
3FEM.SG.SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-3SG.SBJ^{1}-MOM.INTRANS^{0}
'She begins drying off (once).' ~ 'She dries off (once).'

c. *da*^{8}*=toqojiŋ*^{7}*-q*^{5}*-s*^{4}*-qut*^{0}
3FEM.SG.SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-SG.SBJ.RESULT^{0}
'She is dried off (stative-resultative form).'

d. *d*^{8}*=toqojiŋ*^{7}*-q*^{5}*-s*^{4}*-dam*^{0}*-in*$^{-1}$
3SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-PL.SBJ.RESULT^{0}-ANIM.PL.SBJ^{-1}
'They (animate-class) are dried off (stative-resultative form).'

e. *da*^{8}*=toqojiŋ*^{7}*-q*^{5}*-a*^{4}*-b*^{3}*-da*^{0}
3FEM.SG.SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-3INAM.OBJ^{3}-ITER.TRANS^{0}
'She dries it off (many times).'

e. *da*^{8}*=toqojiŋ*^{7}*-q*^{5}*-s*^{4}*-a*^{1}*-dij*^{0}
3FEM.SG.SBJ^{8}=dry^{7}-TC^{5}-PRES^{4}-3SG.SBJ^{1}-ITER.INTRANS^{0}
'She dries off (many times).'

These patterns were originally inceptives based on the root **daq* 'put' and thematic *q*^{5} 'inside' that literally meant 'S puts O into the action of drying', 'S is put into the action of drying', etc. When action nominals were inserted into Ket stems to generate root-initial forms, the P0 base **daq* 'put' became semantically secondary. It also merged phonologically with the circum-root intransitive/resultative prefix **jə-* and suffix **-ej* ~ **-ŋ* to form the aspect-related bases *t* ~ *tn* ~ *qut* ~ *dam*^{5} ~ *da* ~ *dij* seen in (31a-e). Thematic *q*^{5} in these stems also grammaticalized into a opaque stem element: 'inside' → 'inceptive' → 'thematic consonant', as the patterns in question evolved into highly productive means of distinguishing single vs. repeated action and action vs. resultant state. Today, the inceptive meaning remains functional only in certain single-action stems.

The Ket verbal lexicon is atomistic in that stems are formally connected with other stems primarily through the template's position class structure. There is no

[5] The base *-dam*^{0} + animate-class plural agreement suffix *-in*$^{-1}$ has merged to *-damin*^{0} in forms with inanimate-class plural subjects, apparently through reanalysis of *-in*$^{-1}$ as part of the plural base *-dam*^{0} in contrast with the singular base *-qut*^{0}:

(a) inanimate-class plural resultative form
toqojiŋ^{7}*-q*^{5}*-i*^{4}*-b*^{3}*-damin*^{0}
dry^{7}-TC^{5}-PRES^{4}-3INAN.SBJ^{3}-PL.SBJ.RESULT^{0}
'They (inanimate-class) are dried off.'

(b) inanimate-class singular resultative form
toqojiŋ^{7}*-q*^{5}*-i*^{4}*-b*^{3}*-qut*^{0}
dry^{7}-TC^{5}-PRES^{4}-3INAN.SBJ^{3}-SG.SBJ.RESULT^{0}
'It is dried off.'

Outside this pattern, no other verbs with inanimate-class plural subjects contain the suffixal element *-in*.

straightforward affixal derivation of the type commonly found in the more familiar families of Eurasia. Instead, we might speak of 'formulaic derivation', where certain stem formulas, such as those illustrated in (31), share a predictable, schematized relationship.

Conclusion

The discussion has shown that nearly every typological generalization about the Ket verb must be qualified in some way. Alignment is nominative-accusative, but animacy, volition and activeness also play a role in agreement, as would be more typical of a language with active alignment. The verb's formal structure can be interpreted in typologically contradictory ways, as well. The original Yeniseian verb was strongly prefixing, yet Modern Ket has partly adapted the prefixing template into a suffixing arrangement under areal pressure. Under the same pressure, the position P8 subject prefixes have become special clitics in most verb forms but not in all. What has remained steadfast is the position class technique of verb formation itself. The straightjacket of the template even seems to have dictated what changed and how it changed. Metathesis and reanalysis, which are marginal factors in regular affixal morphology, have figured as leading techniques of innovation in Ket templatic verb morphology. The resulting dissonance of form and function in the Ket verb is thus an artifact of areal pressure combined with the conservatism of the inherited template.

Defining a Ket verb stem and listing it in a dictionary poses logistic challenges not encountered with most other Eurasian languages. Verbs in Indo-European, Uralic, Turkic, Mongolic, and Turkic languages are easily treated lexicographically as a word-like stem followed by a slot (or grid of slots) for grammatical endings. This familiar structure can be conveniently alphabetized in a dictionary using a citation form like the infinitive followed by information about conjugation class and any irregular morphological changes, or by listing the bare stem itself, which resembles a pared-down word-form equally amenable to alphabetical ordering. Ket finite verbs, by contrast, must be listed as positional formulas that specify disjunct lexical morphemes, the location of subject and object agreement markers, and the form of tense-mood affixes. Omitting any of this information from the lexical entry of a Ket verb renders the description incomplete and ultimately unusable.

Abbreviations

ANIM	animate-class
DETRANS	detransitive
FEM	feminine-class
IMP	imperative
IMPERF	imperfective aspect
INTRANS	intransitive
MASC	masculine-class
MOM	single action
OBJ	object
PERF	perfective aspect
PL	plural
PRES	present tense
RESULT	resultative
SBJ	subject
SG	singular
TC	thematic consonant
TRANS	transitive

References

Belimov, Eduard. 1991. *Ketskij sintaksis: situatsija, propozitsija, predlozhenie*. Novosibirsk: Nauka.

Butorin, Sergej. 1995. *Opisanie morfologicheskoj struktury finitnoj glagol'noj slovoformy ketskogo jazyka s ispol'zovaniem metodiki porjadkovogo chlenenija.* Unpublished Candidate Degree Dissertation. Novosibirsk.

Dul'zon, Andrej. 1968. *Ketskij jazyk*. Tomsk: Tomsk State University.

Kotorova, Elizaveta and Nefedov, Andrej (eds). 2015. *Comprehensive dictionary of Ket with Russian, German and English translations*. Munich: Lincom Europa.

Krejnovich, E. A. 1968. *Glagol ketskogo jazyka*. Leningrad: Nauka.

Nefedov, Andrey; Malchukov Andrej and Vajda, Edward. 2010. Ditransitive constructions in Ket. In: Malchukov, Andrej; Haspelmath, Martin and Comrie, Bernard (eds.) *Studies in ditransitive constructions: a comparative handbook*. Berlin and New York: Mouton de Gruyter. 352–381.

Nefedov, Andrey; Malchukov, Andrej and Vajda, Edward. 2011. Impersonal constructions in Ket. In: Malchukov, Andrej and Siewierska, Anna (eds.)

Impersonal constructions: a cross-linguistic perspective. Amsterdam and New York: John Benjamins. 439–458.

Reshetnikov, Kirill and Starostin, Georgij. 1995. Struktura ketskoj glagol'noj slovoformy. In: Starostin, Sergej (ed.) *Ketskij sbornik* 4. Moskva: Vostochnaja Literatura. 7–121.

Vajda, Edward. 2001. The role of position class in Ket verb morphophonology. *Word* 52/3. 369–436.

Vajda, Edward. 2003. Ket verb structure in typological perspective. *Sprachtypologie* und *Universalienforschung* 56: 1/2. 55–92. Berlin: Akademie Verlag.

Vajda, Edward. 2004. *Ket*. Munich: Lincom. (Languages of the World/Materials 204)

Vajda, Edward. 2009. Loanwords in Ket. In: Haspelmath, Martin and Tadmor, Uri (eds.) *Loanwords in the World's languages: a comparative handbook*. Berlin: Mouton de Gruyter. 471–494.

Vajda, Edward. 2013. Metathesis and reanalysis in Ket. *Tomsk journal of linguistics and anthropology* 1/1. 14–26.

Vajda, Edward. 2015. Valency properties of the Ket verb clause. In: Malchukov, Andrej and Comrie, Bernard (eds.) *Valency classes in the world's languages*. Berlin and New York: Mouton de Gruyter. 630–668.

Vajda, Edward. (in press l). Ket incorporation. In: Fortescue, Michael; Mithun, Marianne and Evans, Nicholas (eds.) *Handbook of Polysynthesis*. Oxford: Oxford University Press.

Vajda, Edward. (in press 2). Patterns of innovation and retention in templatic polysynthesis. Fortescue, Michael; Mithun, Marianne and Evans, Nicholas (eds.) *Handbook of Polysynthesis*. Oxford: Oxford University Press.

Vajda, Edward & Heinrich Werner. (in preparation). *Etymological dictionary of the Yeniseian Languages*.

Werner, Heinrich. 1997. *Die ketische Sprache.* Wiesbaden: Harrassowitz.

Sayan Turkic reindeer terminology

Elisabetta Ragagnin

Freie Universität Berlin

1. Sayan Turkic languages and peoples

Sayan Turkic is the technical term used in Turcology to refer to the South-Siberian branch of the Turkic language family, which includes standard Tuvan, Tofan and various related languages spoken by small communities in neighbouring Mongolia, China and Buryatia Sayan Turkic can be further divided into two groups: Steppe Sayan Turkic and Taiga Sayan Turkic. To the former belong Standard Tuvan and its dialects together with Altay-Sayan varieties spoken in China and western Mongolia, and Uyghur-Uriankhay (Tuhan) of East Khövsgöl. The traditional lifestyle of Taiga Sayan Turkic peoples used to be based on low-land i.e. steppe pastoralism, characterized by herding of sheep, goats, cattle (cows and yaks), horses and camels.

The Taiga Sayan Turkic branch includes Tofan, the Toju variety of Tuvan and some varieties of the Tere-Khöl area as well as Soyot of Buryatia and Dukhan of Mongolia's northern Khövsgöl Aimag. Reindeer breeding and hunting has characterized the lifestyle of these groups until recently.[1] During the last decades reindeer herding has dramatically decreased in the Taiga-Sayan area - see, for instance, Donahoe and Plumley (2003). Presently, reindeer herding is best maintained among the Dukhan people.

2. The Dukhan people

Secluded in the northernmost regions of Mongolia's Khövsgöl region – bordering in the northeast with Buryatia and in the west with the Tuvan republic – the Dukhan people total approximately 500 people and are divided into two main groups: those of the "West Taiga" (*barïïn dayga*) originate from Tere Khöl, whereas those of the "East Taiga" (*ǰüün dayga*) came from Toju, both of them regions in Tuva. Presently,

[1] For more on the taiga vs. steppe division, though with slight differences from the view presented here, see Žukovskaja et al. (2002).

around 32 Dukhan families are reindeer herders in the surrounding taiga areas, on the southern slopes of the Sayan mountains. They follow what has been called the Sayan-type of reindeer breeding, characterized by small-size herds of reindeer used as pack and riding animals and as a source of dairy products. (For more on the Sayan-type of reindeer herding, see Vainshtein 1980). Hunting used to be an important part of the Dukhan economy. However, the Mongolian government has recently banned hunting and fishing and, in order to balance the impact of these proscriptions, the Dukhan families were granted dwelling in the taiga and a state pension for tending to reindeer, calculated on the basis of the number of family members. The remaining Dukhan families have settled down in the village of Tsagaan Nuur and in the neighbouring riverside areas, abandoning reindeer breeding. Some families, however, regularly return to the taiga in the summer months and tend to reindeer. Although the Dukhan people identify themselves as *du^{h}ha*, in Mongolia they are generally called *Tsaatan*, a rather derogatory term meaning 'those who have reindeer', stressing in this way the fact that they are not like Mongolian herders. Recently the more neutral Mongolian term *tsaačin* 'reindeer herders' has been introduced to refer to them.[2]

Nowadays, Dukhan is actively spoken by the older generation, that is, by speakers older than 40. Younger Dukhans communicate in Darkhat-Mongolian,[3] although they possess a passive knowledge of Dukhan. Furthermore, language shift is more acute in Tsagaan Nuur and riverside areas, where most of the households have already completely switched to Darkhat-Mongolian. For a linguistic description of Dukhan, see Ragagnin (2011).

[2] In the available published materials, Dukhans have been designated by several other names such as "Urianxay", "Taiga Urianxay", "Taigïn Irged" 'peoples of the taiga', "Oin Irged" 'peoples of the forest' and "Soiot" (Badamxatan 1962: 3). Dukhans do not call themselves Uyghur, as claimed in some publications.

[3] The general view among scholars is that the Darkhat people are of Turkic origin and that their language and customs have become Mongol in the past few centuries. For a short survey of Darkhat grammatical features, see Sanžeev (1931) and Gáspár (2006). On Darkhat Mongolian's substrate of Turkic features, see Ragagnin (2012b).

Below I shall alphabetically list and comment on Dukhan reindeer terminology,[4] a very unique part of Dukhan lexicon.[5] The data is supported with materials from the other neighbouring Taiga Sayan Turkic varieties:

(1) *anhay* 'new-born reindeer calf'; cf. Toju-Tuvan *aʿniy* 'reindeer calf' (Čadamba 1974: 63), Tofan *anhay* 'new-born reindeer calf' (Rassadin 2014: 52) and Soyot *aʿnay ~ aʿnhay* 'reindeer calf up to year of age' (Rassadin 2006: 22). Steppe Sayan Turkic displays cognates referring to the young of other animals, e.g. Tuvan *aʿnay* 'young offspring of a goat or mountain goat' (Monguš 2003: 130b). Etymologically, *anhay* is rather obscure. It could be traced back to Turkic *ana* 'mother' augmented with the hypocoristic suffix *-KAy*, thus meaning 'mommy'. Similar expressions are used, for instance, in Turkish (see Ragagnin 2012a: 135–136 for details). For other views, see Rassadin (2014: 54).

(2) *bogana ~ mogana* 'male reindeer castrated at an advanced age'; cf. Tere-Khöl Tuvan *boxana* 'gelded reindeer' (Kuular and Suvandi 2011a: 165); in Toju-Tuvan a cognate of this term occurs in the compound *bogana čarï* denoting an older breeding reindeer. Etymologically, *bogana ~ mogana* may represent a nominal formation built with the Mongolic suffix *-gAnA*, used for names of plants and animals (Poppe 1954: 41), or a deverbal formation involving the Mongolic suffix *-gAn* (Poppe 1954: 45) or the corresponding denominal suffix *-gAn* deriving zoological and botanical names in Old Turkic (Erdal 1991: 87). Moreover, *bogana ~ mogana* may be related to Gagauz *bobana* 'seven/eight-year-old sheep' (Ščerbak 1961: 153). This claim, however, needs further scrutiny.

(3) *döŋhʉr* 'young reindeer buck'; cf. Toju Tuvan *döʿnggür* 'without antlers, with dropped antlers, one of the terms in use for a breeding reindeer' (Čadamba 1974: 64), Tofan *dönggür* 'male domesticated uncastrated ridable reindeer in its third year and first mating season but not ready for mating' (Harrison 2010: 57) and Soyot *döŋkir* 'reindeer buck' (Rassadin 2006: 46). In Tuvan, the primary meaning of *döŋgür* is 'antlerless, bold' (Tenišev 1968: 178a); also cf. Altay Turkic *toŋkur* 'komolyj, bezxvostyj, obrezannyj, kucyj; derevo bez veršiny i vetvej' (Verbickij 2005: 364). Also, see Tatarincev (2002: 235–236) and Monguš (2003: 495b) in this

[4] Some Dukhan reindeer terms are listed in Badamxatan (1962: 9), Somfai-Kara (1998: 18–19), Kuular and Suvandi (2011b) and Ragagnin (2010 and 2012a).

[5] This scarcely documented special lexicon is rapidly getting lost. In my fieldwork sojourns, I could personally see how fuzzy it is among many Dukhans.

respect. It is a moot question whether this term may have originally been applied to a young reindeer buck after its first (autumnal) antlers fell.

(4) *eʰsɨrɨk* ‘new-born reindeer calf’; cf. Toju-Tuvan *eʿzirik* ‘affectionate name for reindeer calf’ (Čadamba 1974: 63) and Soyot *eʿsirik* ‘new-born reindeer calf’ (Rassadin 2006: 85, 204). Steppe Sayan Turkic, on the other hand, displays corresponding items referring to the young of other animals, cf. *ezirik* (*eʿzirik*) ‘goatling, (kid), fawn’ (Tenišev 1968: 608b, Dorlig and Dadar-ool 1994: 242a). Etymologically, *eʰsɨrɨk* is a transparent Turkic agent formed from the verbal stem *ehsɨr-* ‘to get drunk’ (cf. ED 251: *esür-* ‘to be, or become, drunk, intoxicated) and literally means ‘drunkard’. This denomination is most probably is based on the fact that the new-born baby reindeer tumbles like a drunkard. Moreover, it likely belongs to the set of taboo names applied to the young of both humans and animals in order to protect them from evil spirits. It is quite unlikely that evil spirits would take away a drunkard. In addition, from *eʰsɨrɨk* the verbal stem *eʰsɨrɨkte-* ‘to calf/to fowl’ is formed.

(5) *eʰter* ‘breeding reindeer’; cf. Toju Tuvan *eʾder* (Čadamba 1974: 63-64) and Soyot *eʾter* ~ *eʾtər* (Rassadin 2006: 205) ‘reindeer buck’. Cognate forms are documented in neighbouring Buryat Sayan dialects as well: *eteer* ‘breeding reindeer’ (Rassadin 1996: 149). Etymologically, *eʰter* literally ‘screamer in rut’, is a Turkic participial form derived from *eʰt-* ‘to scream in rut’ (cf. ED 39-40 and Röhrborn 2010: 200). Evidently, ‘screaming in rut’ is the most important characteristics of a male reindeer on heat. Dukhan *eʰter* also occurs in the expression *ulɨγ eʰter*, literally ‘big *eʰter’* with reference to an “experienced” reindeer buck.

(6) *guuday* ‘castrated reindeer buck’; cf. Soyot *quuday* ‘domesticated three-year-old reindeer buck’, Tere-Khöl Tuvan *kuuday* ‘small/young male reindeer (general term)’ (Seren 2006: 82) and Tofan *kuuday* ‘reindeer buck of about 2-3 years of age’ (Rassadin 1995: 33a). Ščerbak (1961: 91-92) and Tatarincev (2004: 327), derived the term *kuuday* from Turkic *kuu* ‘grey’ and *day* ‘foal, young horse’.

(7) *hokkaš* ‘young reindeer up to one year of age’; cf. Tere-Khöl Tuvan *xokaš* ‘reindeer calf below one year of age’ (Seren 2006: 81), Tofan *hokkaš* ‘one-year-old reindeer calf’ (Rassadin 1971: 190) and Soyot *hoqaš* ~ *hokkaš* ‘one-year-old reindeer calf (in its second year)’ (Rassadin 2006: 85). Etymologically, *hokkaš* is rather obscure. Arguably it goes back to the diminutive formation *kuškaš* ‘small bird’ from *kuš* ‘bird’ through phonological distortion, not uncommon in taboo

names. Note in this respect that a structurally similar lexeme is documented in Sarig Yugur *go^{h}qaš* ‘small bird’ (Nugteren and Roos 2006: 110).

(8) *hospayak* ‘new-born reindeer fawn rejected by its own mother’; no cognates are documented so far in the rest of Taiga Sayan Turkic. Morphologically, *hospayak* is a nominal form derived from the verbal stem *hos-* ‘to refuse animal’s babies’. Standard Tuvan employs the form *xosturgan* (*xos-* ‘to refuse animal babies’ + causative + part) to characterize a young animal refused by its own mother, e.g. *xosturgan xuragan* ‘rejected lamb’[6].

(9) *ǰarɨ* ‘“calm” riding and packing reindeer older than four years of age’; cf. Tofan *ǰarə*, Soyot *čarï* ‘riding and packing reindeer’ (Rassadin 1971: 194; 2006: 153), Tere-Khöl and Toju Tuvan *čarï* ‘castrated reindeer’ (Seren 2006: 82). Interestingly, Steppe Sayan Turkic varieties show a rather different picture. The standard Tuvan cognate *čarï* refers to a breeding male reindeer (Tenišev 1968: 520a) and in the Uyghur Uriankhay Sayan-Turkic variety of Eastern Khövsgöl *ǰarɨ* is the only existing term meaning ‘reindeer’ (Ragagnin 2010: 142). Note, in this respect, that Sayan western Buryat dialects, which show several Turkic features, displays the cognate *zari* denoting both a gelded reindeer (older than 4 years of age) and a (breeding) reindeer (Čeremisov 1973: 251b). Cognates of this term are quite widespread in Siberia. Yeniseian terms such as for instance Yugh *^{4}sɛ^{h}:r* ‘reindeer’ and Ket *s’ɛl’* ‘reindeer’ may be related items (Werner 2002: 183); also see Khabtagaeva (2015: 116). Historically, a cognate of *ǰarɨ* is documented in Rašīd-ud-Dīn’s Jāmiʿ al-tawārīkh (Compendium of Chronicles). In section 107 dealing with the “forest” Uriangqat tribe, the Ilkhanid Persian historian wrote: “They had no cattle or sheep but raised and caught instead mountain oxen, mountain rams, and *jür* (antelope), which is like a mountain sheep, which they milked and drank” (Thackston 2012: 42, §107). Also see the information supplied by Marco Polo’s 13th century travelogue concerning reindeer herding nomads in the neighbouring Bargu area (cf. Ragagnin 2015). Furthermore, Dukhan *ǰarɨ* occurs in the nominal compound *e^{h}ter-ǰarɨ* which designates a breeding reindeer. Formal and semantic cognates are well documented in other Taiga Sayan Turkic varieties, e.g. Tofan *e^{ʔ}ter ǰarï* (Rassadin 1996: 149-150) and Soyot *e^{ʔ}ter čarï* (Rassadin 2006: 153). In these

[6] I wish to thank my colleague Choduraa Tumat for providing me with this example.

forms, *ǰarɨ* is clearly used as a species collective denomination.[7] The compound noun can thus be interpreted as 'the category of rutting reindeer'.

(10) *iβɨ* general term referring to 'reindeer' (rangifer tarandus sibiricus); close formal and semantic cognates are documented throughout Sayan Turkic, cf. Tofan *ibi* 'reindeer' (Rassadin 1995: 24), Soyot *ivi* 'id.' (Rassadin 2006: 50), Tuvan *ivi* 'id.' (Tenišev 1968: 200). Etymologically, *iβɨ* is possibly related with *iwiq* 'a she-antelope which frequents stony tracks and deserts', recorded in Maḥmūd Al-Kāšγarī's encyclopaedic work Dīwānu l-Luġat al-Turk (Compendium of the languages of the Turks) and glossed with Arabic *ẓabya* (Dankoff and Kelly 1982: 108; also cf. Hauenschild 2003: 100).

(11) *mïndɨ* 'fertile reindeer doe'; cf. Toju Tuvan *mïndï* 'reindeer doe' (Čadamba 1974: 63) and Soyot *mïndï* 'adult reindeer doe; camel cow' (Rassadin 2006: 93).[8] Beyond Sayan Turkic, cognates of *mïndɨ* are widespread across north Asian languages whose speakers are reindeer herders, see Tatár (1985) for examples. Etymologically, *mïndɨ* most likely represents a loanword from Samoyedic; cf. Mator *méinde* 'rangifer ferus'. According to Helimski (1997: 301-302), Mator *méinde* may be traced back to Proto-Samoyedic **məjan-cɐ̨ɜ* (*məjan* 'ground [gen]' + *cɐ̨ɜ* '[tamed] reindeer'). In Dukhan, the term *mïndɨ* also occurs in the following expressions: *haš* ('hairless') *mïndɨ* 'older reindeer doe with little fur' *gïsɨr* ('barren') *mïndɨ* 'dry doe' and *ulɨγ mïndɨ* 'big/mature reindeer doe'.

(12) *mïndɨǰak* 'reindeer doe that has fowled once'; cf. Tofan *mïndïǰaq* 'two-year-old reindeer doe' (Rassadin 1995: 47b), Soyot *mïndïǰaq* 'three-year-old reindeer doe' (Rassadin 2006: 93) and Toju Tuvan *mïndïčak* 'važenka dvux let' (Čadamba 1974: 63). Etymologically, *mïndɨǰak* is a diminutive form of *mïndɨ*.

(13) *öskʉsek* 'motherless fawn'; cognates of this Dukhan lexeme are not documented in the rest of Taiga Sayan Turkic. Morphologically, *öskʉsek* represents a diminutive formation from *öskʉs* 'orphan' (cf. ED 116–117: Old Turkic *ögsüz* 'motherless').

(14) *bir düktəγ guuday*[9] 'three-year-old castrated reindeer' (one hair-der *guuday*). Based on the same syntactic structure, i.e. cardinal number + *dük* 'hair' augmented

[7] In this regard, also see Hauenschild (2003: 105–106).

[8] Tofan, on the other hand, employs the term *iŋgen* (cf. ED 184: Old Turkic *ingēn* 'female camel') with reference to a reindeer doe. On the term *ingēn*, see also Hauenschild (2003: 94).

[9] According to several Dukhan informants, *bir düktəγ guuday* is a synonym for *düktəγ mïs̃*.

by the adjectivalizing suffix *-LXɣ* + *guuday*, *iʰx̌ɨ tüktəɣ guuday* and *üš tüktəɣ guuday* refer, respectively, to ‘two-haired guuday’ and ‘three-haired guuday’, i.e. ‘four-year vs. five-year old gelded reindeer’. Cognates are documented in the other Taiga Sayan Turkic varieties: Tere-Khöl Tuvan *iyi tüktüg kuuday* ‘three-year-old male reindeer’, *üš tüktüg kuuday* ‘four-year-old male reindeer’ and Toju Tuvan *bir düktüg mïyïs* ‘male reindeer of about 3 years of age’, *iyi düktüg mïyïs* ‘male reindeer of about four years of age’, *üš düktüg mïyïs* ‘male reindeer of about five years of age’ (Seren 2006: 82).

(15) *sããrsək mĩĩs* ‘reindeer with one dropped antler’ (*sããrsək* ‘one of two’ and *mĩĩs* ‘antler’).

(16) *dasfanaŋ* ‘wild, i.e. not tamed reindeer’. Among Taiga Sayan Turkic varieties merely Soyot displays the formally and semantically close cognate *daspanaŋ* ‘wild reindeer’ (Rassadin 2006: 43). The term *dasfanaŋ* possibly goes back to *taspan* (see below) and *aŋ* ‘wild animal’.

(17) *daspan* ‘one-year-old young reindeer’; cf. Toju Tuvan *daspan* ‘one-year-old reindeer’ (Sat 1987: 77), Tofan *daspan* ‘two-year-old young wild reindeer’ (Rassadin 1995: 21a). The etymology of *daspan* is obscure; for some proposals, see Tatarincev (2002: 105–106).

(18) *doŋgʉr* ‘approximately 18-month-old male reindeer’; cf. Tere-Khöl Tuvan *toŋgur*/*tuŋxur* ‘approximately two-year-old male reindeer’ (Kuular and Suvandi 2011a: 165), Toju Tuvan *dongur* ‘approximately one-year-old young reindeer’ and Soyot *doŋɣur* ~ *doŋɣïr* ‘two-year-old reindeer buck’ (Rassadin 2006: 45). Etymologically, *doŋgʉr* might be traced back to a rhotacised form of Turkic *toŋuz* ‘pig’. This assumption, however, needs further investigation. In this respect, it should be kept in mind that names of strong and dangerous animals, such as the boar, belong to the set of taboo names in use across Siberia and neighbouring areas.

(19) *doŋgʉy* ‘approximately 18-month-old reindeer doe’; cf. Tere-Khöl Tuvan *tuŋxuy* ~ *tuŋguy* ‘approximately two-year-old reindeer doe’ (Kuular and Suvandi 2011a: 165). Further formal cognates of *doŋgʉy* are Mator *toŋoi* ‘pig’ (Helimski 1997: 365, §1060), Toju Tuvan *doŋay* ‘two-year-old wild reindeer’ and Tuvan *doŋay* (Monguš 2003: 474a) ‘bear cub’. Etymologically, *doŋgʉy* may also ultimately go back to a rhotacised form of Turkic *toŋuz*. Note that in this regard that Helimski (1997: 365, §1060) traced Mator *toŋoi* ‘pig’ back to Proto-Turkic **doŋur*. Further Dukhan expressions including the term *doŋgʉy* are *ǰaš doŋgʉy*, literally ‘young

doŋgʉy', referring to a young doe about to foal, and *hur doŋgʉy*, literally 'last year's *doŋgʉy*' employed for a reindeer doe that has foaled twice.

(20) *düktəɣ mĩĩs* 'three-year-old castrated reindeer'; cf. Soyot *düktɨɣ miis* 'domesticated young reindeer buck in its third year of age' (Rassadin 2006: 47). The expression *düktəɣ mĩĩs* literary means 'hairy antler'.

(21) *uzən but* literally 'long leg'; this term specifically refers to a reindeer buck that will be castrated in the autumn, at least according to some informants; cf. Soyot *uzïn-but* 'young reindeer buck in its third year of age' (Rassadin 2006: 139).

3. Transcription and abbreviations

The transcription employed here follows general principles employed in Turcology, with the following additions: the symbols *ɨ* and *ʉ* represent the high central vowels occurring beyond first syllables, and the super-script ^{h} designates preaspiration of fortis consonants, corresponding to Tuvan and Tofan glottalization/ pharyngealization, marked here with the symbol ^{ʔ}. Abbreviations occurring in the grammatical glosses are: CAUS: causative, DER: derivation, and PART: participle.

References

Badamxatan, S. 1962. Xövsgöliyn Tsaatan ardïn aǰ baydlïn toym. *Studia Ethnographica* II/1, 1–66.

Čadamba, Zoya B. 1974. *Todžinskij dialekt tuvinskogo jazyka*. Kyzyl: Tuvinskoe knižnoe izdatel'stvo.

Čeremisov, K. M. 1973. *Burjatsko-russkij slovar'*. Moskva: Sovetskaja enciklopedija.

ED = Clauson, Sir Gerard 1972. *An etymological dictionary of pre-thirteenth-century Turkish.* Oxford: Oxford University Press.

Dankoff, Robert and Kelly, James (eds. & translators) 1982. *Compendium of the Turkic dialects, by Mahmūd al-Kāshgharī.* (Sources of Oriental Languages & Literatures 7, Turkish Sources VII/I.) Duxbury, MA: Harward University.

Donahoe, Brian and Plumley, Dan (eds.) 2003. *The troubled taiga*. Special Issue of Cultural Survival Quarterly. Spring 27: 1.

Dorlig, C. and Dadar-Ool, B. 1994. *Tïva-Mool tol'*. Ölgiy.

Erdal, Marcel 1991. *Old Turkic word formation. A functional approach to the lexicon*. 2 vols. (Turcologica 7.) Wiesbaden: Harrassowitz.

Gáspár, Csaba 2006. *Darkhat*. München: Lincom. (Languages of the World/Materials 419)

Harrison, David K. 2010. *The last speakers, the quest to save the most endangered languages*. Washington DC: National Geographic

Hauenschild, Ingeborg 2003. *Die Tierbezeichnungen bei Mahmud al-Kaschgari. Eine Untersuchung aus sprach- und kulturhistorischer Sicht*. Wiesbaden: Harrassowitz. (Turcologica 53)

Hauenschild, Ingeborg 2008. *Lexikon jakutischer Tierbezeichnungen*. Wiesbaden: Harrassowitz. (Turcologica 77)

Helimski, Eugen 1997. *Die Matorische Sprache. Wörterverzeichnis – Grundzüge der Grammatik – Sprachgeschichte.* Szeged: Szeged University. (Studia Uralo-Altaica 41)

Khabtagaeva, Bayarma 2015. Some remarks on Turkic elements of Mongolic origin in Yeniseian. *Studia Etymologica Cracoviensia* 20. 111–126.

Kuular, Elena Mandan-Oolovna and Suvandii, Nadežda Daryevna 2011a. Leksika olenevodstva v Tere-Xol'skom rajone. *Novye issledovanija Tuvy* 2011/1: 160–168.

Kuular, Elena Mandan-Oolovna and Suvandii, Nadežda Daryevna 2011b. Polovozrastnye nazvanija olenej v dialektax tuvinskogo jazyka. *Novye issledovanija Tuvy* 2011/4: 146–151.

Monguš, D. A. (ed.) 2003. *Tolkovyj slovar' tuvinskogo jazyka*. Tom 1: A – J. Novosibirsk: Nauka.

Nugteren, Hans & Roos, Marti 2006. Prolegomena to the classification of Western Yugur. In: Erdal, Marcel & Nevskaya, Irina (eds.) *Exploring the eastern frontiers of Turkic*. Wiesbaden: Harrassowitz. 99–130. (Turcologica 60)

Poppe, Nicholas 1954. *Grammar of written Mongolian*. Wiesbaden: Harrassowitz.

Ragagnin, Elisabetta 2010. O nekotoryx osobennostjax sajano-tjurkskix narečij severnoj Mongolii. In: Bičeldej K. A. (ed.) *Tuvinskaja pis'mennost' i voprosy issledovanija pis'mennostej i pis'mennyx pamjatnikov Rosii i central'no-aziatskogo regiona. Materiali meždunarodnoj naučnoj konferencii, posvjaščennoj 80-letiju tuvinskoj pis'mennosti 1-4 ijulja 2010 g.* Kyzyl. Tuvinskij Institut gumanitarnyx issledovanij pri Pravitel'stve Respubliki Tyva. 140–145.

Ragagnin, Elisabetta 2011. *Dukhan, a Turkic variety of Northern Mongolia: Description and analysis*. Wiesbaden: Harrassowitz. (Turcologica 76)

Ragagnin, Elisabetta 2012a. Etymologische Überlegungen zu einigen Rentiertermini im Sajantürkischen. In: Erdal, M.; Kellner-Heinkele, B.; Ragagnin, E. and Schönig, C. (eds.) *Botanica und Zoologica in der Türkischen Welt. Festschrift für Ingeborg Hauenschild*. Wiesbaden: Harrassowitz. 133–140.

Ragagnin, Elisabetta 2012b. Turkic substrates in Darkhat-Mongolian. In: Tumortogoo, D. (ed.) *Proceedings of the 10th International Congress of Mongolists*, Volume I: Prehistoric and historical periods of Mongolia's relations with various civilizations. Ulaan Baatar: IAMS secreteriat. 291–295.

Ragagnin, Elisabetta 2015. Bargu. In: Simion, S. and Burgio, E. (eds.) Giovanni Battista Ramusio: *Dei viaggi di Messer Marco Polo* (Filologie medievali e moderne 5). Venezia. http://edizionicafoscari.unive.it/it/edizioni/libri/978-88-6969-00-06/

Rassadin, Valentin I. 1971. *Fonetika i leksika tofalarskogo jazyka*. Ulan-Ude: Burjatskoe knižnoe izdatel'stvo.

Rassadin, Valentin I. 1978. *Morfologija tofalarskogo jazyka v sravnitel'nom osveščenii.* Moskva: Nauka.

Rassadin, Valentin I. 1996. *Prisajanskaja gruppa burjatskix govorov*. Ulan-Ude: Burjatskij naučnyj centr SO RAN.

Rassadin, Valentin I. 2005. *Tofalarsko-russkij i russko-tofalarskij slovar'*. Sankt-Peterburg: Drofa.

Rassadin, Valentin I. 2006. *Sojotsko-russkij slovar'*. Sankt-Peterburg: Drofa.

Rassadin, Valentin I. 2014. *Tofalarskij jazyk i ego mesto v sisteme tjurskix jazykov*. Elista: Izdatel'stvo Kalmyckogo Universiteta.

Röhrborn, Klaus 2010. *Uigurisches Wörterbuch: Sprachmaterial der vorislamischen türkischen Texte aus Zentralasien. – Neubearbeitung – I. Verben, Band 1: ab- – äzüglä-*. Stuttgart: Franz Steiner.

Sanžeev, G. D. 1931. *Darxatskij govor i fol'klor*. Leningrad: Izdatel'stvo AN SSSR.

Sat, Šuluu Č. 1987. *Tyva dialektologija*. Kïzïl: Tïvanïŋ nom ündürer čeri.

Seren, Polina 2006. *Tere-Xol'skij dialekt tuvinskogo jazyka*. Abakan: Izd-vo Xakasskogo gosudarstvennogo universiteta im. N. F. Katanova.

Somfai Kara, Dávid 1998. A xöwsgöl-i tibák. In: Birtalan, Ágnes (ed.) *Őseink nyomán Belső-Ázsiában II. Hitvilág és nyelvészet.* (Magyar Felsőoktatás Könyvek 10.) Budapest: Nemzeti Tankönyvkiadó. 17–21.

Ščerbak, Aleksandr M. 1961. Nazvanija domašnix i dikix životnyx v tjurkskix jazykax. In: Ubrjatova, Elizaveta I. (ed.) *Istoričeskoe razvitie leksiki tjurkskix jazykov*. Moskva: Izdatel'stvo Akademii nauk SSSR, Institut jazykoznanija. 82–172.

Thackston, Wheeler M. (tr.) 2012. *Classical writings of the Medieval Islamic world: Persian histories of the Mongol dynasties: Rashiduddin Fazlullah* (vol. III). London: I. B. Tauris.

Tatár, Magdolna 1985. Sayan Etymologies. In: *Beşinci Milletlerarası Türkoloji Kongresi. İstanbul, 23–28 Eylül 1985. Tebliğler. I. Türk Dili 1.* İstanbul: İstanbul Üniversitesi Edebiyat Fakültesi Türkiyat Araştırma Merkezi. 243–252.

Tatarincev, Boris I. 2000. *Ėtimologičeskij slovar' tuvinskogo jazyka*. Tom 1: A–B. Novosibirsk: Nauka.

Tatarincev, Boris I. 2002. *Ėtimologičeskij slovar' tuvinskogo jazyka*. Tom 2: D, Yo, I, J. Novosibirsk: Nauka.

Tatarincev, Boris I. 2004. *Ėtimologičeskij slovar' tuvinskogo jazyka*. Tom 3: K, L. Novosibirsk: Nauka.

Tenišev, Ėdgem R. (ed.) 1968. *Tuvinsko-russkij slovar'.* Moskva: Sovetskaja enciklopedija.

Vainshtein, Sevyan [Vajnštejn, Sev'jan I.] 1980. *Nomads of South Siberia. The pastoral economies of Tuva.* Cambridge. [English translation of Vajnšteijn, Sev'jan 1972, Istoričeskaja etnografija tuvincev, by Michael Colenso]

Verbickij Vasilij 2005[2]. *Slovar'altajskogo i aladagskogo narečij tjurkskogo jazyka*. Gorno-Altajsk: Ak Čeček.

Werner, Heinrich 2002. *Vergleichendes Wörterbuch der Jenissej-Sprachen*. Band 2: L-Š. Wiesbaden: Harrassowitz.

Žukovskaja, N. L – Oreškina, M. V. & Rassadin, Valentin I. 2002. Sojotskij jazyk. In: Neroznak, V. P. (red.) *Jazyki narodov Rossii. Krasnaja kniga*. Moskva: Academia. 164–170.

On the Mongolian verb of motion *yav-* 'to go, to travel, to leave'

Veronika Zikmundova

Charles University, Prague

1. Introduction

Since the 1970s, verbs of motion as an important part of the vocabularies of most languages have been extensively studied by cognitive linguists, in particular, Talmy (1975, 1985, 2000), Levin and Rappaport Hovav (1992), Slobin (1996), and others.[1] Based on Talmy's terminological apparatus (Figure, Path, Manner, Ground, etc.), a series of studies have featured the "universals" of the usage of verbs of motion in many languages.

This cross-linguistic research has shown that languages tend to fall into two major classes in terms of how they express meaning connected to motion in space – these are the "verb-framed" vs. "satellite-framed" languages. According to this classification, the Altaic languages generally belong to the former class. Recently, Nakazawa (2007, 2009) has focused on verbs of motion in Mongolian. Some recent studies have emphasized the necessity of both more focused and intense methods of apprehending the verbal systems of motion in various languages (Word-Allbritton 2004: 9–10).

It is my conviction that that the descriptive approach to particular languages, based on careful examination of larger samples of language data, contributes important information to the pragmatics of the studied languages. The description of their semantics, including the grammaticalized and figurative usage of particular verbs of motion, may become one of the clues to comprehending the "linguistic mind" and the ways in which particular languages view and reflect on the world.

The present paper continues a discussion of the topic of verbs of motion in Mongolian and Sibe, which I described in previous articles (Zikmundova 2010,

[1] See the bibliography by Slobin and Matsumoto at http://www.lit.kobe-u.ac.jp/~yomatsum/motionbiblio1.pdf.

2011), and continues the examination of the Sibe verb *yaf-* (2013). The general translation of the Sibe verb *yaf-* and the Khalkha Mongolian verb *yav-* are largely similar, covering the meanings 'to go', 'to travel', 'to depart', and in a more the abstract or figurative sense referring to various aspects of existing.

The verbs of motion which I examined in previous papers (Sibe *gen-* 'to go there, to go' and *ji-* 'to come here', Khalkha Mongolian *oč-* 'to go there', *ir-* 'to come here') form a part of a largely coherent system of verbs of motion consisting of antonymical pairs with distinct spatial semantics, which is typical of Altaic languages in general. However, the verb with the general meaning of motion, examined in the present article, lacks some characteristics of this system, tending rather to resemble the usage of verbs of motion in the "satellite-framed" languages. In particular, while it is difficult to establish an accurate English translation of the verbs 'to come here' or 'to go there' which would encompass the entire range of meanings, both the Sibe verb *yaf-* and the Mongolian verb *yav-*, including much of their figurative meanings, are aptly translated with the English verb 'to go'. An important difference between English on the one hand and both Sibe and Mongolian on the other lies in the emphasis on motion as the basic aspect of life, which is mirrored in the semantics of both of these verbs.

In the present paper I examine the semantics of the Khalkha Mongolian verb *yav-*, based on examples from Modern Khalkha Mongolian. The analyzed data come from various sources: folktales and proverbs retrieved from the Database of Mongolian folklore texts at http://www.signeta.cz/textsearch/, interviews recorded by V. Kapišovská in the Khentii Province of Mongolia in 2014, utterances noted during everyday conversations in 2015, and texts and conversations available on the Internet. The examples are roughly divided into "literal" and "figurative" usage and further classified into tentative groups according to the manner of usage.

2. The verb *yav-* expressing the literal meaning of 'motion'

The basic semantics of the verb *yav-* is connected to the process of motion. Within its range of meanings, the emphasis on 'leaving, or setting off [in motion]' and 'to undertaking a journey' seem to be especially frequent.

2.1. The verb *yav-* designating the process of motion

The general meaning of motion in space seems to be the basic and most important component within the semantics of the Mongolian verb *yav-*. In this literal meaning

the agent is usually either an animate being, or an object for which such motion is a typical action. In this type of usage spatial determination may be either present (sentences 1 and 2) or absent (sentence 3). Unlike in other Mongolian verbs of motion, however, here the semantic emphasis is on the motion itself rather than on its goal or direction.

(1) *Нөгөө хүн тэр явж байна.*
nögȫ̄ xün **ter** **yav-ž** **bai-na**.
that* person **that** **go-CI**[2] **be-IMPRF**
(*in reference to a previously mentioned fact)
'That [previously mentioned] person **is walking/riding over there**.'

(2) *Би сайн морио унаад*
bi sain mori-ō un-ād
I good horse-ACC-POSS ride-CP

саадгүй цаашаа явна.
sādgüi **cāšā** **yav-na.**
without.obstacles further go-IMPRF
'I will ride [on] my good horse and **proceed smoothly**.'

(3) *Улаан ширх **явна**. Үүнийг ална уу?*
ulān širx yav-na. Ün-īg al-na uu?
red louse go-IMPRF this-ACC kill-IMPRF Q
'A red louse is **creeping [towards the location of the speaker**]. Should I kill it?'

2.2. The verb *yav-* – expressing the meaning of a journey with a specific purpose

While designating the process of motion, many of the examined examples can also be interpreted as referring to a whole journey with a certain goal and purpose, which may be either overtly expressed (4)–(7), or understood from the context (9). When the goal of motion is defined spatially, it is mostly in the lative case (4), (9). In

[2] The abbreviations in the interlinear glosses follow, for the most part, the list of standard abbreviations available at https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf (Leipzig glossing rules). The rest is abbreviated from grammatical terms used in Vacek and Luvsandorj (2004).

colloquial usage, however, the case suffix is often omitted (8). When the purpose of motion is expressed by a verb, the latter mostly has the form of a verbal noun in the instrumental case, which is a form typically used for the expression of purpose – (7), (9), or employs an analytic construction. When the purpose of motion is expressed by a noun, it usually stands in the dative-locative (6), (10).

(4) *Та* ***гадагшаа*** ***явсан*** *юм* *уу?*
ta **gada-gšā** **yav-san** yum uu?
you **outside-LAT** **go-NP** PTC Q
'Oh, **you have gone elsewhere**? [You are not here?]'

(5) *Нараа эгч* ***Монгол*** ***явсан***.
Narā egč **Mongol** **yav-san**.
naraa elder.sister **Mongolia** **go-NP**
'Aunt Naraa **went to Mongolia/ is in Mongolia**.'

(6) *Арван гурван жилийн* *анд* *явсан* *аав*
arvan gurvan žil-īn **an-d** **yav-san** āv
ten three year-GEN **hunt-DL** **go-NP** father

хаан *чинь* *хүрээд* *ирэв.*
xān čin‘ xür-ēd ir-ev.
where POSS.2SG arrive-CP come-PRET.PERF
'Your father, who **left for a hunt** of thirteen years**,** has come back.'

(7) *Гадаадад* *сурахаар* *явах* *хүмүүст*
gadāda-d **sura-x-ār** **yava-x** xümű̄s-t
abroad-DL **study-NF-INS** **go-NF** people-DL

банкны *батлагаа* *гаргана.*
bankn-ï batlagā garga-na.
bank-GEN certificate issue-IMPRF
'We issue a bank certificate for those who **go to study** abroad.'

(8) *Аав аа,* *аав аа!* *Ямар* *мориороо* *явах* *вэ?*
āv-ā, āv-ā! Yamar mori-or-ō **yava-x** ve?
father-VOC father-VOC what horse-INS-POSS.REF **go-NF** Q
'Father, with which horse should I **go**?'

(9) *Нөхөр маань Баянхонгор аймаг руу ажил*
nöxör mān' Bayanxongor aimag-rū **ažil**
husband POSS1PL.GEN Bayankhongor aimag-LAT **work**

хийхээр яваад эзгүй байсан.
xī-x-ēr **yav-ād** ezgüi bai-san.
do-NF-INS **go-CP** absent be-**NP**
'My husband **went** to Bayankhongor aimag **to work** and was not at home.'

(10) *Ухнанд явсан хүү ямаа ишиглэсэн хойно ирэв.*
uxnan-d **yav-san** xū̃ yamā išigle-sen xoino ire-v.
ram-DL **go-NP** boy goat give.birth-NP after come-PRET.PERF
'The boy who **went to fetch the ram** returned after the goats gave birth.'

2.3. The verb *yav-* expressing the meaning 'to set oneself in motion, to leave'

In everyday colloquial usage, the verb *yav-* is frequently used in connection with leaving a certain site, setting off, etc.

(11) *За одоо явцгаая!*
za odō **yav-cgā-ya**!
PTC now **go-VP-VOL**
'So, let's **go** now!'

(12) *Ээж явсан юм уу?*
ēž **yav-san** yum ū?
mother **go-NP** PTC Q
'**Has** Mother **left**?'

(13) *Явахаасаа өмнө надаас юу хүссэнээ ав!*
yava-x-ās-ā ömnö nad-ās yū̃ xüs-sen-ē av!
go-NF-ABL-POSS before 1SG-ABL what wish-NP-ACC-POSS take-IMP
'Before you **leave**, take anything you wish from me.'

(14) *Явна явна гэж ямааны мах барав…*
yav-na **yav-na** ge-ž yamān-ï max bara-v.
go-IMPRF **go-IMPRF** say-CI goat-GEN meat finish-PRET.PERF
'Saying "**I am leaving,**" he finished off the meat of an [entire] goat.'

(15) *Тэр охин урьд нь гэрээсээ хоёр,*
ter oxin ur'd n' ger-ēs-ē xoyor,
that girl before POSS3SG home-ABL-POSS two

гурван удаа явж байсан гэнэ.
gurvan udā **yav-ž** **bai-san** ge-ne.
three time **go-CI** **be-NP** say-IMPRF
'They say that formerly this girl **had left** home two or three times.'

(16) *Дахиж гэрээсээ хэзээ ч явахгүй.*
daxiž ger-ēs-ē xezē č yava-x-güi.
again home-ABL-POSS when PTC **go-NF-NEG**
'I will never **leave** home again.'

(17) *Тэнд очоод сураглатал охиныг маань байхгүй,*
tend oč-ōd suragla-tal oxin-ïg mān' baix-güi,
there go.there-CP ask-CT daughter-ACC 1PL.GEN be-NF-NEG

явчихсан гэсэн.
yav-čix-san ge-sen.
go-INT-NP say-**NP**
'When we went there to ask, we were told our daughter wasn't there; she **had left**.'

3. Modal and grammatical usage of the verb *yav-*

3.1. The verb *yav-* expressing a modality of continuous motion

A frequent type of phrase is one where the verb *yav-* is preceded by another verb of motion in the form of an imperfective converb. In such phrases, the function of the verb *yav-* may be interpreted as modal, foregrounding the notion of continuous motion. This kind of usage is typical for dependent phrases with a temporal meaning (sentences 20–22).

(18) *Хүчтэй нарны шуурга Дэлхийг **чиглэн ирж явна**.*
xüčtei narn-ï šūrga delxī-g **čigle-n** **ir-ž** **yav-na**.
strong sun-GEN storm Earth-ACC head-CM come-CI go-IMPRF
'A strong sun storm **is heading towards** the Earth.'

(19) *Хоёр залуу цирк үзээд* ***гарч*** ***явна.***
xoyor zalū cirk üz-ēd **gar-č** **yav-na.**
two youngster circus see-CP go.out-CI go-IMPRF
'Two young men are **leaving** the circus after a performance.'

(20) *Уртын давaа давaад баруун тийшээгээ*
Urtïn davā dav-ād barūn tī-šē-gē
Urtyn pass surmount-CP west there-LAT-POSS.REF

буужx явахад …
bū-ž **yava-x-ad**...
descend-CI **go-NF-DL**
'As they crossed the Urtyn Pass and **were descending** in a westerly direction…'

(21) ***Буцаж*** ***явах*** *замдаа* *осол* *гаргасан.*
buca-ž **yava-x** zam-d-aa osol gar-ga-san
Return-CI **go-NF** way-DL-POSS.REF accident go.out-CAUS-NP
'**On the way back**, he caused a car accident.'

(22) *Гэртээ* ***харьж*** ***явтал*** *гудамжны*
ger-t-ē **xar'-ž** **yav-tal** gudamžn-ï
home-DL- POSS.REF **return.home-CI** **go-CT** treet-GEN

өнцөгт *гоё* *бүсгүй* *зогсохыг* *үзээд*
öncög-t goyo büsgüi zogsox-ïg üz-ēd
corner-DL beautiful woman stand-NF-ACC see-CP
'**As he was returning** home, he saw a beautiful woman standing on the corner of the street…'

(23) *Гурван хүн машинаар чоно* ***хөөж*** ***яваад*** *осолджээ.*
gurvan xün mašin-ār čono **xō̆-ž** **yav-ād** osold-žē.
three person car-INS wolf **hunt-CI** **go-CP** crash-PRET.IMP
'Three people, **hunting** wolves while [driving] in the car, were in an accident.'

3.2 The verb *yav-* expressing a modality of intensity and change

This modal usage is derived from the basic semantics of ‘setting oneself in motion’ or ‘leaving’. In this type of phrases the verb *yav-* adds a modal colouring of an intensive and persistent change of state (Ex. 24–26), or of an intensive action (27).

(24) *Их л ойлгомжгүй* ***болоод*** ***явчихлаа.***
ix l oilgomžgüi **bol-ōd** **yav-čix-lā**.
much PTC incomprehensible **become-CP** **go-INT-PRES.PERF**
‘It **became** very confusing.’

(25) *Сайхан* ***болоод*** *явчихлаа шүү.*
saixan **bol-ōd** **yav-čix-lā** šǖ.
nice/good **become-CP** **go-INT-PRES.PERF** PTC
‘It **became** better; it **improved** a lot.’

(26) *Машины наймаанд яваад амьдрал*
mašin-ï naimān-d yav-ād am’dral
car-GEN business-DL go-CP life
‘*под*’ *хийгээд* ***явчихлаа.***
‘pod’ **xīg-ēd** **yav-čix-lā**.
‘pod’ **do-CP** **go-INT-PRES.PERF**
‘After I [started] doing trade in buying and selling cars, my life **improved.**’

(27) *Унтаж байтал толгой дээр минь*
unta-ž bai-tal tolgoi dēr min’
sleep-CI be-CT head on POSS1SG
нэг юм ***таш баш*** *хийгээд явчихлаа шүү.*
neg yum **taš** **baš** **xīg-ēd** **yav-čix-lā** šǖ.
one thing **tash** **bash** **do-CP** **go-INT-PRES.PERF** PTC
‘When I was sleeping, something **made a bumping noice** above my head.’

3.3. The verb *yav-* in auxiliary function

A situation that describes two speakers encountering each in open space, when one or both speakers are in motion, employs the verb *yav-* as the final member of a verb phrase following other verbs which are non-motion verbs (28–31). In this kind of situation—typical in the life of nomadic herders – the verb *yav-* may possibly be interpreted as expressing a grammatical meaning very close to the auxiliary

existential verb *bai-* 'to be.' In most of the examples below, the verb *yav-* could be replaced by *bai-* without any substantial change in the meaning of the sentence.

(28)

Барон	*Унгерний*	*тухай*	***судлах***	***гээд***	***яваа***	*юм*	*уу?*
baron	Ungern-ī	tuxai	**sudla-x**	**ge-ēd**	**yav-ā**	yum	ū?
Baron	Ungern-GEN	about	**research-NF**	**say-CP**	**go-NI**	PTC	Q

'So you **are travelling with the purpose of studying** Baron Ungern?'

(29)

Та нар	***юунд***	***яваа***	*хүмүүс*	*вэ?*
Ta-nar	**yūn-d**	**yav-ā**	xümūs	ve?
you-PL	**what-DL**	**go-NI**	people	Q

'What **is the purpose** of your journey?'

(30)

Намайг	*алж*	*идэх*	*шахлаа.*
Nama-ig	al-ž	ide-x	šax-lā.
1SG-ACC	kill-CI	eat-NF	press-PRES.PERF

Тэгээд	*би*	*ингэж*	***зугатаж***	***явна.***
tegēd	bi	inge-ž	**zugata-ž**	**yav-na**.
so	I	like.this	**flee-CI**	**go**-IMPRF

'They almost killed me, so **I am running** like this.'

(31)

Бар	*чи*	*юундаа*	*ингэж*	*их*
bar	či	yūn-d-ā	ingež	ix
tiger	you	what-DL-POSS.REF	like.this	much

сандарч	***явна?***
sandar-č	**yav-na**?
be.anxious-CI	**go-IMPRF**

'Why **are you** so **anxious**, tiger?'

(32)

Судалгаа	*хийж*	*явсан*	*ажилтныг*	*хутгалж*
sudalgā	**xī-ž**	**yav-san**	ažiltn-ïg	xutgal-ž
research	**do-CI**	**go-NP**	worker-ACC	stabb-CI

амь	*насыг*	*нь*	*хохироосон*	*хэрэг*	*гарчээ.*
am'	nas-ïg	n'	xoxirō-son	xereg	gar-čē.
life	age-ACC	POSS3SG	harm-NP	matter	go.out-PRET.IMP

'It happened that an employee was stabbed to death **while doing field research**.'

(33) *Үүргээ* ***гүйцэтгэж*** *яваад* *хоёр* *цагдаа*
ӱ̄rgē **güicetge-ž** **yav-ād** xoyor cagdā
task-POSS.REF **fulfill-CI** **go-CP** two policeman
амиа *алджээ*
ami-a ald-žē
soul-POSS.REF lose-PRET.PERF
'Two policemen lost their lives **while in service**.'

(34) *Ихэнхдээ* *хоолыг* *нь* ***дээрэмдээд*** ***явсан*** *юм.*
inenxdē xōl-ïg- n' **dēremd-ēd** **yav-san** **yum**.
mostly food-ACC POSS3SG **rob-CP** **go-NP** PTC
'He **was** mainly **stealing** their food.'

3.4. The verb *yav-* as a member of a complex descriptive expression

The verb *yav-*, when forming part of a complex description of actions, shows lesser variability in its usage than other verbs of motion. This is due to its lack of a concrete deictic function.

3.4.1. The verb *yav-* in initial phrasal position

When placed in initial phrasal position, the verb *yav-* usually describes either a departure with consequent motion (35–38) or a process of motion (39–40). In the latter case, it often adds a modal colouring of slight expectation followed by surprise (39).

(35) *Хөөш* *чи* ***талханд*** ***яваад*** ***ир!***
xō̄š či **talxan-d** **yav-ād** **ir**!
hey you **bread-DL** **go-CP** **come**
'**Go buy bread**, please!'

(36) *Аав аа,* *аав аа,* *би* *алтан* *мөнгөн*
āv-ā, āv-ā, bi altan möngön
father-VOC father-VOC I golden silver
аргайндаа ***яваад*** ***ирье!***
argain-d-ā **yav-ād** **ir-ye**
playing.bone-DL-POSS.REF **go-CP** **come-VOL**

'Father, I will **go fetch** my golden and silver playing bones.'

(37) *Би* *өчигдөр* *Хархорум* *хот* ***яваад*** ***ирсэн.***
Bi öčigdör Xarxorum xot **yav-ād** **ir-sen.**
I yesterday Kharkhorum city **go-CP** **come-NP**
'Yesterday I **paid a visit** to Kharakhorum city.'

(38) *Тэгэхээр* *нь* *хонь* ***явж*** ***байсан***
tege-x-ēr n' xon' **yav-ž** **bai-san**
do.so-NP-INS POSS3SG sheep **go-CI** **be-NP**

газарт ***яваад*** ***иржээ.***
gazar-t **yav-ād** **ir-žē**
place-DL **go-CP** **come-PRET.IMP**
'And so he **arrived** at the place where the sheep **had been** before.'

(39) ***Яваад*** ***очсон*** *чинь* *Дорж* *сууж* *байна.*
yav-ād **oč-son** čin' Dorž sū-ž bai-na.
go-CP **go.there-NP** PTC Dorj sit-CI be-IMPRF
'I **arrived** there and, oops, Dorj was sitting there.'

(40) *Гэрт* *нь* ***яваад*** ***очиж*** *зүрхэлсэнгүй*
ger-t n' **yav-ād** **oči-ž** zürxel-sen-güi.
home-DL POSS3SG go-CP go.there-CI dare-NP-NEG
'He did not dare to **go directly** to her place.'

3.4.2. The verb *yav-* in final phrasal position

When used as a final component of verbal phrases, the verb *yav-* often expresses a spatial meaning of motion directed away from the location of the speaker, or from the scene of the action.

(41) *Дүү* *чинь* *өчигдөр* ***ирээд*** ***явсан.***
dū̃ čin' öčigdör **ir-ēd** **yav-san**.
younger.sibling POSS2SG yesterday **come-CP** **go-NP**
'Your brother **came around** yesterday.'

(42) *Орос,* *буриад,* *хамниган* *хүмүүсүүдийг*
oros, buriad, xamnigan xümū̃s-ū̃d-īg
Russian, Buryat, Khamnigan people-PL-ACC

зэвсгийн *хүчээр* ***аваад*** ***явсан***.
zevsg-īn xüč-ēr **av-ād** **yav-san**.
arm-GEN power-INS **take-CP** **go-PERF**
'By using the might of weapons, he **took away** the Russian, Buryat and Khamnigan people.'

(43) *Сурч* *байсан* *сургуулиас* *нь* ***аваад*** ***явсан***.
sur-č bai-san surgūl'-ās n' **av-ād** **yav-san**.
study-CI be-**NP** school-ABL POSS3SG **take-CP** **go-NP**
'They **took her out** of the school [where] she had been studying.'

(44) "*Дараа* *уулзъя*" *гээд* ***ороод*** ***явчихсан*** *гэнэ.*
"darā ūlz-ya" ge-ed **or-ōd** **yav-čix-san** ge-ne.
later meet-VOL say-CP enter-CP **go-INT-NP** say-IMPRF
'She said "See you later", **entered** the house (**and disappeared**).'

(45) *Мафийн* *гишүүн* *оросын* *шоронгоос* *халбагаар*
Maf-īn gišűn oros-ïn šorong-ōs xalbag-ār
mafia member Russian-GEN prison-ABL spoon-INS

нүх *ухаад* ***гараад*** *явчихаж.*
nüx ux-ād **gar-ād** **yav-čixa-ž.**
hole dig-CP **go.out-CP** **go-INT-PRET.IMP**
'A member of the mafia **escaped** from a Russian prison [by] digging a hole with his spoon.'

(46) *Яаж* *чамайг* *ийм* *хэцүү* *үед* ***хаяад***
yāž čama-ig īm xecű üye-d **xay-ād**
how 2SG-ACC such difficult time-DL **throw-CP**

явж *чадав* *аа.*
yav-ž čad-av ā.
go-CI can-PRET.PERF PTC
'How could he **abandon** you in such a difficult time?'

(47) *Манай* *эндээс* *морь* ***унаад*** ***явсан***.
mana-i end-ēs mor' **un-ād** **yav-san**.
1PL.GEN here-ABL horse **ride-CP** **go-NP**
'He **left** our place **riding** [our] horse.'

(48) *Чөтгөр…* *хоёр* *шавар* *хүн* *үрүү* *эргэн* *тойрон*
čötgör... xoyor šavar xün ürū̄ ergen toiron
demon… two clay person towards around around

шээж ***бузарлаад*** ***яваад*** ***өгч*** *гэнэ.*
šē-ž buzarl-ād yav-ād ög-č ge-ne.
urinate-CI pollute-CP go-CP give-PRET.IMP say-IMPRF
‘The devil **defiled** the two clay people **by urinating** on and all around them, **and left**.’

4. Derived and figurative usages of the verb *yav-*

When considering the range of meanings of the Khalkha Mongolian verb *yav-*, it is nearly impossible to distinguish between basic and derived meanings. The material yields a large portion of examples closely connected with motion; at the same time, the verb *yav-* may also designate a more general meaning. In some cases, it is best translated with existential verbs (cf. section **3.5**).

4.1. Designation of an activity connected with movement

The verb *yav-* is frequently used as a representative expression for an action connected with movement. In these types of phrases, the action is usually specified by a noun — the object or goal of the action — in the dative-locative case.

(49) *Охин* *чинь* ***хичээлдээ*** ***явж*** ***байгаа*** *юу?*
oxin čin’ **xičēl-d-ē** **yav-ž** **baig-ā** yū?
daughter 2SG-GEN **class-DL-POSS.REF** **go-CI** **be-NI** Q
‘Does your daughter **go to school**?’

(50) *Хүү…* ***малдаа*** ***явдаг*** ***болж*** *гэнэ.*
xū̄... **mal-d-ā** **yav-dag** **bol-ž** ge-ne.
boy… beasts**-DL-POSS.REF** **go-NU** **become**-PRET.IMP say-IMPRF
‘The boy **began herding his beasts** regularly.’

(51) *Бид* *охиноо* *алдчихлаа* *гээд…* *аль* *сайн*
bid oxin-ō ald-čix-lā ge-ēd... al’ sayin
we daughter-POSS.REF lose-INT.PRES.-IMP say-CP which good

мэргэч ***төлгөч,*** ***бөө,*** ***ламаар*** ***явсан***.
mergeč **tölgöč,** **bō̄,** **lam-ār** **yav-san**.
good **seer** **shaman** **lama-INS** **go-NP**

‘Since our daughter was missing, we **visited** every good **diviner, shaman and lama**.’

(52) *Тэгэхэд ‘Амьд байна. Нэг хүнтэй* ***хамт***
tege-xe-d ‘am’d bai-na. neg xün-tei **xamt**
do.so-NF-DL alive be-IMPRF one person-SOC **together**

яваад байна’ гэсэн.
yav-ād bai-na’ ge-sen.
go-CP be-IMPRF say-NP
‘And they said: “She is alive, she **is (living, moving) together** with another person.”’

4.2. Derivational meanings of ‘living, spending time’

An extremely frequent abstract usage involves instances where the verb *yav-* is used to refer to life in general, making one’s living, success or lack thereof, one’s conduct, and so on.

(53) *Залуудаа* ***сайн явсан*** *нь*
zalū-d-ā **sain yav-san** n’
young-DL-POSS **good go-NP** POSS3SG

наслахад мэдэгдэнэ.
nasla-xa-d mede-gde-ne.
grow.old-NF-DL know-PAS-IMPRF
‘If you **live well** in your youth, the results become evident when you grow old.’

(54) ***Муу явахад*** *нөхөр хол* ***сайн явахад*** *садан ойр.*
mū yav-ax-ad nöxör xol **sain yav-ax-ad** sadan oir.
bad go-NF-DL friend distant good go-NF-DL relative close
‘When **things aren’t going well**, friends are far away; when **things are going well**, relatives are close by.’

(55) ***Үнэнээр явсан*** *хүн үхэр тэргээр туулай гүйцнэ.*
ünen-ēr yav-san xün üxer terg-ēr tūlai güic-ne.
truth-INS go-NP person cow cart-INS rabbit reach
‘Those who **live in truth** can overtake a rabbit with an oxen cart.’

(56) *Сохор Тарваа үүнээс хойш удтал* ***амьд*** ***явсан.***
soxor tarvā ūn-ēs xoiš udtal **am’d** **yav-san**.
blind Tarvaa this-ABL after long **alive** **go-NP**
‘After that, blind Tarvaa still **lived** a long time.’

(57) ***Тэнүүн*** ***явахад*** *тэмээгээр тусалснаас* ***тэвдэж***
tenūn **yava-xa-d** temē-gēr tusal-san-ās **tevde-ž**
calm **go-NF-DL** camel-INS help-NP-ABL **be at a loss-CI**

явахад *тэвнээр тусалсан нь дээр.*
yava-xa-d tevn-ēr tusal-san n’ dēr.
go-NF-DL awl-INS help-NP POSS3SG above
‘It is better to help those **in need** with an awl than to help those who **have enough** with a camel.’

(58) ‘*Хурмаст тэнгэрээ тайх юмсан’*
‘Xurmast tenger-ē tai-x yumsan’
Khurmast tengri-ACC-POSS.REF worship-NF PTC

гэж ***бодож*** ***явдаг*** *байжээ.*
ge-ž **bodo-ž** **yav-dag** bai-žē.
say-CI **think-CI** **go-NU** be-PRES.IMP
‘He **used to think**: I would like to make an offering to Khurmast Tengri.’

(59) *Ялангуяа нэгдлийн* ***ажилд*** *их* ***явсан*** *даа*
Yalanguya negdel-īn **ažil-d** ix **yav-san** dā.
above.all cooperative-GEN **work-DL** much **go-NP** PTC
‘I **used to work** mainly in cooperatives.’

4.3. Reference to the passage of time

The phrase *yavsār baigād*, literally meaning ‘having been going [on] for a long time’, is idiomatically used to refer to the passage of time. The subject may be either a typical subject of the verb *yav-* (60) or any other subject (61).

(60) *Бид өсөх гэж* ***завгүй*** *явсаар байгаад аав*
bid ösö-x ge-ž **zavgüi** **yav-sār** **bai-gād** āv
we grow-NF say-CI **busy** **go-CP** **be-CP** father

ээжийнхээ хөгширч байгааг анзаардаггүй.
ēž-īnx-ē xögšir-č bai-gā-g anzār-dag-güi.

mother-GEN-POSS.REF grow.old-CI be-NI-ACC notice-NU-NEG
'**Being busy** with [our] growing up, we do not notice how our parents grow old.'

(61) *Дүгрэг гэдэг үг явсаар байгаад*
dügreg ge-deg üg **yav-sār** **bai-gād**
"Dügreg" say-NU word **go-CA** **be-CP**

төгрөг болсон байхаа
tögrög bol-son baixā.
"tögrög" become-NP probably
'It seems that the word "dügreg" evolved into [the word] "tögrög" **over time**.'

4.4. Figurative usage in place of an existential verb

The verb *yav-* may be also idiomatically used for inanimate objects, in which case it is best translated wih an existential verb. Such usage has rich emotional associations, for the most part making the statement more agreeable to the listener.

(62) *Миний түрүүвч чиний цүнхэнд* ***явж*** ***байна*** ***уу?***
Min-ī tŭrŭvč čin-ī cünxen-d **yav-ž** **bai-na** **ū**?
1SG-GEN purse 2SG-GEN bag.DL **go-CI** **be-IMPRF** Q
'**Is** my purse in your bag?'

4.5 Expression of intention, change of direction

In modern texts, the figurative usage of the verb *yav-*, in the sense of 'change of direction, intention', etc. is common. This manner of usage, corresponding to that of European languages and possibly also the result of the influence of the latter, seems nonetheless to conform with the general semantics of the Mongolian verb.

(63) *Нийгэм хаашаа* ***яваад*** ***байна?***
nīgem xāšā yav-ād bai-na?
society where.to **go**-CP **be-IMPRF**
'Where is society **headed**?'

4.6. Metaphor for death

As in many other languages, the verb yav- is used as a euphemistic expression in reference to death.

(64) *Жасрайн Жанцан энэ ертөнцөөс гэнэт явчихаж.*
Žasrain Žancan ene yertönc-ӧ̄s genet **yav-čix-až**.
Jasrain Jantsan this world-ABL suddenly **go-INT-PRET**
‘Jasrain Jantsan has suddenly **left** this world.’

4.7 Problems in delimitation of „literal“ and „figurative“ meaning

Finally, it seems necessary to introduce one important feature of the semantics of the verb *yav-*. Many instances of usage, which a European linguist would most likely – according to the mode of European linguistics – label as figurative, modal, auxiliary, and so on, actually convey the basic and literal meaning of movement. Nonetheless, a semantic overlap with the derivational meanings of existence, living, or the modality of continuous action, still occurs. The following two examples, both taken from folkloric texts, show the verb *yav-* being used in syntactical or logical parallels in both the literal and basic meaning of motion, along with the figurative meaning of mode of existence, behaviour, habit, and so on.

The first series of examples, taken from the folktale known as *Цуут цагаач гүү, цолмон цагаан унага* (Cūt cagāč gū̆, colmon cagān unaga), is taken from the section in which the White Mare advises its foal on how to act in his own independent life; the foal subsequently disregards this advice only to then discover the White Mare’s reasons behind it. Here, examples of both the literal meaning of movement (67), (69), along with the meaning of habitus or mode of existence (65), (66), (68), (70) accumulate in logical paralells, linked by the image of the life of a Mongolian horse, for whom constant motion is the basic mode of life.

(65) *Айлын бууцан дээр битгий **унтаж** **яваарай**.*
ail-īn būcan dēr bitgī **unta-ž** **yav-ārai**.
family-GEN camping.ground on PROH **sleep-CI** **go-IMP**
‘**Do not sleep** in empty camping grounds.’

(66) *Эрүүл газар **унтаж** **яваарай.***
erū̆l gazar **unta-ž** **yav-ārai**.
healthy place **sleep-CI** **go-IMP**
‘**Do sleep** in healthy places.’

(67) *Ижилтэйгээ* ***явахдаа*** *адууны*
ižil-tei-gē **yava-x-d-ā** adūn-ï
alike-SOC-POSS.REF **go-NF-DL-POSS.REF** horse-GEN

захад ***гарч*** ***явваарай***.
zax-ad **gar-č** **yav-ārai**.
edge-DL **go.out-CI** **go-IMP**
'When you **walk/run** with your herd, always **keep to the side** of the herd.'

(68) *Усанд* *орж* *ус* *ууxдаа* *түрүүнд*
usan-d or-ž us ū-x-d-ā türūn-d
water-DL enter-CI water drink-NF-DL-POSS.REF front-DL

нь *орж* ***ууж*** ***явваарай***.
n' or-ž **ū-ž** **yav-ārai**.
POSS3SG enter-CI **drink-CI** **go-IMP**
'When going to the waterside to drink, **be** among the first **to** (step into the water) and **drink**.'

(69) *Адууны* *дунд* ***явсан*** *чинь* *ижлүүд* *нь*
adūn-ï dund **yav-san** čin' iž-lūd n'
herd-GEN among **go-NP** PTC alike-PL POSS3SG

өшиглөж *алчих* *гээд* ***явуулсангүй***.
öšiglö-ž al-či-x ge-ēd **yav-ūl-san-güi**.
kick-CI kill-INT-NF say-CP **go-CAUS-NP-NEG**

'While he was **running** inside the herd, the horses were kicking him, trying to kill him and **did not let him run**.'

(70) *Адууныхаа* *захад* *нь* *гараад* ***явсан***
adūn-ï-xā zax-ad n' gar-ād **yav-san**
herd-GEN-POSS.REF border-DL POSS3SG go.out-CP **go-NP**

аятай *сайхан* ***явдаг*** ***болов*** *гэнэ.*
ayatai saixan **yav-dag** **bol-ov** ge-ne.
pleasant nice **go-NU** **become-PRET.PERF** say-IMPRF
'As he **(learned to) stay** to the side of the herd, he finally **began enjoying** an agreeable **life.**'

The next example is from a folk song. The four rhymes display syntactic parallelism, which is a typical device in Mongolian poetry. In the first and the third lines, as a basis of comparison, the verb *yav-* is used in its more literal meaning of movement, whereas in the second and fourth lines *yav-* functions as a modal verb: its usage is figurative, referring to certain aspects of life and human behaviour.

(71) *Уулаар* ***явдаг*** *угалзын* *зан,*
ūl-ār **yav-dag** ugalz-īn zan,
mountain-INS **go-NU** ibex/muflon-GEN character
'**To roam** in the mountains is the ibexes' custom,

(72) ***Уулзаад*** ***явдаг*** *хүний* *зан,*
ūlz-ād **yav-dag** xün-ī zan,
meet-CP **go-NU** person-GEN character
'**To live meeting others** is a human custom,

(73) *Үүрээр* ***явдаг*** *үхрийн* *зан,*
ūr̄-ēr **yav-dag** üxr-īn zan,
dawn-INS **go-NU** cattle-GEN character,
'**To walk** at dawn is the habit of cattle,

(74) ***Үерхэж*** ***явдаг*** *багын* *зан.*
üerxe-ž **yav-dag** bag-īn zan.
be.friends-CI **go-NU** little-GEN character.
'**To become close (friends**) is the habit of children/youth.'

5. Conclusion

The main semantics of the Khalkha Mongolian verb *yav-* are connected to motion in the general sense. It refers to motion in space with either no concrete spatial definition (**1.1**), or with the emphasis on departing from a certain location (**1.3**). Even when the goal, course and manner of motion are specified in various ways, the emphasis is still placed on the process of motion itself. In more general or abstract usage, the verb *yav-* often represents the meaning of an entire journey with a specific goal or purpose (**1.2**). Syntactically, when the spatial goal of motion is expressed, it does not bear the dative-locative case (used mainly to express the purpose of motion), but, instead, the lative case (or postpositional construction), which corroborates the idea of the lack of inherent spatial features of this verb.

Similarly to other verbs of motion, the verb *yav-* is used as a component of complex descriptive designations of actions, which is a typical feature of most Altaic languages. The scale of usage options, however, is relatively small, being mostly limited to the expression of one of the basic lexical meanings (**2.4.1**) or the spatial meaning of motion away from a certain point (**2.4.2**).

Likewise, the Mongolian verb *yav-* displays, in comparison to other verbs of motion, relatively few modes of modal usage, which may be the result of its lack of concrete spatial and deictic features. The two main modal meanings of the verb, stemming from its two basic "semantic cores", convey the meaning of durativity or continuity of an action (**2.1**), and the modality of intense or abrupt change (**2.2**).

The material used for this study yields examples of usage which could be labelled as auxiliary in the narrower sense (i.e. having predominantly grammatical function) (**2.3**). This feature, absent in the cases of other verbs of motion, reflects the verb's semantic connection to existence.

In its basic figurative meaning, the Mongolian verb *yav-* refers to life, existence and its various modes and manners. While, generally speaking, in many cultures, motion is perceived as the representation of life, in the traditional Mongolian nomadic culture, where constant movement is the basic condition of survival, the association of motion with life seems to be even more implicit and essential. Hence, the figurative meaning of the verb *yav-* covers almost all spheres connected to existence, including staying, dwelling, making one´s living, behavior, degree of success in a given endeavor, and so on. The boundaries between the literal and figurative meanings are often difficult to distinguish when these aspects of life happen to overlap with the actual process of motion, as shown in section **3.7**.

In my two previous papers (Zikmundová 2009 and 2010) I examined two verbs of motion, whose meaning may be roughly translated as 'to come here' and 'to go there/to visit', in both Khalkha Mongolian and Jungarian Sibe. In both languages these verbs are antonyms and form part of a coherent system of verbs of motion with distinct spatial meanings and deictic functions.

The verb *yav-*, while being one of the basic verbs of motion, does lack some features usually characteristic of verbs of this system. It does not convey such inherent spatial meaning. Instead, its greater ambiguity with regard to space, as well as its main semantic contours, both resemble the semantics of equivalent verbs in European ("satellite-framed") languages, in particular, the English verb 'to go'.

Abbreviations

ABL	Ablative
ACC	Accusative
CA	Converbum abtemporale
CAUS	Causative
CC	Converbum conditionale
CI	Converbum imperfecti
CM	Converbum modale
CP	Converbum perfecti
CT	Converbum terminale
DL	Dative-locative
GEN	Genitive
IMP	Imperative
IMPRF	Imperfective verb
INS	Instrumental
INT	deverbal suffix of intensity
LAT	Lative
NEG	Negative
NF	Nomen Futuri
NI	Nomen Imperfecti
NU	Nomen usus
PAS	Passive
PL	Plural
POSS	Possessive
POSS.REF	Reflexive possessive suffix
PRES.PERF	Presens Perfecti
PRET.IMP	Preteritum Imperfecti
PRET.PERF	Preteritum Perfecti
PROH	Prohibitive particle
PTC	Particle
Q	Question marker
SOC	Sociative
TOP	Topic particle
VOC	Vocative
VOL	Voluntative
VP	Verbum pluritativum

References

Levin, Beth and Rappaport Hovav, Malka 1992. The lexical semantics of verbs of motion: The perspective from unaccusativity. In: Iggy, Rocca (ed.) *Thematic structure: its role in grammar*. Berlin: Foris. 247–269.

Lubsangdorji, Jugderiin and Vacek, Jaroslav 2004. *Colloquial Mongolian: an introductory intensive course*. Prague: Triton.

Nakazawa, Tsuneko 2007. A typology of the ground of deictic motion verbs as path conflation verbs: The speaker, the addressee, and beyond. *Poznań Studies in Contemporary Linguistics* 43/2. 59–82.

Nakazawa, Tsuneko 2009. Typology of the path of deictic motion verbs as path conflation verbs: the entailment of arrival and the deictic center. *Poznań Studies in Contemporary Linguistics* 45/3. 385–403.

Slobin, Dan I. 1996. Two ways to travel: the verbs of motion in English and Spanish. In: Shibatani and Thompson, S. A. (eds.) *Grammatical constructions: their form and meaning*. Oxford: Clarendon Press. 195–219.

Talmy, Leonard 1975. Semantics and syntax of motion. In: Kimball, J. P. (ed.) *Syntax and semantics* 4. New York: Academic Press. 181–238.

Talmy, Leonard 1985. Lexicalization patterns: semantic structure in lexical forms. In: Shopen, T. (ed.) *Grammatical categories and the lexicon: Language typology and syntactic description* 3. Cambridge: Cambridge University Press. 57–149.

Talmy, Leonard 2000. *Toward a Cognitive Semantics. Volume 1: Concept structuring systems*. Cambridge, MA: MIT Press.

Word-Allbritton, A. 2004. The Turkmen verb system: Motion, path, manner and figure. *IULC Working Papers Online. Indiana University and The University of Alabama – Huntsville*. https://www.indiana.edu/~iulcwp/pdfs/04-word.pdf.

Zikmundová, Veronika 2010. The function of descriptive verbs in colloquial Sibe 1. The basic pair of verbs of motion *ji*- ‘to come’ vs. *gen*- ‘to visit’. *Mongolo-Tibetica Pragensia ’10*, 3/2. *Ethnolinguistics, Sociolinguistics. Religion and Culture*. Praha: Triton. 95–123.

Zikmundová, V., 2011. The function of descriptive verbs in Khalkha Mongolian 1. The basic pair of verbs of motion *ir*- ‘to come’ vs. *oč*- ‘to visit’. *Mongolo-Tibetica Pragensia ’11*, 4/2. *Ethnolinguistics, Sociolinguistics, Religion and Culture*. Praha: Triton.

Zikmundová, V. 2013. The Sibe verb *yaf-*. *Mongolo-Tibetica Pragensia '13,* 6/1. *Linguistics, Ethnolinguistics, Religion and Culture*. Praha: Charles University and Triton.

Database of Mongolian folklore texts at http://www.signeta.cz/textsearch/

Tungusic loanwords in Yeniseian languages*

Bayarma Khabtagaeva

University of Szeged

1. Introduction

The topic of my current research is the exploration of Altaic elements in the Yeniseian languages[1]. Previous studies only focused on the Turkic elements,[2] whereas the Mongolic and Tungusic loanwords have not been discussed yet.

The paper presents 23 different Tungusic loanwords in Yeniseian from the etymological, phonetic, morphological and lexical aspects. In the current corpus of my research I deal with 124 loanwords with clear Tungusic etymology and I have 30 questionable words in terms of origin. The source of borrowing for Yeniseian languages is the Ewenki dialects, which belong to the Northern Tungusic group.[3]

* I would like to extend my gratitude to Professor Edward Vajda for his valuable remarks and teaching me historical Yeniseian linguistics from April to June in 2012 and in March 2014 when I had an opportunity to carry out research at the Center for East Asian Studies, Western Washington University, in the USA.

[1] Yeniseian languages belong to the Paleo-Siberian language group. The Yeniseian languages are claimed to be related to the Sino-Tibetan, Burushaski (Karasuk) and Caucasian language families. Recently Vajda (2010) has presented a hypothesis that the Yeniseian languages show genealogical connections with the Na-Dené languages of North America, but this question remains open for further discussion. According to the most recent works on historical linguistics such as Starostin (1982), Vajda and Werner (in preparation), and Vajda (personal communication) the Yeniseian languages are divided into at least three sub-branches: Ket-Yugh, Pumpokol, Assan-Kott, with Arin either connected with Pumpokol or Ket-Yugh or representing a fourth sub-branch. Today the Yeniseian language family is represented only by the three surviving dialects of Ket.

[2] See the papers by Timonina (1986; 2004) and Stachowski (1996; 1997). The latter also discussed the Turkic loanwords of Arabic origin (Stachowski 2006). Vajda (2009) is a valuable paper on different loanwords (of Russian, Uralic and Altaic origin) in Ket.

[3] The Ewenki language belongs to the Tungusic languages, traditionally believed to form the Altaic language family together with the Turkic and Mongolic languages, and the northern

Morphologically most loanwords are either nouns or adjectives. In addition, there are also small numbers of verbs, adverbs and particles.

The main source of my research was the *Comparative Dictionary of Yeniseian languages* by Werner (2002/1-3) and his study on 18th century Yeniseian materials (Werner 2005). In these works we can find the whole lexical material of Yeniseian languages published until now. Another source of my work was the *Etymological Dictionary of Yeniseian languages* by Vajda and Werner, which is still at a preparatory stage (Vajda and Werner: in preparation).

2. Etymology

2.1. The Tungusic loanwords with clear etymology

(1) Ket *dɔgbən* 'area between two riverbends' (Vajda and Werner: in preparation) ← Northern Tungusic: Ewenki *dāgwūn* 'crossing, ford across a river' < *dāγ-* 'to cross, pass river' *-wūn* {Ewenki VN}:

> cf. Northern Tungusic: Lamut; Negidal *daw-*; Southern Amuric: Oroch; Udihe; Ulcha *dau-*; Orok *dāu-*; Nanai *dā-*; Southern Manchuric: Jurchen – ; Manchu *dō-* (SSTMJa 1: 187).

The Ket word was obviously borrowed from Ewenki form *dāgwūn*, which was derived from the Common Tungusic verb *dāγ-* 'to cross, pass river' and productive Ewenki VN suffix *-wūn* (On function see Vasilevič 1958: 748). The borrowing from the Ewenki dialect confirms the absence of this suffix in other Tungusic languages.

(2) Northern Ket *kɔlɔmɔ* ~ *gɔ́lɔmɔ* 'a kind of winter dwelling covered with sod' (Werner 2002/1: 439) ← Northern Tungusic: Ewenki *golomo* 'a kind of winter dwelling' < *golo* 'a log, a beam' *+mA* {Ewenki NN}:

> cf. Northern Tungusic: Ewenki dial. *golomo*; Lamut *goloma* < *gol* 'firewood'; Negidal *golo* 'log, beam'; Southern Amuric: Oroch *golo*; Udihe

Tungusic branch, together with the Even (or Lamut) and the Negidal languages. The southern branch is divided into two groups: the Manchuric group (Jurchen or Old Manchu, Manchu and its sole living remnant Sibe ~ Sibo) and the Amuric group (Nanai, Ulcha, Orok, Oroch and Udehe). The Ewenki language in Russia has 51 dialects, which can be divided into northern, southern and eastern groups. The main criterion used during the classification of the dialects is the development of the Common-Tungusic consonant **s* in initial and intervocalic positions. The representations *h, s* and *š* appear in the three groups. Literary Ewenki is based on the Poligus dialect, which belongs to the southern or sibilant group, exhibiting the hissing type (*s-*, *VsV*) (For more details on classification, see Khabtagaeva 2010: 10; 12–13).

golo ‘half-rotten log’; Ulcha *goloŋqo* ‘firewood, stack’; Orok *golo* ‘log’; Nanai *ɣoloŋqo* ‘firewood, stack’; Southern Manchuric: Jurchen – ; Manchu *ɣoldon ~ ɣolton* ‘charred ends of wood, charred stump’; Sibe - (SSTMJa 1: 159b).

The Northern Ket term connected with the kind of winter dwelling has clear Tungusic etymology. The Yeniseian word was borrowed from Ewenki, where it was *golomo,* derived from the Common Tungusic word *golo* ‘a log, a beam’ and productive Ewenki suffix *+mA* (Vasilevič 1958: 769). The devoicing of the Tungusic initial **g-* is a regular change in Ket, which characterizes the early period of borrowing. The initial *g-* is typical only of loanwords in Ket.

2.2. The Tungusic loanwords of questionable etymology

The next Ket word belongs to the category of uncertain etymology, with some problematic aspects. E.g.

(3) Ket *ɛretńek ~ ɛr’etńek* ‘devil’ (Werner 2002/1: 240) ← Northern Tungusic: Ewenki *erūŋit-* ‘to do an evil deed, crime; to say spiteful things; to be squeamish’ < *erū* ‘trouble, adversity; mischief; guilt; harm’ *+ŋi-* {Ewenki NV, see Vasilevič 1958: 780} *-t-* {Ewenki VV, see Vasilevič 1958: 790} ← Mongolic **erūl̄* < *eregü* ‘torture, torment, chastisement; capital punishment’:

cf. Northern Tungusic: Ewenki dial. *erū* ‘trouble, adversity; mischief; guilt; harm’ (Bold.); Lamut – ; Negidal – ; Southern Amuric: Oroch – ; Udihe – ; Ulcha *erule-* ‘to torture; criticize, oppress’; Orok - ; Nanai *erū* ‘torment’; Southern Manchuric: Jurchen - ; Manchu *erun* ‘torture; execution’; Sibe *erulu-* ‘to torture’ (SSTMJa 2: 465b-466a) ← Mongolic **erūl̄* < *eregü* ‘torture, torment, chastisement; capital punishment’: Middle Mongolic: – ; Literary Mongolian *eregüü*; Modern Mongolic: Buryat – ; Khalkha *erūl̄*; Oyrat dial. – ; Kalmuck –.

The Ket word is of unknown etymology. As a hypothesis, I propose that it originates from the Ewenki dialectal verbal form *erūŋit-* ‘to do an evil deed, crime; to say spiteful things’, which is of Mongolic origin (Doerfer 1985: 39; Rozycki 1994: 71). I suppose the Ket taboo word was borrowed from Ewenki and underwent metathesis, which is a peculiar feature of Yeniseian: **erūngit- > *erutnig > ɛretńek* (On metathesis in Yeniseian, see Vajda 2013).

2.3. Tungusic loanwords of Mongolic origin

During my research I found some Ket words of Mongolic origin also exhibiting Tungusic peculiarities, e.g.:

(4) Ket *ába ~ áva* ‘shelf in a tent for storing kitchen utensils’ (Werner 2002/1: 12) < **alba* < **yalba* ← Northern Tungusic: Ewenki **ǰalba* < *dalba* ‘shelf for kitchen utensils’ ← Mongolic **talbūr* < *talbiγur* ‘stand, rest; hanger; low table; board to put things on; to set free, release, let loose’ < *talbi-* ‘to place, put, set, lay or put down; to install’ *-GUr* {Poppe GWM §155}

> cf. Northern: Ewenki dial. *dalba ~ ǰalba* (Bold.), cf. *dalbaptun, dalbur* ‘id.’; Southern Amuric: - ; Southern Manchuric: - (SSTMJa 1: 193b) ← Mongolic: cf. Middle Mongolic: MNT *talbi-*; HY *talbi-*; Muq *talbi- ~ tabi-*; Leiden *talbi-*; Literary Mongolian *talbiγur* < *talbi- ~ tabi-*; Modern Mongolic: Buryat *tabyūr* ‘stand; small table’; Khalkha *tawiur* ‘stand, rack, shelf, easel, stage’; Oyrat dial. *täwῡ̄r ~ täür* ‘cupboard; shelf’ (For more Mongolic data, see Nugteren 2011: 510-511).

The Tungusic source of borrowing is proven by the initial *y-*, which alternates with *ǰ- ~ d-* in the Ewenki dialects. The Ket word was likely borrowed from Ewenki, where it was the dialectal form *yalba*, where the initial *y-* disappeared. This phonetic change occurs in some Tungusic loanwords, e.g. Ket *enna* ‘really?’ ← Northern Tungusic: Ewenki *yēŋan* ‘what; how; really?’; Central Ket *aqtul* ‘spring (water coming out of the ground)’ ← Northern Tungusic: Ewenki *jūkte* ‘spring, brook’, etc. Another important phonetic feature in Yeniseian is the prohibition on consonant clusters, e.g. Ket *mina* ‘pig’ ← Russian *svin’ja*; Ket *kuruk* ‘hook’ ← Russian *kr’uk*; Ket *kola* ‘school’ ← Russian *kola*, etc. (Vajda 2009: 486).

(5) Ket *bʌˀj* ‘friend’ (Werner 2002/1: 156) ← Northern Tungusic: Ewenki *beye* ‘man, person, body’ ← Mongolic *beye* ‘body, physique, organism; health’:

> cf. Northern: Ewenki dial. *beye* ‘the man; male; husband; personality; the body; referring to a man or a woman; generation; age (~ 75 years)’; Lamut *bej* ‘man, personality’; Negidal *beye* ‘man; body; personality; oneself’; Southern Amuric: Oroch *beje* ‘body; oneself; similar’; Udihe *beje* ‘body; oneself; present, real’; Ulcha *beje* ‘body, trunk; oneself’; Orok *beje* ‘body, trunk; stem’; Nanai *beje* ‘body, trunk; corpus; figure; oneself, personal, own, proper’; Southern Manchuric: Jurchen *péi-yè* ‘body’; Manchu *beje* ‘body, trunk; life, being; personality; oneself, own, personal’ (SSTMJa 1: 122a) ← Mongolic *beye* ‘body, physique, organism; health’: Middle Mongolic: MNT

beye ~ be'e; HY *beye*; Muq *beye*; Leiden *biye*; Literary Mongolian *beye(n)*; Modern Mongolic: Buryat *beye*; Khalkha *biye*; Kalmuck *biy* (For more Mongolic data, see Nugteren 2011: 281).

The Tungusic borrowing of Ket word confirms the lexical meaning. Originally in Mongolic it means 'body, physique, organism; health', while in Tungusic it means 'man, person, body'. It seems that the Mongolic word was borrowed early, as it is present in almost all Tungusic languages.

2.4. Hybrid words

An independent group consists of hybrid words, where one of the elements is Tungusic, while the other one is Yeniseian, e.g.:

(6) Northern Ket *aʁidɛ* 'marsh, tundra'; Central Ket *ajgiddɛ*; Southern Ket *ajgitdɛ* 'wooded tundra, pine bog' (Werner 2002/1: 85) ← Northern Tungusic: Ewenki *ayī* 'taiga, tundra, marsh' + Yeniseian *deˀ* 'lake':

cf. Northern: Ewenki dial. *ayī ~ ajī* 'taiga', cf. Stony Tung. 'forest in the plain', Barguzin 'open desert place, steppe', Aldan 'field', Tungir. 'tundra, marsh'; cf. *ayī-* 'to walk in the snow (without a road, without skis); to go past; to hunt'; Lamut *āju- ~ āwi-* 'to walk in the snow'; Negidal *awī-*; Southern Amuric: Udihe *ai-* 'to go through deep snow'; Ulcha *ajī- ~ u-* 'to walk in the snow (without a road, without skis)'; Orok *āwi-* 'to walk in the snow'; Nanai *aoi-* 'to go through deep snow', *āi-* 'to walk in the snow'; Oroch - ; Southern Manchuric: Jurchen - ; Manchu *ajli-* 'to dodge a straight road'; Sibe - (SSTMJa 1: 13a).

The Ket forms probably consist of the Common Tungusic *ayī* 'taiga, tundra, marsh' and Yeniseian word *deˀ* 'lake'.

(7) Ket *dankïjaj* 'rucksack' (Vajda and Werner: in preparation) ← Northern Tungusic: Ewenki *daŋajā* 'shoulder bone' + Yeniseian *aj* 'sack':

cf. Northern: Ewenki dial. *daŋanī ~ daŋańā ~ daŋajā ~ dayańa* 'name of bone (shoulder bone, hip bone, shin bone); thigh'; Lamut - ; Negidal *dayańa* 'hip bone'; Southern Amuric: Oroch - ; Udihe - ; Ulcha - ; Orok *dāna* 'name of bone (shoulder bone, hip bone, shin bone); thigh; shoulder (upper arm until elbow)'; Nanai - ; Southern Manchuric: - (SSTMJa 1: 188b).

2.5. Tungusic loanwords with Yeniseian suffixes

There are some Ket words which were borrowed from Tungusic but are used with native Yeniseian suffixes. They bear plural, collective or nominalizer suffixes, e.g.:

(8) Central Ket *kʌdaŋ ~ kʌdəŋ* ‘marshy place’ (Werner 2002/1: 459) ← Northern Tungusic: Ewenki *kuta* ‘marsh’ *+ŋ* {Yeniseian Collective suffix}:

> cf. Northern:[4] Ewenki dial. *kuta* ‘marsh, bog, clay’, cf. *kuta-* ‘to get bogged down in the swamp’; Lamut *kuta* ‘bog, swamp’; Negidal *kota* ‘bog, swamp’; Southern Amuric: – ; Southern Manchuric: – (SSTMJa 1: 439b).

(9) Northern Ket *dɔktɔraŋ* ‘socks made from reindeer hide’ (Werner 2002/1: 195) < *dɔktɔr +aŋ* {Ket Plural} ← Tungusic *doktokōn* ‘fur stockings’< *doqto +KĀn* {Ewenki NN/Diminutive}:

> cf. Northern: Ewenki dial. *doktokōn ~ dektekēn* ‘fur shoes, worn on boots’; Lamut *dōten* ‘fur stockings’; Negidal *dokton* ‘stockings (fur, cloth, cotton)’; Southern Amuric: Oroch *dokton* ‘fur or leather stockings’; Udihe *dokti* ‘fur stockings’; Ulcha *doqto* ‘fur or cotton stockings’; Orok *doqto* ‘fur or cotton stockings’; Nanai *doqto* ‘fur or cotton stockings’; Southern Manchuric: – (SSTMJa 1: 213a).

The Ket word is obviously connected with the Common Tungusic word *doqto* ‘fur stockings’, which is present in almost all Tungusic languages. The Ewenki borrowing proves the Diminutive suffix *+KĀn* (see Vasilevič 1958: 759). Unlike the Tungusic form, the Yeniseian form acquires the plural suffix, and the *-k-* > *-r-* change occurred.

3. Phonetic peculiarities

3.1. Syncope

Ket words are usually monosyllabic, and syncope is typical of polysyllabic loanwords, e.g. Russian *nedel’a* ‘week’ in Ket is *nela*; Russian *samovar* in Ket is *sambar*, etc. (see Vajda 2009: 486). There are some Yeniseian examples where syncopation occurred:

(10) Ket *kɔɣɔ́n* ‘myth. forbidden jewelry in a snake’s nest; copper pendant of shaman’s costume’ (Werner 2002/1: 445) ← Northern Tungusic: Ewenki *kulitkān* ‘the image of snake in the shaman’s costume’ < *kulīn* ‘snake’ *+tkĀn* {Ewenki NN/Diminutive}:

> cf. Northern: Ewenki dial. *kulitkān* ‘the image of snake in the shaman’s costume’< *kulīn* ‘snake’ (Bold.); Lamut *qulin ~ quličān ~ qolisān ~ kuličan ~ quličān* ‘mosquito’; Negidal *kolixān ~ kulikān* ‘worm, bug’; Southern

[4] Northern Tungusic → Turkic: Yakut *kuta* ‘bog, peat’ (SSTMJa 1: 439b).

Amuric: Oroch *kulæ* ‘worm (common name for worms, snakes and caterpillars)’; Udihe *kuliga* ‘id.’; Ulcha *qoli* ‘kind of aquatic insect’, *qula* ‘worm’; Orok *qola* ~ *qolia* ~ *qoliɣa* ‘insect, worm’; Nanai *qolã* ‘worm; caterpillar; insect’; Southern Manchuric: - (SSTMJa 1: 428b).

The Ket word was likely borrowed from the Ewenki, where it is a diminutive form *kulitkān* (on its function, see Vasilevič 1958: 791). The Tungusic origin proves the lexical meaning of the base *kulīn* ‘snake’.

(11) Ket *áʁses* ‘bear trap’ (Werner 2002/1: 56) < **áʁse* *+s* {Yeniseian nominalizer} ← Northern Tungusic: Ewenki *amākākse* ‘bear’s skin; bear’s flesh’ < *amā* ‘father’ *+kā* {Ewenki NN/Diminutive} *+kse* {Ewenki NN/Adj.}:

cf. Northern Tungusic: Ewenki *amākākse* ‘bear’s skin; bear’s flesh’; cf. Negidal *amaj* ‘father’; Southern Amuric: Oroch, Udihe, Ulcha, Nanai *ama*; Orok *ama* ~ *amma*; Southern Manchuric: Jurchen *ʻá-mînʼ*; Manchu *ama* (SSTMJa 1: 34b-35a).

Possibly the Ket word used with Yeniseian nominalizer *+s*. It was likely borrowed from Tungusic form *amākākse* ‘bear’s skin; bear’s flesh’, which is derived from the Common Tungusic word *amā* ‘father’ with the Ewenki diminutive suffix *+kā* (Vasilevič 1958: 758) and the Ewenki productive denominal noun suffix *+kse*, deriving the adjective forms (Vasilevič 1958: 763). This is a good example, where the Tungusic taboo word ‘bear’ was originally formed from the word with the meaning ‘father’.

3.2. Aphaeresis

Another important typical phonetic feature of Tungusic loanwords in Yeniseian is aphaeresis. There are some Ket words where the Tungusic initial *jX-* regularly disappeared, e.g.:

(12) Ket *enna* ‘really?’ (Vajda and Werner: in preparation) ← Northern Tungusic: Ewenki *yēŋan* < *yē* ‘what; how; really?’ *+ŋĀn* {Ewenki NN}, cf. Common Tungusic *yē-* ‘quest. verb what to do?’:

cf. Northern: Ewenki dial. *jēŋan* ‘really?’; Lamut *ā̄-* ~ *ē-* ~ *ʲā̄-* ‘[quest. verb] what to do?’; Negidal *ē-* ‘id.’; Southern Amuric: Oroch *ya-* ~ *ye-* ‘[quest. verb] what to do?’; Udihe *ya-* ~ *ye-* ‘id.’; Ulcha – ; Orok – ; Nanai – ; Southern Manchuric: Jurchen – ; Manchu *ya* ‘what, who; which, which kind?’; Sibe *ya* ‘id.’ (SSTMJa 1: 286).

The Ket adverb was clearly borrowed from Ewenki, where it is *yēŋan*. The base of the word is the Common Tungusic interrogative pronoun *yē* and Ewenki NN *+ŋĀn*

(on its function, see Vasilevič 1958: 778). Etymologically the Tungusic word is probably connected with the Mongolic interrogative pronoun *yaγun* ‘What? What kind? Which?’.

(13) Central Ket *aqtul* ‘spring (water coming out of the ground)’ (Werner 2002/1: 55) < *aqtu* + ? *l* ← Northern Tungusic: Ewenki *jūkte* ‘spring, brook’[5] < *jū-* ‘to go out, come out’ *-ktA* {Ewenki VN}:

cf. Northern: Ewenki dial. *yūkte ~ ǰūkte ~ jūktu ~ ńūkte* ‘spring, brook’ < *jū- ~ ǰū- ~ ńū-*; Lamut *ńō̄- ~ jō- ~ ńu-* ‘to get out; to leave; to rain’; Negidal *jū- ~ ńū-* ‘to go out; to float’; Southern Amuric: Oroch *ńū-* ‘to get out, to pop out; to rise (sun)’; Udihe *ńū-*; Ulcha *ńie- ~ ńē-*; Orok *nē-*; Nanai *ńie- ~ ńē- ~ ńiu-*; Southern Manchuric: – (SSTMJa 1: 348b-349a).

The Ket word originated from the Common Tungusic verb *jū-* ‘to go out, come out’. The borrowing from Ewenki is evident from the presence of the Ewenki productive deverbal noun suffix *-ktA* (on its function, see Vasilevič 1958: 764), which is absent in other Tungusic forms.

3.3. Excrescence of internal *-ŋ-*

The next phonetic feature which is typical of Tungusic elements, is excrescence. There are some words where the internal *-ŋ-* usually appeared usually before the dental consonant *-t-*, e.g.:

(14) Ket *daŋtakan*; Northern Ket *deŋtiɣin ~ däŋtiɣin* ‘marsh’ (Werner 2002/1: 184) ← Northern Tungusic: Ewenki *detkēn* ‘marsh’ < *det* ‘tundra, marsh’ *+kĀn* {Ewenki NN/Diminutive, see Vasilevič 1958: 759}:

cf. Northern: Ewenki *detkēn* ‘tundra; marsh; mossy glade’ (Bold.); Lamut *det*; Negidal *det ~ detkēn*; Southern Amuric: Oroch *detu*; Udihe – ; Ulcha *detu*; Orok *detu*; Nanai *detu*; Southern Manchuric: – (SSTMJa 1: 238b).

(15) Ket *laŋtei ~ laŋtegai ~ laŋteŋai* ‘evil’ (Werner 2002/2: 3) < **lagataki* ← Northern Tungusic: Ewenki *lawadapki* ‘evil spirit’ < *lawādā-* ‘to carry, haul something by mouth; to grab by beak or teeth’ *-pki* {Ewenki VN}:

cf. Northern: Ewenki dial. *lawadapki* (Bold.); Lamut – ; Negidal *lawādā-* ‘to carry, haul something by mouth (about animal)’; Southern Amuric: – ; Southern Manchuric: – (SSTMJa 1: 485b).

[5] Northern Tungusic → Turkic: Yakut *ǰūkta* ‘spring (water coming out of the ground); polynya (unfrozen patch of water in the midst of ice)’ (SSTMJa 1: 348b–349a).

The Ket forms *laŋtegai ~ laŋteŋai* ‘evil’ are problematic. As a hypothesis, I assume that this Ket word has Tungusic etymology. The borrowed form was possibly **lagataki*, which can be connected with the Ewenki word *lawadapki* ‘evil spirit’. And finally, metathesis occurred in Ket. Another important fact for Tungusic corroboration is the initial *l-*, which is peculiar to loanwords in Ket.

3.4. The final *-l* of unknown origin

During my research I found several Tungusic loanwords with a final consonant *-l* of unknown origin. This must be a Ket peculiarity, but there is no any information about this feature. E.g.

(16) Ket *e²tl* ‘colour’ (Vajda and Werner: in preparation) < **e²t +l* ← Northern Tungusic: Ewenki *ičede* ‘colour’ < *iče-* ‘to see, watch, look, notice’ *-dA* {Ewenki VN, see Vasilevič 1958: 752}:

cf. Northern: Ewenki *ičeden ~ ičede ~ ičedi* ‘view, look, appearance; reflection; color’ < *iče-* (Bold.); Lamut *it-* ‘to see, notice’; Negidal *iče-*; Southern Amuric: Oroch *iče-*; Udihe *isesi-*; Ulcha *iče-*; Orok *it- ~ ite-*; Nanai *iče-*; Southern Manchuric: – (SSTMJa 1: 334b-335b).

(17) Ket *saŋɔl* ‘chimney of a dug-out’, cf. Central Ket *sɔnal* ‘smoke hole of a dug-out’ (Vajda & Werner: in preparation) < **sona +l* ← Northern Tungusic: Ewenki *sōna* ‘chimney’:

cf. Northern: Ewenki dial. *sōna ~ sōŋa ~ suona ~ hōna ~ šōna* ‘smoke hole of a dug-out; chimney of a dug-out’; Lamut *hōnan ~ hōnān* ‘smoke hole of a dug-out; rafter’; Negidal *sōna* ‘id.’; Southern Amuric: Oroch *sōno* ‘smoke hole of a dug-out’; Udihe - ; Ulcha *sōn* ‘roof rafter’; Orok *sōno* ‘smoke hole of a dug-out; rafter’; Nanai *sõ* ‘roof rafter; pole’; Southern Manchuric: Jurchen - ; Manchu *son* ‘pole’; Sibe - (SSTMJa 2: 110).

4. Morphology

From a morphological point of view, the majority of loanwords are nouns and adjectives. Additionally, a small number of loanwords are verbs, adverbs and particles.

4.1. Verbs

The Ket verbal system is highly complicated, and it is not typical of Ket to borrow verbal forms (for details on the Ket verbal system, see Vajda 2004; Vajda in press). Some verbs were easily recognizable, e.g.:

(18) Ket *alepqaj* 'inf. flare up, fly into a rage' (Werner 2002/1: 25) ← Northern Tungusic: Ewenki *alipkī-* 'to be angry' < *ali-* 'to be angry, to anger, make angry' -*pkī* {Ewenki VV, see Vasilevič 1958: 784}:

> cf. Northern: Ewenki *ali-* 'to be angry, to anger, make angry' (Bold.); Lamut *alel-* ~ *ālil-* 'to be angry'; Negidal *ali-* 'to be angry, hate'; Southern Amuric: – ; Southern Manchuric: – (SSTMJa 1: 32ab).

(19) Ket *ɔq-tɛt* 'to come loose and fall down' (Vajda and Werner: in preparation) ← Northern Tungusic: Ewenki *ugďe-* 'to become lean, gaunt':

> cf. Northern: Ewenki *ugďe-* 'to become lean, gaunt', cf. *ugďe* 'lean, gaunt'; Lamut – ; Negidal *ugdexe* ~ *ugduxo* 'few; small; weak'; Southern Amuric: Oroch – ; Udihe – ; Ulcha – ; Orok *ugde* 'quiet, slow, weak'; Nanai *ugǰe-* 'vanish, disappear (about fish in lake)'; Southern Manchuric: – (SSTMJa 2: 245a).

The Ket verb has unclear etymology. It likely originates in the Tungusic verb *ugd'e-* 'to become lean, gaunt'. Here the Tungusic etymology strengthens the lexical meaning.

4.2. Other parts of speech

Besides verbs, a small number of Tungusic adverbs and particles were borrowed in Ket, e.g.:

(20) Ket *báŋa* 'under no circumstance, never' (Werner 2002/1: 102) ← Northern Tungusic: Ewenki *bān* 'refusal, repudiation; failure' < *bā-* 'to be unable, to resist, to refuse' -*n* {Ewenki VN, see Vasilevič 1958: 777}:

> cf. Northern: Ewenki *bān* 'refusal, repudiation; failure' < *bā-* 'to be unable' (Bold.); Lamut *bā-* 'to be lazy, to sit back'; Negidal – ; Southern Amuric: Oroch *bāki* 'lazy'; Udihe – ; Ulcha – ; Orok *baja* ~ *bāju* 'lazy, idler, loafer'; Nanai *bāqi* 'lazy, loafer'; Southern Manchuric: – (SSTMJa 1: 60b-61a).

Usually the source of borrowing is Ewenki, but in some cases it is questionable because of the word's absence in Ewenki. The Ket example of the intensive negative particle *a:na* 'not even, don't even' probably has a Tungusic origin. But despite the

fact that the word is present in almost all Tungusic languages, it is absent in the Ewenki dialects:

(21) Ket *aːna* ‘intensive negative particle (not) even, (don’t) even’ (Werner 2002/1: 34) ← Common Tungusic *ana* ‘intensive negative particle (not) even, (don’t) even’:

cf. Northern: Ewenki – ; Lamut *ān ~ āŋ ~ jān* ‘not having someone or something; without someone or something’; Negidal – ; Southern Amuric: Oroch ana ‘not, missing’; Udihe – ; Ulcha *ana* ‘net’; Orok *ana ~ anā ~ anaɣa* ‘not having someone or something; not, missing’; Nanai *anā* ‘not, missing’; Southern Manchuric: – (SSTMJa 1: 41a).

5. Lexicology

From a semantic point of view, the meaning of the loanwords taken from Tungusic is usually preserved. There appear to be some cases in which a change in meaning occurs. There are two Ket words, which belong to the taboo category:

(22) Central Ket, Northern Ket *ulla ~ úlle* ‘ruble, the monetary unit in Russia’ (Vajda and Werner: in preparation) ← Northern Tungusic: Ewenki *ulukī* ‘squirrel’:

cf. Northern: Ewenki dial. *ulukī*; Lamut *öliki ~ úlikí ~ uliki ~ ul’ki*; Negidal *ölöxī ~ ölukī ~ eluxi*; Southern Amuric: Oroch *oloki*; Udihe *oloxi*; Ulcha *xolo*; Orok *xolo ~ xulu*; Nanai *xulu ~ uluki*; Southern Manchuric: Jurchen - ; Manchu *ulhu* ‘squirrel; ermine; squirrel fir’ (SSTMJa 2: 263-264).

The Ket word likely has a Tungusic etymology; it was borrowed from Ewenki, where it is *uluki* ‘squirrel’ form. My assumption can be confirmed by the fact that the price of squirrel hides in the early 20th century was one ruble (Dolgikh 1934: 91). It is a well-known fact that Tungusic, Turkic and other native Siberian people paid fur animals like squirrel and sable as tribute and tax to the Russians.

(23) Southern Ket *ɯjiŋes* ‘spring’ (a euphemism for the taboo word *ɯr̀* ‘spring’) (Werner 2002/2: 421) < Tungusic: Ewenki *ije* ‘horn’ +*ŋ* {Yeniseian plural} + Ket *ēːs* ‘weather’:

cf. Northern: Ewenki *ije* ‘horn’; Lamut *īj ~ īje*; Negidal *īje*; Southern Tungusic: Oroch *ije*; Udihe *jē*; Ulcha, Orok *huje*; Nanai *hujī*; Southern Manchuric: Jurchen *wúh-yè-hēi*; Manchu *ujhe ~ wejhe* (SSTMJa 1: 298b-299a).

The Ket word belongs to the hybrid words. In ethnographic works the word is mentioned as a euphemism replacing the Yeniseian taboo word *ɯ̀r*. The word likely

consists of the Tungusic word *ije* ‘horn’ with the Yeniseian plural suffix and the Yeniseian word *ēˑs* ‘weather’. My assumption can be confirmed by the observation that reindeer shed their antlers during the spring season.

6. Yeniseian loanwords in Tungusic

During my research I discovered new etymologies for some Ewenki words, e.g.:

(24) Northern Tungusic: Ewenki *homōtī* ‘bear’ < **homō* *+tī* {Ewenki NN/Adj.} ← Proto-Yeniseian **qoːħmə* < *qom* ‘taboo’ *+əŋ* {Yeniseian NN/Adj.}:

cf. Southern Ket *qɔ̀m*; Central Ket, Northern Ket *qɔ̀ːmə*; Yugh *χɔːħm* ‘taboo, sinful, ritually forbidden’ (Vajda & Werner: in preparation).

The Ewenki word *homōtī* ‘bear’ is possibly connected with the Common Yeniseian word *qom* ‘taboo’. Morphologically the Ewenki word was used with the Tungusic suffix *+tī*, which forms nouns that designate possession or a connection with something (for more on its function, see Vasilevič 1958: 793).

(25) Northern Tungusic: Yerbogachen Ewenki *kivšim* ‘fine snow’ < **kibti* *+m* {Ewenki NN/Adj.} ← Proto-Yeniseian **kibedj* < *kiˀ* ‘new’ + **beˀdj* ‘snow’:

cf. Ket *kívɛt*, Yugh *kíbɛtj* ‘new-fallen snow’ (Vajda & Werner: in preparation).

The Ewenki word probably was borrowed from Ket, where it is *kívɛt* ‘new-fallen snow’, with the Ewenki denominal noun suffix *+m*, which forms adjectives (for more on its function, see Vasilevič 1958: 769).

7. Conclusion

It is important to note that for the Tungusic loanwords in Yeniseian the source of borrowings were the Ewenki dialects. There are several Mongolic and Turkic words borrowed from Ewenki. The independent group of loanwords consists of hybrid words. Most loanwords are changed according to Yeniseian phonetic features (involving amalgamation, syncope, aphaeresis, metathesis, excrescence, Yeniseian suffixes as plural, the collective suffix, and the nominalizer, etc.). Morphologically they are mostly nouns and adjectives, but we also have a few verbs, adverbs and particles. In semantics the majority of Tungusic loanwords belong to the tundra, reindeer, hunting and house terminology.

References

Bold. = Boldyrev, B. V. 2000. *Ėvenkijsko-russkij slovar'*. 1-2. Novosibirsk: Izdatel'stvo SO RAN, filial GEO.

Buryat = Čeremisov, K. M. 1973. *Burjatsko-russkij slovar'*. Moskva: Sovetskaja Ėnciklopedija.

Doerfer, G. 1985. *Mongolo–Tungusica*. Wiesbaden: Harrassowitz.

Dolgikh, B. O. 1934. *Kety*. Moskva and Irkutsk: Poligrafkniga.

Georg, S. 2007. *A descriptive grammar of Ket (Yenisei-Ostyak). Introduction, phonology, morphology*. Folkestone/Kent: Global Oriental.

Kalmuck = Ramstedt, G. J. 1935. *Kalmückisches Wörterbuch*. Helsinki: Société Finno- Ougrienne.

Khabtagaeva, B. 2010. Mongolic elements in Barguzin Evenki. *Acta Orientalia Academiae Scientiarum Hungaricae* 63 (1). 9–25.

Khalkha = Bawden, Ch. 1997. *Mongolian–English dictionary*. London & New York: Kegan Paul International.

LM = Lessing, F. D. 1996. *Mongolian–English dictionary*. Bloomington: The Mongolia society, Inc.

Nugteren, H. 2011. *Mongolic phonology and the Qinghai-Gansu languages*. Utrecht: Landelijke Onderzoeschool Taalwetenschap.

Oyrat dial. = Coloo, J. 1988. *BNMAU dax' mongol xelnii nutgiin ayalguunii tol' bičig. Oird ayalguu.* Ulaanbaatar.

Poppe, N. 1964. *Grammar of written Mongolian.* Wiesbaden: Harrassowitz.

Rozycki, W. 1994. *Mongol elements in Manchu.* Bloomington.

SSTMJa = Cincius, V. I. (ed.) 1975-1977. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov*. 1-2. Leningrad.

Vajda, E. J. and Werner, H. (in preparation): *Etymological dictionary of the Yeniseian Languages.*

Vajda, E. J. (in press): Patterns of innovation and retention in templatic polysynthesis. In: Fortescue, M., Mithun, M. and Evans, N. (eds.) *Handbook of Polysynthesis*. Oxford: Oxford University Press.

Vajda, E. J. 2000. *Yeniseian peoples and languages. A history of Yeniseian studies with annotated bibliography and a source guide.* Richmond: Curzon Press.

Vajda, E. J. and Zinn, M. 2004. *Morphological dictionary of the Ket verb*. Tomsk: Tomskij gosudarstvennyj pedagogičeskij universitet.

Vajda, E. J. 2009. Loanwords in Ket, a Yeniseian language of Siberia. In: Haspelmath, M. and Tadmor, U. (eds.) *Loanwords in the world's languages: A comparative handbook.* Berlin: Mouton de Gruyter. 471–494.

Vajda, E. J. 2013. Metathesis and reanalysis in Ket. *Tomsk Journal of Linguistics and Anthropology* 1/1. 14–26.

Vajda, E. J. 2015. Yeniseian. In: Stekauer, P. and Lieber, R. (eds.) *Handbook of derivation*. Oxford: Oxford University Press. 509–519.

Vasilevič, G. M. 1958. *Ėvenkijsko-russkij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.

Werner, H. 2002. *Vergleichendes Wörterbuch der Jenissej-Sprachen*. 1–3. Wiesbaden: Harrassowitz.

Werner, H. 2005. *Die Jenissej-Sprachen des 18. Jahrhunderts*. Wiesbaden: Harrassowitz.

Yakut = Slepcov, P. A. 1972. *Jakutsko-russkij slovar'*. Moskva: Sovetskaja ėnciklopedija.

Lexical review of disease names in the Udmurt language

Rebeka Kubitsch
University of Szeged

1. Introduction

This paper discusses the lexical review of the disease names in the Udmurt language. The paper introduces what kind of crosslinguistic tendencies occur in the naming of diseases, which factors can motivate the naming, what morphological features are typical for the disease names, and how all these features appear in the Udmurt language. The contacts of the names with other lexical areas are also mentioned.

Udmurt is a Uralic language, part of the Permic subgroup. It is spoken by ethnic Udmurts in the Russian constituent republic of Udmurtia, were it is co-official with Russian. According to Ethnologue (2010), there are 324,000 native speakers and the population of ethnic Udmurts is 554,000.

There is very little research on disease names and there is not much general literature about them. Furthermore, these sources introduce just a smaller part of the topic (and the authors are usually doctors, not linguists: for example, Berde 1940, who developed the folk approximation of the disease names). Galgóczi (1981) deals only with those names which have turned into curses or swear words. Jääsalmi-Krüger (1990) introduced the disease names of the Khanty language. Kicsi (1999) wrote about disease names that involve personal names or refer to causes, and, finally, Magyar (2000) discussed those disease names which have come from animals' names. In any case, there is no previous study on this topic with regard to the Udmurt language.

2. Definition

For the examination a definition was needed of what we can regard as a disease name. During my work I have treated as a disease name every expression which

describes a pathological condition. A condition is pathological if an organ does not operate properly (Elekes 2010: 1). According to this definition it was problematic whether some symptoms (fever, cough), "aches" (headache, toothache) and cutaneous diseases (pimple, rash) could be considered as diseases. In the examination I treated them as such,[1] because although a present speaker of the language would not consider a word like wound as a disease name, however, in this case some kind of aberration can be identified compared to the healthy condition.

Furthermore, it is quite a circumstantial (even impossible) exercise to define where the line is since an expression is not a symptom but a disease name. During this examination I did not make this separation because, in my opinion, the answer to this question depends on the individual user and there is no general solution. I do not consider important the distinction between symptoms and diseases, because I am examining folk terminology, not the technical, medical one. According to Frake (1980: 72), folk medicine always searches for the individual reason and event of the disease, hoping that by putting an end to the cause the problem will go away. Based on this, at the perception of a wound or pain the main question could be "What has caused this?" That is, certain symptoms have not been considered as traces of a disease, but, as I mentioned before, have been searched for their individual reasons.

3. Corpus

In the study 56 names of diseases are examined altogether. Some of these are from the book Votják népköltészeti hagyományok [Udmurt folk poetry traditions] by Bernát Munkácsi (1887), while others are names elicited from a native speaker. The corpus includes the names of general diseases (e.g. flu, diarrhea), many kinds of skin conditions (e.g. furuncle, rash), and illnesses causing epidemics (e.g. smallpox, dysentery) are included. Further names of diseases were included on the basis of naming patterns in Hungarian and other languages. For the analysis I used the dictionaries of Bernát Munkácsi (1896) and István Kozmács (2002) and the assistance of a native speaker (Ekaterina Suncova, Udmurt instructor at the Finno-Ugric Department, University of Szeged)

[1] However, I treated them such that I did not collect all of the variants of "aches" in the Udmurt language. On the one hand because I will fully introduce their construction later, and on the other hand, by analogy we can "create" several disease names, but the aim of the study is not to collect more expressions but to introduce the naming patterns.

During the lexicological research I did not take into consideration which dialect the word belongs to.

4. Lexicological research

During the lexicological analysis I have examined the following features: (1) whether the basis of the naming is a metaphor or a metonymy; (2) whether the name focuses on the cause or the symptom of a disease; (3) whether the disease name is expressed with a stem, a derived noun or compound word; (4) whether there is evidence of contact with other lexical fields. In this study I do not discuss all 56 disease names, only examples of all the types.

The disease names in the folk medicine are predominantly based on substitution. This is connected mainly to taboos, because diseases were considered as deviant, sickening phenomena in most of the cases, and speaking about them was considered inappropriate (even nowadays we can observe this attitude among the speakers regardless which language they speak). This is the reason for the metaphorical and metonymical nature of disease names. Furthermore, it is a universal tendency in the languages of the world that an abstract phenomenon is expressed with the help of the lexemes of more concrete conceptual fields (Kövecses and Benczes 2010: 91). In addition to folk medicine, beliefs are important in the naming of diseases as well.

4.1. The metaphorical transfer

The use of a metaphor generates a substitution between two different objects or phenomena according to analogy and/or similarity (Szathmári 2008: 309). During the metaphorical transfer the identifier (a concrete entity) and the identified (an abstract entity) originate from two different, "distant" points of our conceptual system (Kövecses and Benczes 2010: 93). In addition, the identifier has an earlier, more evidential meaning, which can be connected to basic physical experiences (because of this it could be even an identifier) (Szikszainé 2007: 428).

Among disease names, the metaphorical ones represent the oldest layer (Magyar 2000: 162). It is quite common that these names are linked to the names of animals, plants or everyday objects. To this can be connected the belief that creeping animals (snakes, frogs, beetles, spiders) sneak into the human body through orifices and cause many kinds of problems (Magyary and Kossa 1929: 285–287). Here are some Hungarian examples: the frog which disturbs the belly (there is a Hungarian saying, according to which if someone drinks too much water, *béka nő a hasában* 'a frog grows in his/her belly'), or the beetle which moves in the head and causes

disturbance (another Hungarian saying is that if someone behaves in a strange way, *bogara van* "he/she has a beetle"). Another reason for metaphorical transfer is the above mentioned universal tendency that people tried to explain unfamiliar events or processes with the known environment, for example they compared the outer or inner attributions of an animal to the attributions of a disease (Kuna 2010). For example, English *chickenpox* ("chicken + pox"), Finnish *vesirokko* 'chickenpox' ("water + pox") and Hungarian *bárányhimlő* 'chickenpox' ("lamb + pox"). The analogy is the same in all three names: the words *chicken*, *vesi* and *bárány* refer to the fact that this kind of pox is less dangerous than the others (e.g. smallpox), since chickens, water and lambs are harmless, so the attribute of these things are similar to the attribute of the disease.

In the Udmurt corpus ten items were based on a metaphor. These are the following: *гижло* ("nail" + nmlz[2]) 'eczema, ringworm, scabies'; *горд пужы* ("red" + "freckle embroidery") 'rash'; *гыжло* 'sty'; *дыж* 'cataract'; *курег син* ("hen" + "eye") 'callus, glaucoma'; *потос* ("come out" + nmlz) 'carbuncle'; *пужы* 'measles'; *пуэд* ("tree" + nmlz) 'eczema, ringworm, scabies';flu'; *рак* 'cancer'; *сьӧдкыль* ("black" + "infectious disease") 'pox, typhoid fever'.

Of these expressions, I will discuss the following disease names in detail: *гижло* 'eczema, ringworm, scabies', *горд пужы* 'rash', *дыж* 'cataract', *курег син* 'callus, glaucoma'.

The basis of the name *гижло* 'eczema, ringworm, scabies' is the word *гижы* 'claw, nail' (Kozmács 2002: 90). The element *-ло* is a nominalizer, which is presumably used in dialects[3] (personal communication, Ekaterina Suncova, 2015). On eczematous skin pimples with a dry, scaly and hard surface can occur, and in some cases the skin can cornify by degrees. The basis of the metaphor is that the touch of these pimples is similar to the hardness of nails. As we can see, the word has many meanings: in fact this word can be used for every type of skin disease when the pimple surfaces are hard.

In the name *горд пужы* ("red + freckle, embroidery") 'rash' the color name refers to the color of the rash. As we can see, the word *пужы* has several meanings: 'freckles, embroidery, pattern, measles' (Kozmács 2002: 353). The basis of the metaphor is the similarity between the arrangement of the rashes and freckles. In

[3] I base my presumption on the fact that Csúcs (1990) mentions the most important nominalizers, which are basically typical of the literary language, and among these it does not occur.

addition, the pimply or freckled skin can look like embroidery or pattern. This is supported by the ‘measles’ meaning of the word.

The primary meaning of *дыж* ‘cataract’ is ‘shell, thin skin, membrane’. The origin of the naming is the same as in the Hungarian disease name (cf. Hungarian *hályog* in the sense of ‘membrane, bladder’), namely, the idea that a membrane covers the eye and this causes the visual impairment. The basis if this idea is the fact that during this disease eyes look as if a thin skin covered them.

In the case of *курег син* (“hen” + “eye”) ‘callus, glaucoma’ the metaphor originates from the formal similarity between the skin disease and the organ of the animal. In the middle of the typically round calluses a protrusion can be seen, so it is reminiscent of an eye. The same basis for the name of the condition occurs in Hungarian and other languages too (German *Hühnerauge*; French *oeil de poule*; Italian *occhio pollino*). The reason for this is that the expression is the calque of the Medieval Latin *oculus pullinus* (*oculus* ‘eye’; *pullinus* ‘hen’) (Magyar 2000: 175). The other meaning of *курег син* is ‘glaucoma’. The base of this naming is that one symptom of the glaucoma is that the lens of the eye becomes progressively opaque, resulting in blurred vision. When this happens the eye looks like the eyelid of a chicken.

4.2. Metonymical transfer

In metonymical transfer the connection between two phenomena is always a logical correlation. In this case instead of the conventional phrase we use a term which originally refered to the denomination of another thing. (Szathmári 2008: 410), Furthermore, in metonymical transfer the identifier (the concrete entity) and the identified (the abstract entity) originate from the same area of our conceptual system (Kövecses and Benczes 2010: 76).

In most cases the contiguity between two phenomena is a spatial or causal relation. A good example for the spatial connection is when a nation’s typical disease (e.g. which occurs typically in a certain country or among the members of a certain nation) is named after the name of the nation (Kuna 2010). The Finnish *espanjantauti* (“Spanish” + “disease”) ‘Spanish flu’ represents this type. The Spanish flu is the name of a special sort of flu which has caused the most destructive worldwide epidemic. The basis of the name is that the first news regarding the disease in 1918-19 came from Spain. A causal connection occurs if the disease is named after one of its causes or consequences, for example, Hungarian *teniszkönyök* (“tennis” + “elbow”) ‘tennis elbow’. The special inflammation in the elbow is

typical of tennis players. Because of this it is named after the sport which causes the condition.

In the Udmurt corpus 19 items were based on metonymy. These are the following: *берган кыль* (“rotation” + “infectious disease”) ‘mad cow disease’; *булык* ‘a kind of disease connected to eating too much; disease causing spirit’; *гордак* (“red” + nmlz) ‘measles’; *дэй* ‘serious disease, hernia, disease causing spirit’; *калера* ‘infectious disease, plague, cholera’; *кöт кыль* (“belly, stomach” + “infectious disease”), ‘diarrhea, dysentery’; *кыль* ‘infecious disease, evil spirit’; *кыльдэй* (“infectious disease” + “infectious disease”) ‘infectious disease, disease causing spirit’; *кынмон* (“cold” + nmlz) ‘cold’; *кырыж кук* (“crooked” + “leg”) ‘rickets’; *мыж* ‘gland disease, disease causing spirit’; *ныр вия* (“nose” + “flow”) ‘cold’; *пöськы* (“hot” + derivational suffix) ‘pimple, rash’; *синтэм* (“eye” + car) ‘cataract, blind’; *сьöдун* (“black” + nmlz) ‘rickets’; *сюлэм чер* (“chest + disease”) ‘chest disease'; *чер* ‘serious disease, infection, bad, disease causing spirit’; *ӧужектон* (“sallow” + nmlz) ‘jaundice’; *ымпотос* (“mouth + carbuncle”) ‘angina, herpes’.

Of these expressions, I will discuss the following disease names in detail: *булык* ‘a kind of disease connected to eating too much, disease causing spirit’; *калера* ‘infectious disease, plague, cholera’; *кыль* ‘infecious disease, evil spirit’; *кыльдэй* ‘infectious disease, disease causing spirit’; *пöськы* ‘pimple, rash’; *синтэм* ‘cataract, blind’.

The name *булык* ‘a kind of disease connects to eating too much, disease causing spirit’ expresses the disease and the spirit which causes the disease (personal communication, Ekaterina Suncova, 2015). In certain dictionaries the meaning ‘pox, rash’ is given, for example it can be found in compound words: *пушнер булык* (“nettle” + “pox”) ‘hives’. The basis of the metonymy is the causal connection, which in the disease is named after the putative causer. It is presumable that one kind of meaning adhesion has happened, because there is also a separate expression for the causer of this type of disease: *булык пери* ‘disease causing, evil spirit’ (Kozmács 2002: 48).

The word *калера* ‘infectious disease, plague, cholera’ is obviously the borrowing of the Russian *холера* ‘cholera’, but it has other connotations in the Udmurt language. In this case we can speak about a metonymical extension, because besides the ‘cholera’ meaning it is used in the more general sense of ‘infectious disease’. The basis of the metonymy is that the name of a concrete disease is widened to refer to the a type of disease.

The disease name *кыль* ‘infectious disease, evil spirit’ is also used in a general sense. This happens to almost every kind of infectious or serious disease (see above, *берган кыль* (“rotation” + “infectious disease”) ‘mad cow disease’; *кöт кыль* (“belly, stomach” + “infectious disease”) ‘diarrhea, dysentery’). The basis of substitution is that the disease name is referred to with the name of the cause. In this case the importance of beliefs in folk medicine can be seen and in the second paragraph mentioned “searching for reason” attitude shows up too. The compound word *кыльдэй* ‘infectious disease, disease-causing spirit’ is connected to this lexeme. The first part of the construction is the above mentioned *кыль*, while the other is *дэй* ‘serious disease, hernia, disease causing spirit’. As can be seen, both words have similar meanings (both meanings are general and express the causer spirit), and the compound’s meaning is the same like its elements. According to Munkácsi’s dictionary (1896: 164), *кыльдэй* is a spirit, which demands sacrifices from people at abandoned places. The basis of the metonymy is the same as in the previous case – the disease name is substituted with name of the spirit. All names introduced in this paragraph, derived from their ‘evil spirit’ meaning, are quite archaic.

The stem of the word *пöськы* ‘pimple, rash’ is *пöсь* ‘hot’. The element *-кы* is presumably a derivational suffix (personal communication, Suncova 2015).[4] The basis of the metonymy is that the disease is named after a symptom, because the inflamed skin is much warmer than its environment.

In this chapter the last introduced disease name is *синтэм* ‘cataract, blind’. The meaning of *син* is ‘eye’ and *-тэм* is a very productive caritive suffix (Csúcs 1990: 59), so it literally means “the one without eyes”. The basis of the metonymy is the experience that if someone has cataract (s)he can’t see (or only very poorly) – the name of the disease is substituted with the name of the body part in question. The same is valid for the word’s other meaning, ‘blind’. Figure 1 below summarises the results.

[4] Regarding to the exact meaning of *-кы* there is no further information.

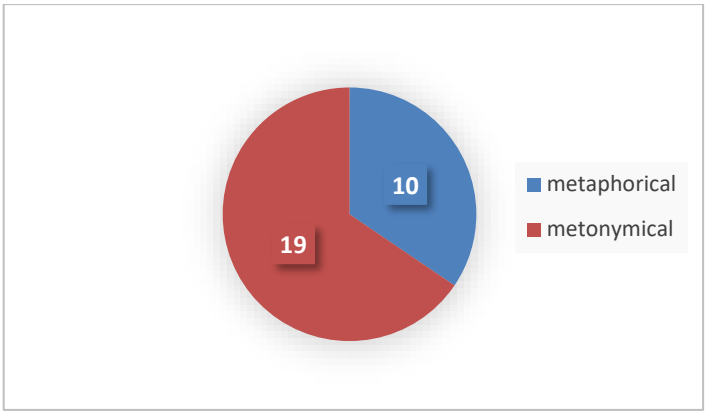


Figure 1
The distribution of metaphorical and metonymical disease names

4.3. Cause- and symptom-focused disease names

The essence of the cause- and symptom-focusing attitude is that during the naming process the disease is identified according to the actual or putative cause, or according to one of its typical symptom.[5] Berde (1940: 95–6) has noticed this phenomenon during his research into folk dermatology. Although Berde has made these statements about folk dermatology, they are also valid for everyday approximation of disease names (Kicsi 1999: 108). For example, the English cold 'cold, sniffle' is a cause-focused disease name, because the problem is named after its cause. But Hungarian *sárgaság* ("yellow" + "nmlz") 'jaundice' disease name focuses on the symptom, because jaundice is a typical symptom of hepatitis.[6]

Depending on which type of disease name predominates in a given language it can, then, be labeled as cause- or symptom-focusing one. For example, in the 1940's

[5] It is important to mention that although there is a causal connection in the cause- or symptom focusing attitude, a cause- or symptom-focused disease name is not automatically a metonymical one. The type of the attitude is also identifiable among the metaphorical disease names.

[6] The cause- or symptom-focused approach is not identifiable in the case of every disease name. This can have many reasons: the word form cannot be segmented, there is not enough information about the etymology, and in the localizing compounds (4.4) none of the approximations prevail, because the localization helps in the identification of the disease, and is not the reference to its cause or symptom.

Berde identified the Hungarian disease approach as a cause-focusing one (Berde 1940: 84), but Kicsi in 1999 claimed it to be symptom-focusing (Kicsi 1999: 190). The phenomenon which was mentioned in **4.2**, namely, when the disease is named after the causer spirit, also belongs to the cause-focusing aspect. This mentality is one of the reasons for the archaic nature of cause-focused disease names.

Below, I will discuss the disease names analysed in chapters **4.1** and **4.2** according to the cause- and symptom focused aspect.

From the metaphorical group, *гижло* ‘eczema, ringworm, scabies’, *горд пужы* ‘rash’ and *курег син* ‘callus, glaucoma’ are symptom-focused. The explanation is obvious: in the case of name *гижло* ‘eczema, ringworm, scabies’ a hard surfaced peeling is a typical symptom of this skin disease; in the case of name *горд пужы* ‘rash’ the name shows the colour and arrangement of pimples which are obviously the symptoms of the disease; and in the case of name *курег син* ‘callus, glaucoma’ as has been mentioned before, the distinctive mutation on the feet is the symptom of the problem. Nevertheless, *дыж* ‘cataract’ is a cause-focused disease name, because the name shows the putative source of visual impairment, that is, a membrane covering the eye.

From the metonymical group, *булык* ‘a kind of disease connected to eating too much, disease causing spirit’, *кыль* ‘infectious disease, evil spirit’, and *кыльдэй* ‘infectious disease, disease causing spirit’ are cause-focused names, because, as has been mentioned before, they are named after the causing spirit.

From this group *пӧськы* ‘pimple, rash’ and *синтэм* ‘cataract, blind’ are symptom-focused disease names. In the case of name *пӧськы* ‘pimple, rash’ it is obvious that the warm skin surface is the symptom of the inflammation, and in the case of name *синтэм* ‘cataract, blind’ the loss of vision is the consequence of the disease.

In the Udmurt corpus altogether nine items were cause-focused and 19 items were symptom-focused (Figure 2). Of the cause-focused names only one was metaphorical, eight were metonymical, while of the group of symptom-focused names five were metaphorical, and 14 were metonymical.

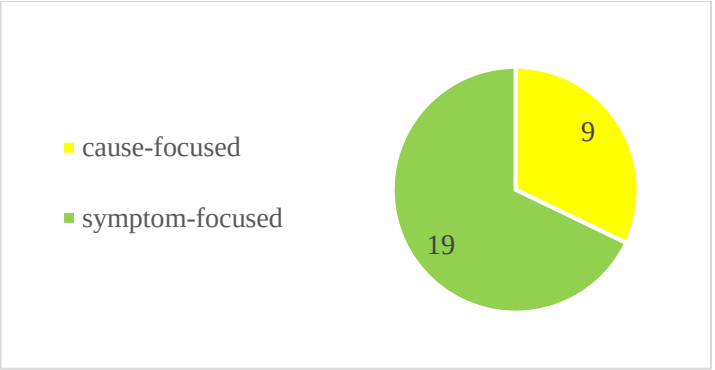


Figure 2. The distribution of cause-and symptom-focused disease name

4.4. Morphological features

The items of the corpus were classified according to the morphological structure of the lexemes: stems, derived words and compound words.

In the Udmurt corpus 13 stems occurred. Most frequently their meanings are 'disease' or 'wound', for example: *мыж* 'gland disease, disease-causing spirit'; *шек* 'serious wound'; *чер* 'serious disease, infection, bad, disease causing spirit'; *яра* 'eczema, ulcer, deep wound'. In several cases either the motivation of the naming process or the prevailing approximation cannot be identified. However, these stems quite often occur in compounds (*сюлэм чер* 'chest disease', *сюлпуш яра* 'stomach ulcer, duodenal ulcer').

In the corpus 13 derived words were found. In section **4.1** and **4.2** I have mentioned most of the derivational suffixes. In addition, the nouns derived from verbs are the most common, whereas disease names derived from nouns or adjectives are less common.

The most productive derivational suffix is *-он*/*-ён* deverbal nominaliser, which derives nouns from abstract verbs (Csúcs 1990: 59), for example, *кынмон* 'cold' from the verb *кынмыны* 'chill', or *ӧужектон* 'jaundice' from the verb *ӧужектыны* 'sallow'.

The *-ос*/*-ёс* derivational suffix, which occurs in *потос* 'carbuncle', is less frequent. In addition the meanings of two more derivational suffixes could be identified exactly: the caritive suffix *-тэк* (Csúcs 1990: 59) and *-ак*/*-як*, which expresses a reduced owning of an attribution (personal communication, Ekaterina Suncova, 2015). In the other cases, *-эд*/*-ед* (*пуэд* 'flu' – "tree" + nmlz), *-ло* (*гижло* 'eczema, ringworm, scabies' – **4.1**), *-кы* (*пӧськы* 'pimple, rash' – **4.2**) and *-ун*

(*сьӧдун* ‘rickets’ – “black” + nmlz) the meanings of the derivational suffixes are not sure. In many cases these disease names were the only examples for the use of these derivational suffixes.

In the corpus 19 disease names are expressed with compound words. It is typical that a word joins to a ‘pain, bad, wound, disease’ meaning word (Jääsalmi-Krüger 1990: 9). Within this, a subgroup can be identified, that of localizing compound words. In this group the disease name somehow localises the source of the problem. The composition ‘body part’ + ‘disease, ache’ is the most frequent in general (e.g. headache, brain tumour) (Kuna 2010).

Udmurt *сюлдэй* ‘gripes, diarrhea’ is a good example of localizing compound words. The first part of the compound is *сюл* ‘gut’, while the second is the above mentioned *дэй* ‘serious disease, hernia, disease-causing spirit’ (**4.2**). The disease name identifies the location of the problem, because this type of disease concerns the gut. In the Udmurt corpus 9 of the 19 compound words were localizing.

It is not uncommon that the name of an animal, plant or color is compounded with the word meaning ‘disease, ache’ (for an explanation, see chapter **4.1**), for example, chickenpox and red measles.

In the corpus there were 13 stems, 13 derived words and 19 compound words altogether (Figure 3).

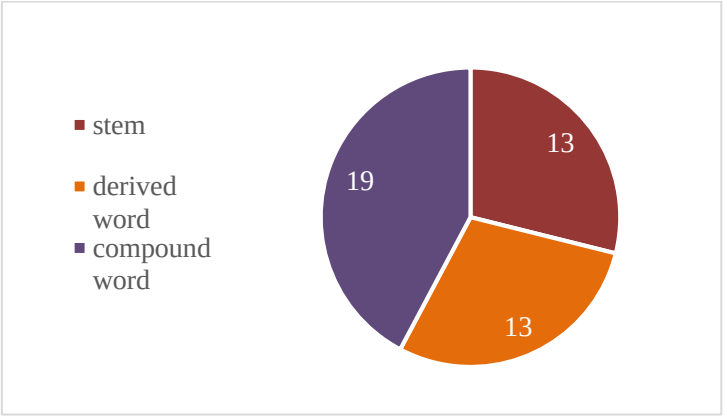


Figure 3. Morphological features

4.5. Unidentified disease names

From the Udmurt corpus a total of 12 items were not examined. The reason is that there was not enough information about either the motivation of naming or the

morphological characteristics. Among the 12 disease names seven items are of Russian origin. These words have been left in the corpus because in spite of their foreign origin, they are used in the Udmurt language. However, I did not examine them because they do not have any further meaning in Udmurt. The disease name *калера* ‘infectious disease, plague, cholera’ is an exception, because it has further meaning (‘infectious disease’) in addition to the original one (**4.2**).

4.6. Contacts with other lexical fields

There’s a contiguity between the lexical field of disease names and the group of animal, plant and color names. Examples are mentioned in chapter **4.1**, and the reasons of connection are explained as well. In addition, this lexical field often contacts with the field of body parts, mainly in compound words, because of the claim of localization (**4.4**). The group of several kinds of gods, spirits names is frequently appears among the contacts, too. This connects to the belief that diseases are sent by gods or spirits (**4.2**, **3**). The lexical fields mentioned in this paragraph are the most typical of the naming of disease names.

Nevertheless, we have to take into consideration those lexical fields too which name their elements after disease names. During the research two such groups were identified. The first is the group of plants names, the other is the group of metaphorical, idiomatic expressions. Into the group of plants names usually the names of herbs are involved. It is a frequent phenomenon, not only in Udmurt that a herb is named after the disease it cures. For example, the Udmurt *кöткыльсяська* (“diarrhea, dysentery” + “flower”) ‘silverweed’, which is used for diarrhea, intestinal bleeding, and other kinds of stomach diseases. Or *булыктурын* (“a kind of disease connected to eating too much, disease-causing spirit” + “herb”) ‘salsify’, which can ease stomach ache, so it could (or still can) be used to cure this kind of problem.

The usage of disease names in metaphorical, idiomatic expressions is the topic of another paper (Kubitsch 2015).

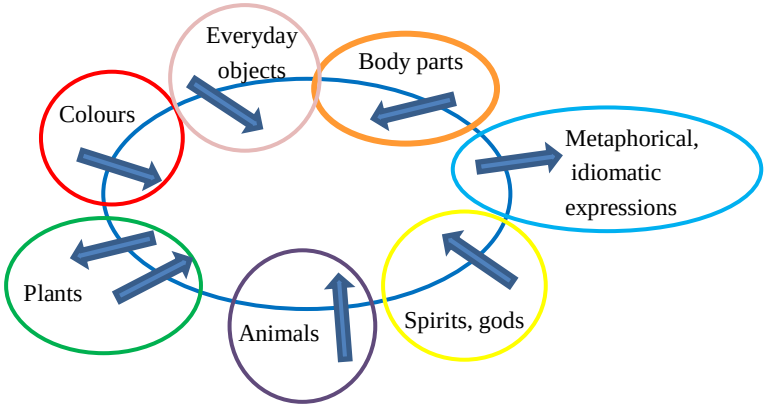


Figure 4. Contacts with other lexical areas

5. Conclusion

In the lexicological research reported on in this paper altogether 56 disease names were examined. Although the research cannot be considered exhaustive, some – in my opinion general – features of the naming methods and the morphological structures can be identified, for example, the predominance of metonymic disease names compared to the metaphorical ones. Of the 56 disease names, in 29 cases the basis of the naming could be identified. Of these 29 items 19 were metonymic and ten disease names were metaphorical (Figure 5). In my opinion the predominance of metonymic disease names is a general phenomenon in the Udmurt language and similar results would be found in a bigger corpus as well. I think that the reason for this is the development of medicine – during the identification of a disease its similarity to a well-known thing is less important, we rather identify it according to causal connections. As Magyar (2000) remarked, metaphorical names represent the oldest layer of disease names.

Another general phenomenon is that the symptom-focused names are more common than the cause-focused ones. This is true the metaphorical and metonymic disease names as well (**4.3**). In the corpus in 28 cases was the disease approach identifiable. Of these, in nine cases the approximation is cause-focusing and in 19 cases it is symptom-focusing (Figure 6). I have mentioned before that the cause-focused names are usually archaic (**4.3**) and the cause seeking attitude is typical of folk medicine (chapter **2**). Because of the development of medicine this need fades away and preferably we focus on the symptoms during the identification of a disease. That is why I think that the predominance of symptom-focused names

would be the same in a bigger corpus too. Thus, it might be plausible to claim that the Udmurt disease approximation is symptom-focusing just like the Hungarian.

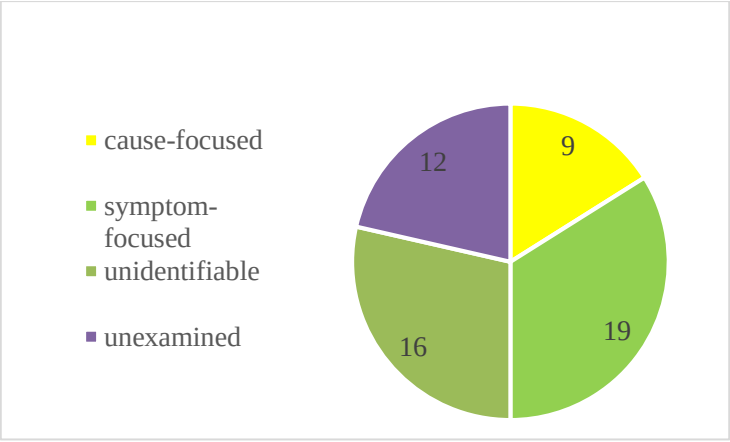


Figure 6. The cause- and symptom focused disease names

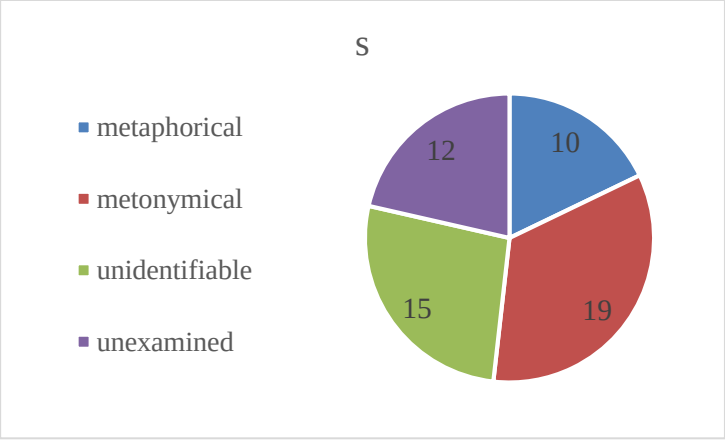


Figure 5. The metaphorical and metonymic disease names

Among the morphological features the dominance of compound words has been observed. The morphological structure could be identified in 45 cases. Of these, 13 were expressed with stems, 13 were derived words, and 19 compound words (Figure

7). Among the compound words nine items were localising compounds (Figure 8). We can see that the claim of localization is quite strong. In my opinion in a bigger corpus the proportion of localising compounds would be very similar. It is also important that in many cases neither the disease approximation nor metaphorical or metonymic based naming could be identified among the localising compounds. The reason could be that in these cases the localising helps in the identification, but it is not necessary that we identify it according to its similarity to something or according to its cause or symptom.

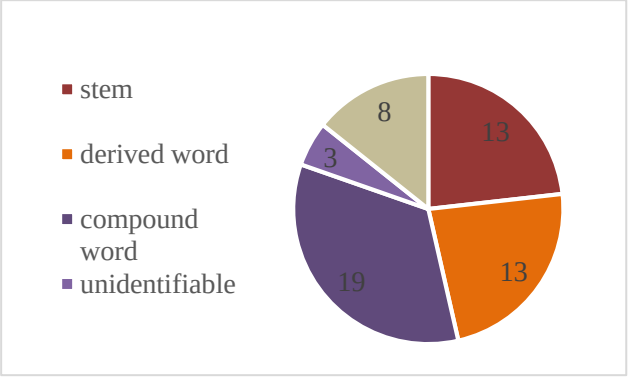


Figure 7. Morphological features

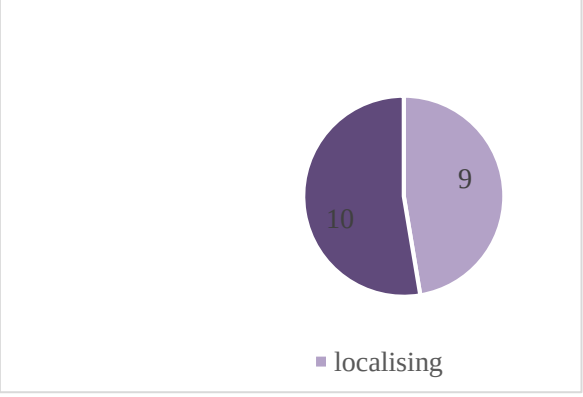


Figure 8. Localising and other compounds

Abbreviations

CAR	Caritive
NMZL	Nominalizer

References

Berde, Károly 1940. *A magyar nép dermatológiája* [The dermatology of the Hungarian people]. Budapest: Magyar Orvosi Könyvkiadó Társulat.

Csúcs, Sándor 1990. *Chrestomathia Votiacica*. Budapest: Tankönyvkiadó.

Elekes, Zsuzsanna 2010. *Egészségkárosító magatartások és mérési módszerek* [Health endangering behaviors and methods of measurement]. www.tankonyvtar.hu

Ethnologue 2010. http://www.ethnologue.com/language/udm

Frake, Charles 1980. *Language and cultural description*. Stanford: Stanford University Press.

Galgóczi, László 1981. Szitokszóvá vált betegségneveink [Disease names that became curse words]. *Magyar Nyelv* 77: 188–196.

Kozmács, István 2002. *Udmurt–magyar szótár* [Udmurt–Hungarian dictionary]. Szombathely: Savaria University Press.

Jääsalmi-Krüger, Paula 1990. *Ostjakische Krankheitsnamen und deren Benennungmotivation. Eine onomasiologisch-semasiologische Untersuchung.* [Khanty disease names and motivation of their naming. An onomasiologic-semasiologic examination.]. (Dissertation). Wiesbaden: Veröffentlichungen der Societas Uralo-Altaica 32.

Kicsi, Sándor András 1999. A hagyományos magyar népi okcentrikus betegségszemléletről [On the traditional Hungarian cause centric view of diseases]. *Orvostörténeti közlemények* 166–169: 107–119. www.orvostortenet.hu/tankonyvek/tk-05/pdf/4.25/1999_166_169_kicsi_sandor_hagyomanyos_magyar.pdf

Kubitsch, Rebeka 2015. The metaphorical usage of disease names in the Udmurt language, *National Students' Scientific Conference,* Budapest. ms.

Kövecses, Zoltán and Benczes, Réka 2010. *Kognitív nyelvészet* [Cognitive Linguistics]. Budapest: Akadémiai Kiadó.

Kuna, Ágnes 2010. *Szent Antal tüzétől a Münchausen-szindrómáig* [From St. Anthony's fire to the Munchausen's syndrome].

http://e-nyelvmagazin.hu/2010/12/06/szent-antal-tuzetol-a-munchausen-szindromaig-%E2%80%93-regi-es-a-modern-betegsegneveink-nehany-jellegzetessege/

Magyar, László András 2000. Állatok és betegségnevek [Animals and disease names]. *Orvostörténeti közlemények* 170–173: 161–177. www.orvostortenet.hu/tankonyvek/tk-05/pdf/2.13/2000_170_173_magyar_laszlo_allatok_betegsegnevek.pdf

Magyary-Kossa Gyula 1929. *Magyar orvosi emlékek* II [Hungarian medical records II] Budapest: Eggenberger-féle Könyvkereskedés.

Munkácsi Bernát 1990. *A votják nyelv szótára* [A dictionary of the Votyak language]. Az Uralisztikai Tanszék Kiadványai 1. Pécs.

Munkácsi Bernát 1887. *Votják népköltészeti hagyományok* [Votyak folk poetry traditions]. Budapest: Magyar Tudományos Akadémia.

Szathmári István 2008. *Alakzatlexikon* [A dictionary of figures of speech]. Budapest: Tinta Könyvkiadó.

Szikszainé Nagy Irma 2007. *Magyar stilisztika* [Hungarian stylistics]. Budapest: Osiris Kiadó.

The Nganasan lexicon from a diachronic onomasiological point of view: The case of metonymy[1]

Sándor Szeverényi
University of Szeged

1. Introduction

The aim of this paper is to provide a solution that has been developed to address an issue of categorization which has occurred during work on the diachronic cognitive onomasiological dictionary of Nganasan. The issue concerns the definition of the relationship between the meanings of two lexemes which have been categorized here as a type of conceptual contiguity (a metonymic relationship). One of the questions is whether this type of relationship indeed realizes a metonymic relationship (to oversimplify it: is the correct terminology used for it?), whereas the other question is along what principles and patterns this category can be differentiated further. Further differentiation is necessary because there are much more instances of conceptual contiguity than of other relationships of meaning, which makes the correct interpretation of metonymy and conceptual contiguity in the HeNg-On dictionary of utmost importance.

In this paper I present my own system, also discussing the theoretical underpinnings that are relevant to the argumentation. The topic is also relevant in view of the fact that several new works and analyses of lexical and word formation metonymy and of metonymy in general have been published in recent years, especially within the field of cognitive linguistics.

[1] Supported by the Hungarian National Research Fund, OTKA, grant number K100854 (2012–2015) and the Bolyai Scholarship of the Hungarian Academy of Sciences (2014–2017). I also used the Hamburg Nganasan Corpus that has been built Corpus based grammatical studies on Nganasan, supported by the DFG.

2. HeNg-On

HeNg-On is a diachronic cognitive onomasiological dictionary of the Nganasan language, a "historical etymological thesaurus" of sorts, with two major goals targeted in making it. The first has been to outline the lexical typological/cognitive onomasiological profile of Nganasan; defining the relationship of the meanings of various lexemes and working out the framework for these relationships is the most difficult part of this task. The second one has been to form groupings of Nganasan lexical items by origin (items created language internally, loanwords, lexical continuity, and unknown). Within the latter goal, uncovering the origin of lexemes so far categorized as unknown has also been set as an aim, especially as far as the northern Samoyedic languages are concerned.

All of this has been done on a dynamic website which makes it possible to continuously enlarge and refine the database and to carry out quick or complex searches, thereby aiding future lexicological, etymological and other linguistic research on these languages. The dynamic website can be found at www.hengon.arts.u-szeged.hu and can be used without registration. All relevant important information regarding the project (publications, links, user instructions etc.) has been made available here. Once the Lexicographical Program is created, the database can be extended in several ways, by increasing the number of languages, lexemes, concepts, semantic domains, analytical parts etc. The present paper relies on data that have been uploaded to the website by March 2015.

3. The Nganasan people and language

Nganasan belongs to the Northern branch of the Samoyedic group of Uralic languages. It is the northernmost language of Siberia and probably of the world. Officially classified as a moribund language, Nganasan is very close to extinction with slightly more than a hundred adult speakers. According to the 2010 Russian Census, the number of the ethnically Nganasan population was 839, with 125 speakers of the language.

The Nganasans live at the Taymyr peninsula in the Russian Federation. They live in a semi-nomadic way, mostly in two ethnically mixed settlements, Ust-Avam and Volochanka. The speakers of Nganasan are all bilingual in Russian, and in the past decades the process of language shift and language loss have accelerated to such an extent that at present they seem irreversible (for more on this, see Ziker 2002, and Wagner-Nagy and Szeverényi 2011). Nganasan has two main dialects, the Avam

and the Vadeyev dialects: the former is spoken in the western part of the Taymyr Peninsula, the latter in the eastern part. The differences between the two dialects are primarily in the phonology and lexicon – although, admittedly, studies of dialectal differences in Nganasan are rather limited. And because linguistic fieldwork has been done mostly in the western part of the language area, the Avam dialect is much better documented than the Vadeyev dialect is.

Nganasan is (still) regarded as an underdocumented language. The first relevant linguistic materials were collected by the Finnish scholar Matthias Alexander Castrén (1813–1852) in the 1840s. He produced a work which was not only a dictionary but an outline of a descriptive grammar, primarily of the morphology of Nganasan. Castrén's is a very valuable and precise collection of language materials which had no match in the next one hundred years. Soviet scholars published some materials on Nganasan after World War II, and some texts were published in Hungary by Mikola (1970) – but all of these were primarily texts in the language, mainly folkloric in nature. As far as the grammatical description of the language is concerned, Tereshchenko's 1979 grammar of Nganasan provided the next considerable step, followed by E. A. Helimski's work, which provided an increasing amount of carefully collected materials on the language beginning with the 1980s. But the main emphasis in this work was still on collecting texts, with language data collected via questionnaires lagging behind, although increasing in amount in absolute terms over the years. Because of this, descriptions of Nganasan grammar were for a long time based solely on collected texts. The work describing Nganasan received renewed impetus in the 1990s when the main morphophonological rules of the language were identified (e.g. Helimski 1994, Wagner-Nagy 2002). Two comprehensive chrestomathies (a collection of texts, grammar, and dictionary) have been published in Hungarian and German (Wagner-Nagy 2002 and Katzschmann 2008, respectively).

4. The Nganasan word formation

In order to have a clear view of the relationships between meaning and form in Nganasan, it is important to say a few words about word formation in this language.

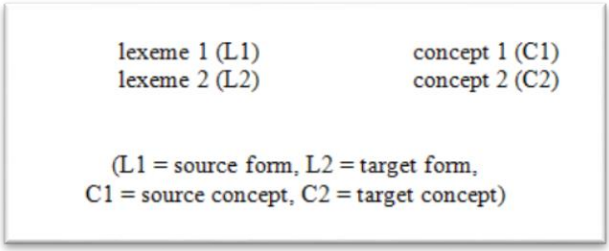
Nganasan is an agglutinative language, with inflections being suffixes exclusively. The most widely used method of Nganasan word formation is derivation: the number of both nominal and verbal derivational suffixes is relatively large (cf., for instance, Wagner-Nagy 2002). A less frequent but still important method is semantic derivation (Zalizniak 2008) or conversion, as it is traditionally

called, that is, zero derivation, or derivation without the use of overt formal morphological markers. Compounding is not typical in the language at all, only a handful of phrases are "suspected" to be compounds in Nganasan lexicography (e.g. *tuj* 'fire' + *ŋənduj* 'boat' > *tuu* [Gen] *ŋənduj* 'steam boat'). There are numerous loanwords in Nganasan, with the majority of recent lexical borrowings being, not surprisingly, of Russian origin. It has to be noted that almost the entire Nganasan speech community is characterized by bi- and multilingualism, and because of the rapid language shift that the community is undergoing it is often difficult to tell whether a given word is a Russian loanword or a codeswitch. For this very reason, words of Russian origin are currently not included in the dictionary. Russian loanwords in Nganasan have not been investigated in a comprehensive way yet, although several studies (by Futaky, Anikin, and Helimski) examined established loans. Other methods of word formation (e.g. reduplication or serial verbs etc.) are not used in Nganasan.

5. Diachronic cognitive onomasiology (DCO)

The theoretical framework of the dictionary is provided by DCO, as I have discussed in previous publications in detail (Szeverényi 2012, 2014). The most important points of this framework as are follows.

The dictionary classifies and systematizes relationships between lexemes, from the point of view of both formal relationships between two lexemes and the semantic relationship between their meanings:



All of this requires basic (etymological, lexicological etc.) research. It is a crucial question how conceptualization at the onomasiological level can be modeled.

The theoretical framework used here is provided primarily by Štekauer's onomasiological theory (1998, 2005). In his approach "the general linguistic background is that of the functional-structural approach of the Prague School of Linguistics. Therefore, the form-meaning unity, i.e., the bilateral nature of

morphemes is regarded as the fundamental principle", furthermore, word-formation occurs as an independent component in the following way. Štekauer presents a model where the word-finding process is divided into the following levels (see also Grzega 2008):

(1) the conceptual level, where the concept to be named is analyzed and conceptually categorized in the most general way – i.e. "SUBSTANCE, ACTION (with internal subdivision into ACTION PROPER, PROCESS, and STATE), QUALITY, and CONCOMITANT CIRCUMSTANCE (for example, that of Place, Time, Manner, etc.)";
(2) the semantic level, where the semantic markers or semantic components are structured;
(3) the onomasiological level, where one of the semantic components is selected as the onomasiological basis (representing a class like agent, object, instrument etc.) and another as the "onomasiological mark" of this basis (the mark can further be divided into a determining constituent — sometimes distinguishing between a specifying and a specified element — and a determined constituent) (= naming in a more abstract sense);
(4) the 'onomatological' level (with the Morpheme-to-Seme-Assignment Principle, where the concrete morphemes are selected (= naming in a more concrete sense); and
(5) the phonological level, where the forms are actually combined, respecting morphological and suprasegmental rules.

I concentrate on the first and fourth levels, that is, the conceptual and the onomatological. The second and third levels are difficult to use in the historical semantic and diachronic onomasiological framework, since our sources and linguistic competence do not make it possible to use them, allowing only for a speculative analysis. Furthermore, as has been pointed out by Grzega, differentiating between the first and second levels is problematic: "We may ask, however, on what cognitive or psycholinguistic results this model was constructed. The distinction between the conceptual and the semantic level is not corroborated by psycholinguistic analyses. These rather tells us that we should depart from what we could call a 'perceptual level', where both the more general, 'global' features and the more specific, 'local' features of a concept are processed at the same time" (Grzega 2008: 77).

This also foreshadows that in in the semantic analysis presented here only a general, abstract system can be made used.

5.1. Semantic innovations in DCO

There are different alternative ways of analyzing cognitive relationships between meanings and different depths to which such an analysis can extend. On the one hand, it is useful to employ terms used in lexicography internationally, while on the other it is important to take into account the characteristics of a given language or group of languages, with special attention to those languages that will be included in the database later. In the present project work by colleagues from Tübingen have been used – it is important to state that different lists of terms were used by them in several publications (e.g. Blank 2001, Gévaudan and Weibel 2004, Gévaudan 2007, Koch and Marzo 2007, Koch 2008). The basic system is as follows:

semantic relation	**semantic process**
identity	identity (verbum proprium)
taxonomic inclusion	specification
	generalization
contiguity	metonymy
similarity	metaphor

Table 1. Semantic relations and processes (Gévaudan 2007: 110)[2]

The same process/relationship can be categorized under different headings, for instance:

(1) PS **tuj* 'fire' Noun (SW 166) > Ng. *tusajkuə* 'black' Adj (KMZ 181)

1. lexical continuity: PS 'fire' > Ng. 'fire' (conceptual identity)
2. compounding: 'fire' + 'sand' > 'ashes, coal' (conceptualcontiguity: kind of)
3. suffixation: 'ashes, coal' > 'black' (conceptual similarity: color of)

At the same time, only the most notable characteristic is captured, for instance:

(2) *bi̇ni̇* 'rope, cord' Noun > *bi̇ni̇-d'i* (Infinitive) 'to domesticate (a reindeer), to teach a reindeer to wear a harness' Verb (KMZ26)

1. derivation: denominal verbal (conceptual contiguity: Object for Action)

[2] Some researchers treat taxonomic relations as part of contiguity due to the fact that they typically express metonymic (part/whole) relations. I follow Gévaudan's classification in this respect.

Doing the analysis without context has several drawbacks, one of which that is very relevant in this case is that it is difficult to analyze lexemes that are clearly connected but, without the context, the nature of their connection can only be established hypothetically.

6. Metonymy in HeNg-On

6.1. On the definition of metonymy

Cognitive linguistics treats metonymy as a conceptual process and metonymic relationships as conceptual relations (for more detail, see, for instance, Bencze 2009). Metonymy is one of the most innovative and most productive method, which, compared to the metaphor, had been seen as much less "interesting" by researchers for a long time but has become the focus of a number of monographs and volumes of studies lately (e.g. Denroche 2015, Littlemore 2015).

In diachronic cognitive onomasiology the category/term of contiguity/metonymy is used (e.g. Koch 2001, 2008, Blank 2001, Gévaudan 2007 etc.). The general definition of contiguity – as cited often and in various places – is a continuous mass, or a series of things in contact or in proximity. Contiguity metonymy is identified among the imaginative capacities of cognition (Langacker 1993). Metonymy is responsible for a great proportion of the cases of regular polysemy (Cruse 2000: 211).

From the perspective of metonymy, on the one hand, "[m]etonymy is a cognitive process in which one conceptual entity, the vehicle, provides mental access to another conceptual entity, the target, within the same idealized cognitive model" (Radden and Kövecses 1999: 21).

On the other hand, "the meaning relationships considered in the traditional study in linguistics of 'relational semantics', such as 'hyponymy', 'superordinacy', 'synonymy' and 'antonymy', are necessarily metonymic, because meaning relations described by them must involve some degree of semantic overlap" (Denroche 2015: 60). Furthermore, "the relationship between the superordinate vehicle and its hyponyms, e.g. car, bus, lorry, van, is metonymic; the relationship between the synonyms little/small, over/above, expert/specialist etc. is metonymic, because synonym pairs share denotational meaning, if not connotational meaning; and the relationship between 'complementary antonyms', such as on/off, open/closed, dead/alive, 'gradable antonyms', such as big/little, fat/thin, rich/poor and 'reversive

antonyms', such as start/stop, husband/wife, borrow/lend, are metonymic, as they also share complementary features."

Koch distinguishes three degrees of metonymic effects (Koch 2004):

I. non-literal ad hoc metonymies relying on implicatures at the universal level of (cognitive) speech rules;

II. non-literal discourse-ruled metonymies relying on conventional (or generalized) implicatures at the historical level, defined by discourse rules;

III. literal (lexicalized) metonymic polysemies relying on explicatures at the historical level, defined by language rules.

DCO focuses on the third type of effects, that is, it does not rely on contextual meaning but on historically fixed changes of meaning (which, of course, does not exclude it being morphologically motivated). Gévaudan (2007: 88–95, 1999) applies the phenomenon of conceptual contiguity/metonymy within the framework of DCO, thus building primarily on the tradition of historical semantics and rhetoric. According to him, all of the examples below exhibit metonymic relationships:

a. polysemy: Ger. *Glas* 'material' / 'drinking vessel'

b. change of meaning: Lat. *testimonium* 'testimony' > Fr. *témoin* 'witness'

c. suffixation: Esp. *toro* 'bull' > *torero* 'matador, bullfighter'

With the help of "Frame" categories (cf. Fillmore 1975), Gévaudan identifies metonymic relationships – this is the context in which the two meanings are connected. For instance, to refer to the last example: the connection between the meanings of toro and torero is contiguity which belongs under the frame BULLFIGHTING. As Denroche (2015: 60–61) remarks: "Fillmore's concept of the 'frame', closely equivalent to terms favoured by other scholars, such as schema, script, scenario and cognitive model, is a theory of understanding categories which relies on metonymic processing". Denroche quotes Fillmore, according to whom a frame is a collection of interrelated concepts: "I have in mind any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits"; and access to one of them allows access to the others: "when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available" (Fillmore 1982/2006: 373)."

6.2. Lexical and word-formation metonymy

So far, context free, primarily lexicographic material has been processed for the HeNg-On dictionary. From this it follows that it focuses mostly on the basic, primary meanings of lexemes and, due to a lack of a suitable corpus, it does not analyze special meanings, investigating lexicalized, literal metonymic relations. The reason for this is that, due to a lack of early sources, historical changes can only be reconstructed, and that the Nganasan linguistic data is not suitable for an analysis of linguistic creativity. However, the investigated relations include also those where the source is a reconstructed element belonging to an earlier historical layer, and, because of this, the relationship itself can only be hypothetical and reconstructed. There are 26 such relationships at present.

Defining the relationships between meanings is much more problematic than that of formal relationships. The largest group of problematic relations is that of conceptual contiguity (metonymy). Providing an exact definition is problematic, not only in terms of the present project, but also in the cognitive linguistic literature. In defining contiguity/metonymy, I have relied on Géavudan and Koch's system, also taking into account Janda (2011), Haselow (2011) and Štekauer's (2005) onomasyological theory, the common element of all of these being that, behind processes of word formation, they presuppose cognitive processes, some of them of the kind implied by derivational suffixes themselves and interpretable as Source+Target pairs of metonymic relationships. Most metonymic pairs were marked with a "metonymic pattern" label in the Comments field, which refers to basically conceptual categories. Two such examples are as follows:

(3) *səənə* 'foolish, stupid, silly' Adj > *səəna-m-sa* (Infinitive) 'to become foolish, to become stupid, to become silly' Verb

word-formation: denominal verbalizing derivational suffix (translative)
semantic relation: conceptual contiguity (Property for Result)

(4) *basa* 'iron, metal' > *basa* 'money'

word-formation: semantic change
semantic relaton: conceptual contiguity (Material for Object)

It is important to discuss the relationship between derivation and conceptual categories separately at the lexicological and morphological levels. We have relied on Haselow (2011) in this, who has investigated the interrelationship of suffixation

and conceptual categories from a historical aspect. His analysis is compositional, although it is clear that compositionality may be lost in processes of lexicalization. As Štekauer (2005: 212) also recognized, "word-formation deals with productive and rule-governed patterns (word-formation types and rules, and morphological types) used to generate motivated naming units in response to the specific naming needs of a particular speech community by making use of word-formation bases of bilateral naming units and affixes stored in the Lexical Component."

A heated debate has taken place recently about the issue of metonymic relationships and suffixation in the journal Cognitive Linguistics, centering on the categorization of word-formation metonymy and lexical metonymy: Janda (2011, 2014) argued that derivational affixes themselves can express metonymic relationships, as is exemplified in Table 2 (Janda 2014: 345):

no derivation	*milk* n.	lexeme	contained for container	as in *The milk tipped over*
zero derivation	*milk* v.	conversion	product for action	as in *The farmer will milk his cows*
overt derivation	*milker* n.	morphological derivation	action for agent	as in *She is good milker*

Table 2

Janda's stance can be summarized as follows: (1) the focus of most works on metonymy has been on lexical metonymy, how to describe it, and how to distinguish it from the metaphor. "Metonymy is an inferential relationship between two concepts: a source concept is overtly named and provides mental access to a target concept in a given context" (Janda 2011: 360). (2) According to her, there is no fixed boundary between lexical metonymy and word-formational metonymy since they coexist in the lexicon-grammar continuum. And finally, (3) context, whether it be a suffix or other cues, is always a factor in metonymy. The following, then, applies in word-formation (Janda 2011: 360):

the source: word that the derivation is based on
the context: the affix (for the metonymic relationship)
the target: the concept associated with the derived word

Analyzing linguistic data from Russian, Czech, and Norwegian, Janda claims that "actually there are more types of metonymy patterns in word-formation than in the lexical use of metonymy" (Janda 2011: 362).

Janda's theory has been criticized by Brdar and Brdar-Szabó (2013, 2014), whose point relevant to the present discussion is as follows: "While both the base and the suffixation is nominal, i.e. the metonymic vehicle is manifest as a noun and the putative metonymy is a noun, as the suffix is word-class maintaining, the verbal base (bake) can hardly be believed to provide simultaneous access to both the concept of 'baking' as activity and 'baker' as the participant in the activity" (Brdar and Brdar-Szabó 2013: 45).

6.3. Conceptual categories and suffixation

Reference to Janda (2011 and 2014) in the present discussion is made relevant by the fact that in Nganasan, where derivation is the most frequent process of word formation, several derivational suffixes exist that follow certain metonymic patterns – although I cannot and do not want to take a stand on whether this really presents a "context" for metonymy. In this, I side with Brdar and Brdar-Szabó, "Metonymic shifts do not arise in the course of derivation, but either operate on the end-result of word-formation" (Brdar and Brdar-Szabó 2013: 45), however, Janda's works have demonstrated that derivation contributes to the realization of a given metonymic relationship.

The term "metonymic pattern" is used here as the type of the relation between conceptual categories such as Person, Object, Action, Instrument etc. I argue that – following Janda (2011, 2014), and especially Haselow (2011) – the word formation processes determine cognitive processes as well. I have applied some basic conceptual categories to describe metonymic relations. These are general, abstract categories which are not the same as the notions of "schema" or "frame" (e.g. Fillmore 1975) but are more general:[3] Person, Object (Material, Instrument), Action (Motion, Event etc.), Characteristic/Property, Abstract (Manner, Result, Goal, Category, Possession etc.), Place.

The starting point is that the meanings of source and target forms can be categorized into conceptual categories (schemas), but the abstract categories of target forms can be consistently defined by certain productive suffixes. Suffixes

[3] Haselow applies five conceptual categories that are assumed to compose the schema of a particular situation: Person, Object, Location, Action (event), and Abstract (Result, Goal) (Haselow 2011: 56).

indicate particular conceptual categories, e.g. teach-er, surf-er ‘person who performs V’: Action > Person. We can clearly see that certain productive derivational suffixes behave consistently: forms derived with them form metonymic relations with their sources, and the result of the process also falls into a certain conceptual category, such as in the following cases:

-m- translative suffix (Noun > Verb)

sʲiər ‘cause, reason’ N > *sʲiərɨ-m-sɨ* (infinitive) ‘to be guilty’ V (KMZ154-155)
conceptual contiguity: Object for (Change of) Property

-ə relational adjectival suffix (Noun > Adjective)

ńersəgə ‘enemy, foe’ > *ńersəgə-ə* ‘hostile’ (KMZ111)
conceptual contiguity: Person for Property

inflectional prolative suffix *-mənu* (Adjective > Adverbial)

(5) *ərəkərə* ’beautiful’ Adj > *ərəkərəmənu* ‘beautifully, well’ Adv (KMZ219)

conceptual contiguity: Property for Manner pattern

Some derivational suffixes do not show such consistency. One reason is that some non-productive suffiexes are analysed as well.

(6) *labsə* ‘cradle’ > *labsə-kəə* ‘the youngest child in the family’

derivation: the derivational suffix *-kəə* is a non-productive adjective forming suffix
conceptual contiguity: Characteristic for Person

The dictionary contains the following main metonymic patterns:

Source	Goal
Action	Characteristic
	Object (e.g. Instrument)
Characteristic	Object
	Person
	Manner
	Material
	Person
	Result

	Category
Material	Action
	Object
	Property
Object	Action
	Property
	Category
	Motion
	Place
State	Action
	Possession

6.4. Metonymic relations in HeNg-On

In the analysis, relations between meanings and relations between forms are treated separately. The former define semantic relations, accompanied by the manner of the formal process.[4] Aiming to carry out an investigation of the entire basic vocabulary, this way it is possible to model what word-formation process typically accompanies what semantic process in Nganasan processes of lexicalization. This is in accord with the primary aims of DCO: we can get closer to creating the motivation profile of a language (cf. Koch 2001, Koch and Marzo 2007, Koch 2008).

At present there are 576 cases of conceptual contiguity where a Nganasan lexeme is the source. From the point of view of form, the proportions are as follows:

derivation:	481	(a total of 586)
lexical continuity:	26	(a total of 328)
conversion:	70	(a total of 125)
loan:	1	(a total of 24)

(A semantic relation is characterized by more than one morphological relation.)

The smaller proportion of the last three groups can be explained by the following:

[4] It sometimes (admittedly very rarely) occurs that it is difficult to define which one is the original form, and which one is the target. This can happen in cases of semantic change without change in form, or in cases suspected to be formed through re-analysis (these are usually relational adjectives).

Most elements inherited from earlier historical layers have preserved their original general meanings according to the reconstructions. There are few examples in the etymological literature where forms of a daughter language appear with different, derivated meaning:

(7) PS **kåptə̂-* ‘to castrate’ V (SW60) > Ng. *kobta-ʔa* ‘deer buck, castrated male deer’ N (KMZ66)

1. lexical continuity
(2. deverbal nomen (augmentative)
conceptual contiguity Action for Category (Property for Type)
conversion

(8) Ng. *ŋənduj* ‘boat’ N > *ŋəntəusa* ‘to ride a boat’ V

1. zero derivation
Instrument for Motion (Object for Action)

A considerable number of the analyzed loanwords did not undergo meaning change but were, instead, borrowed together with their original meanings (22 of the 24 examined forms).

7. Conclusion

Returning to the original question, namely, the investigated relations can be analysed as metonymy or they are something other. As we have seen, there is no general, unambiguously applicable notion of metonymy in onomasiology or in cognitive linguistics, and using the broad notion of metonymy as proposed by Janda might be the solution. Since the present analysis examines the result rather than the progression of the process, it is not of primary concern whether metonymic relations are expressed by general and frequent suffixation in Nganasan or, instead, the the meaning of the derivated form (stem + suffix) is crucial. At the same time, we can also see that some suffixes consistently trigger a change of conceptual categories.

Abbreviations

Adj	adjective
Adv	adverbial
N	noun
V	verb

References

Bencze M. Ildikó 2009. Mi a metonímia? Elméletek és kísérletek a metonímiakutatásban [What is metonymy? Theories and experiments in research into metonymy]. *Magyar Pszichológiai Szemle* 64/4: 677–696.

Blank, Andreas 2001. *Words and concepts in time: Towards diachronic cognitive onomasiology*, [www.metaphorik.de 01/2001]

Brdar, Mario and Brdar-Szabó, Rita 2013. Some reflections on metonymy and word-formation. ExELL (ExELL_1_040_Brdar_Brdar_Szabo.pdf) Date of access: 11/12/2014.

Brdar, Mario and Brdar-Szabó, Rita 2014. Where does metonymy begin? *Cognitive Linguistics* 25/2: 313–340.

Denroche, Charles 2015. *Metonymy and language: A new theory of linguistic processing.* New York: Routledge.

Fillmore, Charles 1975. *Lectures on deixis*. Indiana University Linguistics Club, Bloomington, IN.

Fillmore, Charles 2006 (1982). Frame semantics. In: Geeraerts, Dirk (ed.) *Cognitive linguistics: Basic readings*. Mouton de Gruyter, Berlin. 373–400.

Gévaudan, Paul and Wiebel, Dirk 2004. Dynamic lexicographic data modelling: A diachronic dictionary development report. LREC 2004. http://homepages.unituebingen.de/paul.gevaudan/lrec_gevaudanwiebel2004.pdf.

Gévaudan, Paul 1999. Semantische Relationen in nominalen und adjektivischen Kompositionen und Syntagmen. *PhiN* 9: 11–34.

Gévaudan, Paul 2007. *Typologie des lexikalischen Wandels*. Stauffenburg, Tübingen.

Grzega, Joachim 2008. Onomasiological approach to word-formation. http://www.skase.sk/Volumes/JTL03/08.pdf. Date of access: 15/06/2016.

Haselow, Alexander 2011. *Typological changes in the lexicon: Analytic tendencies in English noun formation*. Berlin: Walter de Gruyter.

Janda, Laura A. 2011. Metonymy in grammar: Word-formation. *Cognitive Linguistics* 22(2): 359–392.

Janda, Laura A. 2014. Metonymy and word-formation revisited. *Cognitive Linguistics* 25(2): 341–349.

KMZ = Kosterkina, N. T. – Momde, A. Č. – Ždanova, T. Ju. [Костеркина, Н. Т. – А. Ч. Момде – Т. Ю. Жданова] 2001. *Словарь нганасанско-русский и русско-нганасанский*. Санкт-Петербург: Просвещение.

Koch, Peter 2001. Lexical typology from a cognitive and linguistic point of view, in Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.): *Linguistic typology and language universals. Handbook of Linguistics and Communication Science* 20/2, Berlin: Mouton de Gruyter. 1142–1176.

Koch, Peter 2004. Metonymy between pragmatics, reference, and diachrony. http://www.metaphorik.de/sites/www.metaphorik.de/files/journal-pdf/07_2004_koch.pdf. Date of access: 15/06/2016.

Koch, Peter 2008. Cognitive onomasiology and lexical change: Around the eye. In: Vanhove, Martine (ed.) *From polysemy to semantic change: Towards a typology of lexical semantic associations*. Studies in Language Companion Series 106. Amsterdam and Philadelphia: John Benjamins. 107–137.

Littlemore, Jeannette 2015. *Metonymy: Hidden shortcuts in language, thought and communication.* Cambridge: Cambridge University Press.

Radden, Günter and Kövecses, Zoltán 1999. Towards a theory of metonymy. In: Panther, Klaus-Uwe and Radden, Günter (eds.). Metonymy in language and though. Amsterdam and Philadelphia: John Benjamins. 17–60.

Štekauer, Pavol 1998. *An onomasiological theory of English word-formation*. Amsterdam and Philadelphia: John Benjamins.

Štekauer, Pavol 2005. Onomasiological approach to word-formation. In: Pavol Štekauer and Rochelle Lieber (eds.), *Handbook of word-formation*. Dordrecht: Springer. 207–232.

SW = Janhunen, Juha 1976. *Samojedischer Wortshatz*. Helsinki.

Wagner-Nagy, Beáta and Szeverényi, Sándor 2011. Visiting the Nganasans in Ust'-Avam. In: Grünthal, Riho and Kovács, Magdolna. *Ethnic and Linguistic Context of Identity: Finno-Ugric Minorities*. (Uralica Helsingiensia 5). Helsinki: SUS. 385–404.

Zalizniak, Anna 2008. A catalogue of semantic shifts: Towards a typology of semantic derivation. In: Vanhove, Martine (ed.) *From polysemy to semantic change: Towards a typology of lexical semantic associations*. Studies in Language Companion Series 106. Amsterdam and Philadelphia: John Benjamins. 217–232.

Ziker, John 2002. *Peoples of the Tundra*. Long Grove: Waveland Press.

The first workday or the Moon's day? Germanic and Slavic traditions in naming the days of the week in the Finnic languages

Kasperi Hasala

University of Szeged

The aim of this paper is to describe the naming traditions of the days of the week in the Finnic languages. It is well known that the Finnish names of the days originate from the Scandinavian and Germanic languages. It is also obvious that the Estonian system concerning the naming of the days is different, having its origins in the Baltic and the Slavic cultures and languages.

In the present paper, I attempt to categorize the Finnic languages from the point of view of the names of the days. Furthermore, I seek to answer the question whether the Finnish names of the days date back earlier than it has usually been assumed.

This paper has a highly practical motivation. Actually, it is a summary of a chapter from a study of the European names of the days by Sándor Maticsák. When translating the book – *Vándorló napok* – into Finnish I considered that it would be more useful to write an entirely new chapter about the names of the days in the Finnish and Finnic languages instead of about the Hungarian ones.

As a first step, I am comparing the names of the days in all seven Finnic languages. Further, the etymology of the names of the days is compared with historical as well as historico-cultural information.

1. The Scandinavian *-tai* days

The names of the days of standard Finnish can easily be connected with the present and old Scandinavian system:

Finnish	Swedish	Old Swedish	Old Norse etc.
maanantai	*måndag*	*mānadagher*	OHG *mânatag*
tiistai	*tisdag*	*tīsdagher*	*týsdagr*
keskiviikko	*onsdag*	*ōþinsdagher*	*miðvikudagr*
torstai	*torsdag*	*þōrsdagher*	*þórsdagr*
perjantai	*fredag*	*frēadagher*	*frjádagr*
lauantai	*lördag*	*löghardagher*	*laugardagr*
sunnuntai	*söndag*	*sunnodagher*	OLG *sunnundag*

(OHG = Old High German, OLG = Old Low German)

The only exception is Old Swedish Wednesday *ōþinsdagher*, which does not fit the pattern. Finnish *keskiviikko* means 'the center of the week', while the Swedish equivalent has the meaning 'the day of the Odin (Wotan)'. However, the Old Norse (*miðvikudagr*) and the Old Middle German (*middeweke*) forms both had an equivalent naming for Wednesday with a similar meaning ('the center of the week').

These names of the days of Scandinavian-Germanic type predominate in Finnish only. Several dialect variants are even closer to the Old Swedish form having the suffix *-taki* (*maanantaki, tiistaki, torstaki* etc.).

The Old Swedish names had the following meanings:

Old Swedish	meaning
mānadagher	the Moon's day, as in Latin (*dies Lunae*)
tīsdagher	the Týr's day (West Germanic *Tiw*, *Tiu* > English *Tuesday*)
ōþinsdagher	the Odin's day (West Germanic Wotan > English *Wednesday*)
þōrsdagher	the Thor's day
frēadagher	the Freya's (Frigg's) day
löghardagher	the bathing day
sunnodagher	the Sun's day, as in Latin (*dies Solis*)

The idea of naming the days after gods and mythological figures comes from the ancient Romans and Greeks. Therefore, it is plausible to claim that the Germanic language(s) only transmitted the system to the other Northern peoples, the Finns and the Samis.

Ancient Finns understood the Germanic names as compound words, but their meanings were not very clear for them. This is why they were not translated – except the name of Wednesday, which was translated as 'the center of the week'. The other six names were only phonologically adapted (*mānadagher* > **mānantaki*, *þōrsdagher* > **torstaki* etc.). The shortest names were thought as simple stems. This is why the element *-taki* was weakened following to the rules of Finnish consonant gradation: **tīstaki* > **tīstaγi* > *tiistai*, **torstaki* > **torstaγi* > *torstai*.

2. The ordinal-based names

The naming principle of the days in the rest of the Finnic languages is mainly based on the ordinal numbers. The meaning of the names is literally 'the first day', 'the second day', etc.

	Karelian	**Izhor**	**Vepsian**	**Votic**
Mon	*enzimäinarki, ensiarki*	*enspäivä, ensimmässarki*	*ezmäžnarg*	*esimespäivä, tuhkapäivä*
Tue	*toiniarki*	*toispäivä, toisssarki*	*tožnarg*	*tõinõpäivä*
Wed	*kolmaspäivä*	*kolmaspäivä*	*koumanpäi*	*kõlma(i)späivä*
Thu	*ńeľľäšpäivä*	*neljääspäivä*	*ńeľľańźpäi*	*nellä(i)späivä*
Fri	*(piätinčä)*	*viijespäivä*	*(pätnič)*	*vijjespäivä*

This naming principle is possibly adopted from the Slavic or Baltic languages, since both language groups have similar systems:

Russian: *вторник* 'Tuesday', literally: 'the 2nd day', *четверг* 'Thursday', literally: 'the 4th day', *пятница* 'Friday', literally: 'the 5th day'.

Lithuanian: *pirmadienis* Mon, *antradienis* Tue, *trečiadienis* Wed, *ketvirtadienis* Thu, *penktadienis* Fri, *šeštadienis* Sat, *sekmadienis* Sun. Latvian has an identical model with one exception, namely, Sunday has a name with a religious motivation: *svētdiena* 'Holy day'.

It can be noticed that the ordinal based names in Karelian, Vepsian and Izhor have two different endings: *arki* and *päivä*. Kustaa Vilkuna explains that originally all workdays were named *arki*. However, due to the Orthodox Christian influence, Wednesday could not be called merely 'the third workday', because it had a great religious importance as a fast day. So it had to be named more neutrally as

kolmaspäivä, the ‘third day’. Afterwards, the ‘fourth’ and the ‘fifth workday’ analogically became *neljäspäivä* ‘the fourth day’ and *viiespäivä* ‘the fifth day’.

There are identical ordinal-based principles also in Estonian and Livonian, but the names of Friday, Saturday and Sunday are based on other principles. It is very possible that the ordinal based naming system was borrowed from the Baltic languages.

	Estonian	**South Estonian**	**Livonian**
Mon	*esmaspäev*	*ǰispäiv*	*eʾžžõmpääva*
Tue	*teisipäev*	*tõõsõpäiv*	*tuuožnapääva*
Wed	*kolmapäev*	*kolmapäiv*	*kuolmõndpääva*
Thu	*neljapäev*	*nelläpäiv*	*nel'l'õndpääva*
Fri	*(reede)*	*(`riǰde)*	*(breed'õg)*
Sat	*laupäev*	*(puul'päiv /-püha)*	*(puuolpääva)*

In many cases, Wednesday is named as ‘the center of the week’. This can be seen in the Germanic languages, like German (*Mittwoch*), Icelandic (*miðvikudagur*) and Old Norse (*miðvikudagr*). Slavic languages also have a similar naming principle: Russian *среда* (< Old Church Slavonic *sr̀ěda*) originally meant ‘the heart of the week’.

Russian *среда* is widely adopted in the Eastern Finnic languages: Izhor *serreeta*, Karelian *śeŕoda*, Votic *sereda*. It usually replaced the older name meaning ‘the 3rd day’.

Finnish *keskiviikko* appears to be an original Finnish or Finnic name, but probably it is a loan translation from Old Norse (*miðvikudagr*). In Old Swedish, the name of Wednesday was ‘the day of Odin’. However, as the *middeweke* of the Middle Low German indicates, old Scandinavians could also have such name for Wednesday, which was adopted by the Finns and later replaced by ‘the day of Odin’.

The form *kesknädala* or *keskviiko* ‘center of the week’ predominates also in North Estonian dialects, unlike in Standard Estonian, where *kolmapäev* ‘the 3rd day’ is used.

It is Friday that most usually has a name borrowed from Russian (Karelian *piätinčä,* Izhor *päätetsä,* Votic *päätnittsa* < Russian *пятница*) and even if it has an indigenous name, it is not *arki* but *päivä*. However, it has been observed that the most remote Izhor and Votic dialects have ‘the fifth day’ rather than ‘the fifth

workday'. In northern Karelian *neljäspäivä* 'the fourth day' meaning Thursday has also been attested.

3. The weekend

	Friday	Saturday	Sunday
Finnish	*perjantai*	*lauantai*	*sunnuntai, pyhä*
Karelian	*piätinč(č)ä*	*šuovatta, suovattu*	*pühäpäivä*
Izhor	*päätetsä, päättentsä viijespäivä*	*soovatta*	*pühä*
Vepsian	*pätnič*	*sobat*	*pühapäi*
Votic	*päätnittsa, vijjespäivä*	*lauk(o)päivä, subott(a)*	*pühä(päivä)*
Estonian	*reede*	*laupäev*	*pühapäev*
Livonian	*breed´õg*	*puuolpääva*	*pivaapääva*

Saturday in Finnish, Estonian and Votic is adopted from the Germanic peoples (< Old Norse *laugardagr*, Proto-Germanic **Laugō dagaz*). Its meaning was 'bath day' (it is interesting that Saturday is still the most important day for sauna in Finland).

The rest of the Finnic languages have the Slavic-Russian *суббота*: *suovatta, šuovatta, suovattu, soavattu* in the Karelian dialects; *sobat* in Vepsian and *soovatta* in Izhor. Russian (and Slavic) *суббота* has its origins in the Hebrew *Shabbat*.

The most interesting variants are Southern Finnic 'half days': *puuolpääva* in Livonian and *puul'päiv* / *puul'püha* 'half day' / 'half holiday' in Southern Estonia. These names demonstrate the idea of Saturday being only a semi-workday before the great weekly holiday of Sunday.

Finally, the most important day of the week, the weekly holiday, Sunday. In Finnic languages there are two naming principles for it: *sunnuntai* in Finnish and *pyhä/püha/pühä* in the rest of the Finnic languages. The form *sunnuntai* is used only by Finns in Finland or in the historical Ingria (by Ingrian Finns in Ingria, in Finnish: *Inkeri,* in Russian: *Ingermanlandiya*).

The other name for Sunday, *pyhä* 'holy, sacred (day)' is also widely used in the Finnish dialects. In Finland, the difference between *sunnuntai* and *pyhä* depends on

the usage: usually Sunday was called *pyhä* in Finnish, but if there happened to be another holiday or feast-day directly before or after it, they were distinguished from each other by calling that Sunday *sunnuntai*. This means that every *sunnuntai* was *pyhä*, but not every *pyhä* was *sunnuntai*.

In the other Finnic languages, Sunday is exclusively *pühä/püha*. In spite of the strong Russian influence, the only Finnic language that has borrowed Russian Sunday *nedel'nik* is Votic, and even there, *nätilpäivä* is not as common as *pühä*.

4. Dating the names of the days

According to the ethnologist Kustaa Vilkuna (1959), the oldest name of the day in the Finnic languages is the *pyhä* 'Sunday'. It was the sacred day, taboo day, with several social restrictions. It was the most important day of the week: therefore, it got a name before the other days. Originally, according the lunar calendar, *pyhä* was the day without the Moon, and as the new moon appeared, the new week had begun (and this day was called 'the Moon's day' or 'the first day' in several languages).

Most likely the old Finnic peoples had a name only for Sunday while the rest of the days were numbered, as it is still the case on the periphery of the Finnic languages.

Traditionally, the names of the days in Finnish are considered to be Old Swedish loanwords dating from between the 9th and the 13th centuries. However, three of them – *lauantai*, *perjantai* and *sunnuntai* – were considered even older Germanic loanwords. Hakulinen (1946) found it strange that the rest of the names of the days would have been adopted later from Old Swedish (which developed from Old East Norse in the early 13th century).

In Old Swedish, Saturday was *löghardagher*, as in Finnish it is *lauantai* (in Estonian: *laupäev*, in Votic: *laukopäivä*). The Old Norse diphthong *au* was monophtongized to *ö* in the 11th century and, therefore, *lauantai* had to be adopted before then.

The *-n-* in *maa**n**tai*, *perja**n**tai*, *laua**n**tai* and *sunnu**n**tai* is etymologically problematic. There was no *-n-* in the corresponding forms of Old Norse (*mánudagr, frijādagr*, *laugardagr*, *sunnudagr*), yet there was one in **sunnu**n**dag* in Old Low German and in *sunnû**n**tag* in Old High German and apparently also in **mānu**n**dag* in Old Low German. Old Low German (and also Old Saxon) was spoken between the 9th and the 13th centuries on the southern coasts of the North Sea and the Baltic Sea. This is why we can assume that the ancient Finns adopted the names of the days in the period of Old East Norse (between 800 and 1100) – or even before then.

Phonologically, the most problematic is *perjantai*. If it had been borrowed from Old (East) Norse (*frijādagr*) or Old Swedish (*frēadagher*), it would be phonologically adapted as "reejatai", "reijatai" (or something like that, as Estonian *reede*). It would be expected that the consonant cluster *fr* in the beginning of the word only lose the first element, as it happened with the rest of the Scandinavian clusters as they were borrowed into Finnish (*strand* > *ranta, slaktare* > *lahtari, stall* > *talli*).

According to Bentlin (2008), the problem of *perjantai* can be solved with the hypothesis of Old Low German **perindag* (in Old Bavarian there was also *pferintag*). The problematic *-n-* mentioned above originates probably from *perjantai* and *sunnuntai*, and the *-n-* in *maanantai* and *lauantai* could be an analogy (otherwise it would be expected that *maanantai* was *maanatai* and *lauantai* was "laua**r**tai", cf. Old Norse *lauga**r**dagr* and Old Swedish *mānadagher*, *lögha**r**dagher*).

In the historico-(cultural) context, the adoption of the names of the days can be seen as part of the Scandinavian influence, which became intense at the same time as the Viking Age had begun. The Western Finnish tribes were subordinated by the 12th century, but the Christian influences had arrived even earlier. It is known that there were Christians living in Swedish Birka in the 9th century.

The ordinal based naming principle of the days predominates in the rest of the Finnic languages. This is also why it has been connected with the Baltic or Slavic influence. However, it is possible that it is a relic of the indigenous Finnic week system, as mentioned above.

By the 6th century, Proto-Finnic consisted of three main dialects, the Western (in present-day Southern Finland), the Southern (present-day Estonia), and the Eastern (on the west coast of Lake Ladoga). Eastern Proto-Finnic developed later into the Karelian and Izhor languages.

The proto-Karelians were highly mobile, even expansive people, and they populated present-day Karelia and the west coast of the White Sea by the 12th century. At the same time, or even earlier, the first Slavic influences were identified in the archaeological digs. They arrived from Kievan Rus, Christianized in the 10th and 11th centuries, so the first Christian influence arrived in Karelia probably also in the 11th century. Russian domination was established slowly by the 13th century. The first border between Sweden and Novgorod was drawn in 1323 by the Treaty of Pähkinäsaari (Nöteborg, Oreshen). The Pähkinäsaari border separated the Finns and the Karelians both linguistically and culturally from each other.

The new border was not only political, it was also cultural, economical and religious between the Roman Catholic and the Orthodox Churches.

References

ALFE = Tuomi, Tuomo (ed.) 2004. *Atlas Linguarum Fennicarum I. Itämerensuomalainen kielikartasto. Läänemeresoome keeleatlas. Ostseefinnisher Sprachatlas. Лингвистический атлас прибалтийско-финских языков.* Helsinki: Suomalaisen kirjallisuuden seura.

Bentlin, Mikko 2008. *Niederdeutsch–finnische Sprachkontakte*. Helsinki: Suomalais-Ugrilainen Seura.

Hakulinen, Lauri 1946. *Suomen kielen rakenne ja kehitys. Toinen osa: sanasto-ja lauseoppia*. Helsinki: Kustannusosakeyhtiö Otava.

Häkkinen, Kaisa 2004. *Nykysuomen etymologinen sanakirja*. Helsinki: WSOY.

Maticsák, Sándor 2006. *Vándorló napok. A hét napjainak elnevezései az európai nyelvekben*. Debrecen: Kossuth Egyetemi Kiadó.

Nirvi, R. E. 1971. *Inkeroismurteiden sanakirja*. (Lexica Societatis Fenno-Ugricae XVIII). Helsinki: Suomalais-ugrilainen Seura.

Oja, Heikki 2013. *Aikakirja 2013. Helsingin yliopiston almanakkatoimisto*. Helsinki. (Also in the Internet: http://almanakka.helsinki.fi/images/aikakirja/Aikakirja2013kokonaan.pdf)

SKES = Toivonen, Y. H. et al. (eds.) 1955–1978. *Suomen kielen etymologinen sanakirja*. (Lexica Societatis Fenno-Ugricae XII) Helsinki: Suomalais-ugrilainen Seura.

SSA = Erkki, Itkonen and Ulla Maija, Kulonen (eds.) 1992–2000. *Suomen sanojen alkuperä.* Helsinki: Suomalaisen Kirjallisuuden Seura.

Vilkuna, Kustaa 1959. Viikon vanhinta historiaa. *Kalevalaseuran vuosikirja* 39. Helsinki: WSOY, 257–291. [Also in: Vilkuna, Kustaa (ed.) 1989: *Kansatieteilijän työpöydältä. Kokoelma Kustaa Vilkunan tutkimuksia ja kirjoituksia.* Helsinki: Suomalaisen Kirjallisuuden Seura. 130–158.]

VKS = Grünberg, Silja (ed.) 2013. *Vadja keele sõnaraamat. Vaďďaa tšeele sõnatširja. Словарь водского языка*. 2., täiendatud trükk. Eesti Keele Instituut, Eesti Keele Sihtasutus, Tallinn.

The possessive plural marker in the Burgenland dialect of Hungarian in Austria

Hajime Oshima

Research Institute for Languages and Cultures of Asia and Africa (ILCAA)
Tokyo University of Foreign Studies (TUFS)

1. Introduction

1.1 Historical Background of Burgenland

Historically, Hungary has had a deep and complex relationship with Austria, including the period of the Austro-Hungarian dual monarchy (1867–1919). Before World War I, the province of Burgenland was Hungarian territory within Austria-Hungary, but in the aftermath of the war, the area, together with its Hungarian speakers, came under the rule of a newly independent Austria. However, a Hungarian minority still exists in the province, the members of which are bilingual in Hungarian and German (because the largest ethnic group in Burgenland is German and the national language of Austria is German). According to Austria's 2001 national census, there were 6,641 "speakers of Hungarian" in the province, or about 2.4% of its total population.[1] Hungarian speakers are mainly found in the town of Oberwart (Felsőőr in Hungarian) and the neighboring towns (Unterwart/Alsóőr and Siget in der Wart/Őrisziget). The dialect has been surrounded by the German language, ie. the Indo-European languages. In other words, this Hungarian dialect of Oberwart (Felsőőr) and neighboring towns is a "language island" in the sea of different language family.

The Hungarian speaking population of Oberwart has been declining recently, as is evident from the census data shown in Table 1.

[1] http://www3.umiz.at/de/index.php/ueberuns/ungarn-im-burgenland/demografie

Year	Hungarian	German	Mixed	Croatian or other	Total
1880	2,701	999			3,700
1910	3,039		1,148		4,187
1920	3,138		965		4,103
1934	2,176	2,008			4,833
1939	1,482				
1951	1,603	2,854	577		4,713
1961	1,206	3,011	424	99	4,740
1964	1,934	2,726			
1971	1,486		4,175		5,661

Table 1. Number of speakers of each language in Oberwart (1880–1971)
(Gal 1979: 26)

In 1920, 76.5% of the total population of Oberwart spoke Hungarian, while only 26.3% did in 1971. This downward trend has continued in more recent times as well, as Table 2 shows:

Year	Hungarian	German	Croatian	Romany	Total
2001	1,169	4,889	233	84	6,696

Table 2. Number of speakers of each language in Oberwart
(National census in Austria, 2001)

The Hungarian speaking minority in Oberwart has already fallen to about 17.5% of the total population and can be expected to continue decreasing in proportion.

1.2. Bilingualism in Burgenland

As indicated above, the Oberwart Hungarian speakers are bilingual in Hungarian and German. They live in a German speaking environment and use German in public places, for example, when accessing government services, visiting German speaking doctors or shops, etc. Meanwhile, they speak Hungarian in more private environments, for instance, at home or with friends who are also Hungarian speakers. This kind of linguistic situation is called diglossia. In a diglossic environment, the language variants used in public and private are typically referred

to as High (H) and Low (L), respectively; in this case, German is H and Hungarian is L (Gal 1979).

A point to be noted in regard to this sociolinguistic situation is that the number of Hungarian speakers is in decline. The shift from a rural type of life to urban life is probably one of the main reasons for the decrease of the Hungarian speaking population in Burgenland. When someone moves to the city, he or she will typically have enough incentive to learn the dominant language used there. At the same time, monolingual Germans in Burgenland would find it very difficult to learn Hungarian even if they wanted to since it does not belong to the Indo-European language family. Furthermore, there is the problem of language prestige, as in Burgenland, Hungarian is the L. In these circumstances, it is only natural that younger Hungarians will tend to favor German as the more urban, career enhancing, and business facilitating language over the more "agrarian" and "poor" Hungarian (Romaine 1994: 52). The same kind of reasoning is often an important factor in selecting a spouse. In Oberwart, the term "exogamous marriage" essentially means a German speaker marrying a Hungarian Calvinist. There has been a marked increase in this kind of marriage in the post-war years, as Table 3 demonstrates:

Years	Percentage of Exogamous Marriages	Total Number of Calvinist Marriages
1896-1900	20	66
1901-1905	15	65
1906-1910	22	63
1911-1915	31	45
1916-1920	25	80
1921-1925	23	57
1926-1930	31	59
1931-1935	37	51
1936-1940	29	59
1941-1945	34	47
1946-1950	27	111
1951-1955	38	66
1956-1960	48	64
1961-1965	50	58

1966-1970	82	66
1971	79	14
1972	65	17

Table 3. Percentage of exogamous marriages of Calvinist Oberwarters
(Gal 1979: 52)

Naturally, this trend has additionally sped up the process of Hungarian native speakers moving away from its use.

1.3. Phonetics in the Burgenland Dialect

The greatest phonetic difference between the Burgenland dialect and Standard Hungarian is that the Burgenland dialect distinguishes between narrow [ë] and wide [e]. Figure 1 shows the vowel inventory of the Burgenland dialect.

Vowels				
Short				
		Rounded		
front	*i* [i]	*ü* [y]	*u* [u]	back
	ë [ə]	*ö* [ø]	*o* [o]	
	e [ɛ]	*a* [ʌ]		
Long				
		Rounded		
front	*í* [iː]	*ü* [y]	*ú* [uː]	back
	ië [iə]	*üö* [yø]	*uo* [uo]	
			á [ɑ]	

Figure 1. Vowels in the Burgenland dialect of Hungarian
(Gal 1979: 80)

For example, the indefinite article is *ë*/*ëdzs*, and the definite is *e*/*ez*.[2] Moreover, in terms of consonants, the Hungarians in Burgenland pronounce *cs* [tʃ] / *dzs* [dʒ] instead of *ty* [c] / *gy* [ɟ].

[2] The former precede words beginning with a consonant, and the latter precede words beginning with a vowel.

(1) *ëdzs asszom mëg ë liány*
a woman and a girl
(egy asszony meg egy lány)[3]
'a woman and a girl'

(Imre 1973: 11)

(2) a. *Hun e dzserëk?*
where the child
(Hol van a gyerek?)
'Where is the child?'

b. *Mëk-harap e kucsa!*
PERF-bite the dog
(Megharap a kutya!)
'The dog will bite you!'

(Imre 1973: 21)

The sounds of *ty* [c] / *gy* [ɟ] of Standard Hungarian are very characteristic, at least the German language does not have the sounds. The phenomenon might be result of the language contact between the Hungarian and the German.

2. Possessive plural

The Hungarian possessive is marked for singular and plural. The markers are added to the possessed noun as suffixes. In the next examples, the Standard Hungarian *gyerek* 'child' is the possessee. The possessive singular is exemplified in (3), and the possessive plural in (4).

(3) a. *a Péter gyerek-e*
the Peter child-POSS
'Peter's child'

b. *a gyerek-em*
the child-POSS.1SG
'my child'

[3] The sentence in parenthesis is Standard Hungarian. (The same is done in the following examples).

(4) a. *a* *Péter* *gyerek-e-i*
the Peter child-POSS-PL.3SG
'Peter's children'

b. *a* *gyerek-e-i-m*
the child-POSS-PL-1SG
'my children'.

2.1. Earlier studies on the possessive plural in the Burgenland dialect

In research on the Burgenland dialect, the major earlier studies are by Imre (1971a, 1973). These studies deal with the dialect as spoken in Oberwart/Felsőőr in South Burgenland. The Burgenland dialect is included in the western dialect group of Hungarian (Imre 1971b, Kiss 2001). However, we can find differences between it and Standard Hungarian not only in phonetics and the lexicon but also in some grammatical points, including the possessive plural marker.

Below, I present Table 4 comparing possessive plural forms in Standard Hungarian and the Burgenland dialect (i.e. the Oberwart dialect in Burgenland). The data of the Burgenland dialect is cited from Imre (1971a, 1973) in this table.

	Standard	Burgenland	Standard	Burgenland	Standard	Burgenland
	gyerek 'child'		*lúd* 'goose'		*tehén* 'cow'	
1SG	*gyerekeim*	*dzserëkëmi͡ëk*	*lúdjaim*	*ludami͡ëk*	*teheneim*	*tehenëmi͡ëk*
2SG	*gyerekeid*	*dzserëkëdi͡ëk*	*lúdjaid*	*ludadi͡ëk*	*teheneid*	*tehenëdi͡ëk*
3SG	*gyerekei*	*dzserëkeji*	*lúdjai*	*luddzsaji*	*tehenei*	*tehennyeji*
1PL	*gyerekeink*	*dzserëkünki͡ëk*	*lúdjaink*	*ludanki͡ëk*	*teheneink*	*tehenünki͡ëk*
2PL	*gyerekeitek*	*dzserëkëti͡ëki͡ëk*	*lúdjaitok*	*ludatoki͡ëk*	*teheneitek*	*tehenëti͡ëki͡ëk*
3PL	*gyerekeik*	*dzserëkcsëki͡ëk*	*lúdjaik*	*luddzsoki͡ëk*	*teheneik*	*tehennyëki͡ëk*

Table 4. Contrast between Standard Hungarian and Burgenland dialect possessive plurals

2.2. Observed data

As seen in Table 4, the possessive plural marker *-i͡ëk* is used by Hungarian speakers in Burgenland. In the next example, (5), the two speakers are Hungarians who live in Unterwart/Alsóőr, which is adjacent to Oberwart.

(5) a. *Hogy valaki mond-ja, hogy van kettő vagy három?*
how someone say-3SG.DEF that to be two or three
'How does someone say that you have two or three [children]?'

b. *Van gyerek-em-i͡ëk?*
to be child-POSS.1SG-PL
'I have children?'

To form the possessive plural of the Burgenland dialect, the special marker *-i͡ëk* is added to the possessive singular (e.g. *dzserëk-em* 'my child' + *-i͡ëk*). Therefore, *-i͡ëk* can be called a plural marker. Evidently, it is very different from the possessive plural in Standard Hungarian, *-i*, which was shown in (4).

3. What is *-i͡ëk*? Comparison with the Standard Hungarian *-ék*

As Imre (1971a, 1973) describes and as seen in Table 1, speakers of the Oberwart dialect in Burgenland add *-i͡ëk* to create the possessive plural form. Here, I point out that *-i͡ëk* in the Burgenland dialect as pronounced in Unterwart corresponds to the associative plural marker *-ék* in Standard Hungarian (see Section **3.1**).

3.1. Associative plural in Standard Hungarian

In Hungarian, we find two types of plural. One is the ordinary or additive plural, and the other is the associative plural, meaning noun *X* and his or her family, friends, or associates. (The associative plural also exists in other languages; see 6a below.) As noted above, the marker for the associative plural is *-ék*[4]; it can be added only to human nouns (6), not to non-human animate or inanimate nouns (7).

[4] In Hungarian, the associative plural is called heterogén többség or heterogén többes szám, meaning 'heterogeneous plural(ity).' The suffix *-ék* can be analyzed as the third person singular possessive marker *-é* 'one's' plus the additive plural *-k*. However, nowadays, Hungarians do not recognize the relationship of this form to its original meaning (Balogh 2000: 185).

(6) a. (Japanese)
Tanaka-tachi
Tanaka-ASSOC.PL
'Tanaka and his [her] family, friends, or associates'

b. (Hungarian)
Péter-ék
Peter-ASSOC.PL
'Peter and his family, friends, or associates'

(Moravcsik 2003: 469)

(7) a. **kutyá-ék*
dog-ASSOC.PL
Int. 'the dog and its family, friends, or associates'

b. **szemüveg-ék*
glass-ASSOC.PL
Int. 'the glasses and their family, friends, or associates'

3.2. Conditions for the associative plural

As seen above, the associative plural can be added only to human nouns, and as Balogh (2000) points out, the noun to which it is added must be the focal element of the sentence.

> „Vagyis heterogén többséget alkot minden olyan csoport, amelynek egyik tagja kiemelt személy, s ez a személy az egész csoport reprezentánsának tekinthető."
> (In other words, the heterogeneous plural is formed from every group where one member is a particular person who is considered to be representative of the whole group.)
>
> (Balogh 2000: 186)

In (8a), the word *ember* 'human' is suffixed by the additive plural (Corbett 2000: 102) since the word is not the focal element but homogeneous. *Micimackó* 'Winnie-the-Pooh' is the representative character in the famous story.

(8) a. **ember-ék* *BUT* *ember-ek*
human-ASSOC.PL human-PL (ADDITIVE)
'humans'

b. *Micimackó-ék*
Winnie.the.Pooh-ASSOC.PL
'Winnie-the-Pooh and his friends'

In principle, the associative plural is used with human nouns, as mentioned above. Corbett (2000) accounts for the structure of plurals in Hungarian using the concept of the Animacy Hierarchy (Silverstein 1976) (see Figure 2).

	1 >	2 >	3 >	kin >	human >	animate >	inanimate
range of plural			■■	■■■■■■	■■■■■■■■	■■■■■■■■	■■■■
range of associative plural	■■■■	■■■■■	■■■■	■■■■■■	■■■■		

Figure 2. Associative plurals in Hungarian
(Corbett 2000: 104)

3.3. Associative plural in the Burgenland dialect

In this section, I present examples of the associative plural in the Burgenland dialect of Hungarian to which Imre (1973) referred. In (9) through (13), the associative plural marker is used as a typical case, i.e. the meaning is 'X and his/her family, friends, or associates'.

(9) *Itt aluo, e Szabuo Lajos-i͡ëk-ná má*
here below the Szabó Lajos-ASSOC.PL-ADE already
kü-csap-ott e Pinka.
out-overflow-PST.3SG the Pinka.River
(Itt alul, a Szabó Lajoséknál már kicsapott a Pinka.)
'Here below, the Pinka River has overflowed at Lajos Szabó's house.'[5]

(10) *Be-jár-t e pap-i͡ëk-ho is.*
in-come-PAST.3SG the priest-ASSOC.PL-ALL too
(Bejárt a papékhoz is.)
'S/He regularly visted the priest's house, too.'

[5] *Szabó* is a family name, and *Lajos* is a first name in Hungarian.

(11) *Mi a Böcskör-iëk-je má esztendüö-t át*
we the Böcskör-ASSOC.PL-INST already year-ACC over

összö-segit-ennünk.
together-help-1PL
(Mi a Böcskörékkel már esztendőt át összesegítünk.)
'We have already been helping the Böcskör family for over a year'

(12) *Ha vuona hel-ëk e zsup-nak, e szomszid-iëk*
if to be-COND space-PL the thatch-DAT the neighbor-ASSOC.PL

së szuomáz-ná-ják e rozs-ot.
not not.bind-COND-3PL.DEF the rye-ACC
(Ha volna helyek a zsúpnak, a szomszédék se szalmáznák a rozsot.)
'If there are spaces for the thatch, the neighbors should bind rye into sheaves.'

(13) *Valami vidiki firfi mën-d be e Fülöp-iëk-ho.*
some country man go-PAST into the Fülöp-ASSOC.PL-ALL
(Valami vidéki férfi ment be a Fülöpékhez.)
'Some country man went into Philips' house.'

In (14) through (16), we can find the associative plural with the possessive marker. It appears that the meanings are the possessive plural.

(14) *Fizet-nek röndössen e të árëndás-od-iëk?*
pay-3PL correctly the you leaseholder-POSS.2SG-ASSOC.PL

Ez ennyiëm-ek-je uannyi gond van!
the mine-POSS.1SG-INST so worry to be
(Fizetnek rendesen a te árendásodék? Ez enyémekkel annyi gond van!)
'Do your leaseholders pay on time? There are so many problems with mine!'

(15) *Ha je-dzsü-nek ez unoká-m-iëk, ë*
if away-come-3PL the grandchild-POSS.1SG- ASSOC.PL a

sütis pogácsa nëm is ëlig.
baking scone not too enough

- Mijelüöt je-mën-t-ëk, je-dzsü-jj-etëk!
before away-come-PAST-2PL away-come-IMP-2PL

(Ha eljönnek ez unokámék, egy sütés pogácsa nem is elég. - Mielőtt elmentetek, eljöjjetek!)
'When my grandchildren visit, one ovenful of scones is not enough. Before you leave, you should visit me!'

(16) *Ez innep-ëk-re haza-dzsü-nek e*
the holiday-PL-SUB home-come-3PL the

fi-jam-i͡ëk is.
son-POSS.1SG- ASSOC.PL too
(Ez ünnepekre hazajönnek a fijamék is.)
'Also, my sons come home for the holidays.'

However, in (17), the expression *apádi͡ëk* 'your fathers' might not have a plural possessive meaning. Generally, a person has one father. Consequently, the meaning has to be associative.

(17) *Apá-d-i͡ëk miëg nincsen-nek ithon?*
father-POSS.2SG-ASSOC.PL yet not- PL at home
(Apádék még nincsenek itthon?)
'Are your father and associates not at home yet?'

4. Data of the Burgenland dialect

As seen above, the associative plural *-ék* can be added to human nouns but not to non-human animate or inanimate nouns. In my fieldwork, I collected data to see whether this concept applies to the Burgenland dialect or not. These data are presented in Table 5.

	Word	**Standard**	**Burgenland**
Person		*Péter-ék*	*Ernő-i͡ëk*
Kin	*anya* 'mother'	*anyám-ék*	*anyám-i͡ëk*
Human	*barát* 'friend'	*barátom-ék*	*barátom-i͡ëk*
Animate	*kutya* 'dog'	**kutyá-ék*	*kutyám-i͡ëk*
Inanimate	*szemüveg* 'glasses'	**szemüveg-ék*	*szemüvegem-i͡ëk*

Table 5. Associative plurals in Standard Hungarian and the Burgenland Dialect

Imre (1973) shows examples in which the associative plural can be added or not to non-human animate or inanimate nouns.

(18) *E* *füöd-em-iëk,* *múta* *árëndá-ba* *van-nak,*
the land-POSS.1SG-ASSOC.PL since lease-INE to be-PL

egissze *je-guazosu-t-ak.*
entirely PERF-be.weedy-PAST-3PL
(A földemék, amióta árendában vannak, egészen elgazosultak.)
'Since my lands were leased they got completely weedy.'

(19) *E* *vatkörti* *egissze* *je-vásluo-t-a* *e*
the wild.pear completely away-wear-PAST-3SG.DEF the

fog-am-iëk-at.
tooth-POSS.1SG-ASSOC.PL-ACC
(A vadkörte egészen elváslalta a fogamékat.)
'The wild pear has worn away my teeth completely.'

(20) *E* *burgëndi-m-iëk* *má* *ojjanok,*
the mangel.wurzel-POSS.1SG-ASSOC.PL already such

hom *mëk* *kë* *típ-nyi* *ük-et.*
that PERF must tear-INF they-ACC
(A burgendimék már olyanok, hogy meg kell tépni őket.)
'My mangel-wurzels are already such that they have to be torn.'

(21) *Lë-szánt-ott-ad* *má* *e* *torruo-d-iëk-at?*
down-plow-PAST-2SG.DEF already the stubble-POSS.2SG-ASSOC.PL-ACC
(Leszántottad már a torrúdékat?)
'Have you already plowed your stubbles?'

(22) *Aziër* *e* *pápistá-k-nak* *izs* *van-nak* *szíp*
therefore the papist-PL-DAT too to be-PL beautiful

ínëk-csëk-iëk.
song-POSS.3PL-ASSOC.PL
(Azért a pápistáknak így vannak szép énekeikék.)
'Therefore, the papists have their beautiful songs.'

4.1. Problem of the form

In the Burgenland dialect, not only the associative plural (*-i̯ëk*) but also the possessive singular (*-m* 'my ...s') is attached to all nouns except proper nouns (PERSON in Table 5). However, see (23).

(23) a. *anyá-m-i̯ëk*
mother-POSS.1SG-ASSOC.PL
'my mothers' (= mother, grandmother, great-grandmother, and so on)
OR
'my mother and her family, friends, or associates'

b. *?anyá-i-m-i̯ëk*
mother- PL-POSS.1SG-ASSOC.PL

Generally, a person has only one *anya* 'mother.' The additive meaning of *anyá-m-i̯ëk* is contradictory. For this reason, the meaning of (23a) has to be a special one or associative.

Additionally, my consultants[6] said that the form with the possessive singular and associative plural is more natural than the one with the possessive plural (23b). However, the plural form *anyáim-i̯ëk* is conceivable. Because speakers use the standard form *anyáim* in public space, and the associative plural *-i̯ëk* can be added to it, the result of contact with Standard Hungarian means that *anyáim-i̯ëk* might be used.

4.2. Applicability to non-humans

As seen in Table 5 above, *-i̯ëk* can be used not only with human nouns but also non-human nouns. In this case, the meaning must be one of possession, as in (24).

(24) a. *macská-m-i̯ëk*
cat-POSS.1SG-ASSOC.PL
'my cats'

b. *cipő-m-i̯ëk*
shoe-POSS.1SG-ASSOC.PL
'my shoes'

[6] Three Hungarian speaking interviewees in Oberwart, two in Unterwart; one in his 20s, the others in their 50s and older. The young Hungarian speaker said that the associative marker can be attached only to human nouns, as in Standard Hungarian.

c. *haj-am-i͡ëk*
hair-POSS.1SG-ASSOC.PL
'my hair[s]'

d. *szem-em-i͡ëk*
eye-POSS.1SG-ASSOC.PL
'my eyes'

e. *föld-em-i͡ëk*
land-POSS.1SG-ASSOC.PL
'my lands'

Given my data on the associative plural in Hungarian, the figure by Corbett (2000), presented as Figure 2 above, can be changed as below, in Figure 3.

	1 >	2 >	3 >	kin >	human >	Possession (animate, inanimate)
range of associative plural	■■■■■■	■■■■■■	■■■■■■	■■■■■■	■■■■■■■■	■■■■■■■■■■■■■■■■■■■

Figure 3. Associative plurals in the Burgenland dialect

4.3. Hypothesis of *-i͡ëk*

Essentially, non-human animate or inanimate nouns have no focal element. Therefore, the meaning of the examples in (24) is not associative. This result is parallel with (23a) for human nouns, although proper nouns (PERSON) have to be able to carry the meaning of the associative plural, as demonstrated by (25).

(26) *Ernő-i͡ëk*
Ernő-ASSOC. PL
'Ernő and his family, friends, or associates' (not 'persons by the name of Ernő')

To summarize the present discussion, I have found that *-i͡ëk* is not only a plural possessive marker but also an associative plural marker. The criteria are (1) whether the noun is human or not, (2) whether the possessive marker may attach or not, and (3) whether the noun has a focal element or not.

Noun	Possessive	Associative Plural	Meaning	Examples
Human		-i͡ëk	Associative	(9)-(13)
	POSS	-i͡ëk	Additive or Associative	(14)-(16), (23)
Non-Human		-i͡ëk	Additive	—
	POSS	-i͡ëk	Plural Possessive	(18)-(22), (24)

Table 6. The uses and meaning of -*i͡ëk* in the Burgenland dialect

5. Conclusion

This analysis has examined the possessive plural marker -*i͡ëk* in the Burgenland dialect of Hungarian as spoken in Oberwart and Unterwart, Austria, and showed that it has different meanings from the ordinary possessive plural marker used in Standard Hungarian, namely, that -*i͡ëk* is also used as an associative plural marker. In Standard Hungarian, the associative plural marker is -*ék*, and it can be attached only to human nouns, not (other) animate or inanimate nouns. By contrast, -*i͡ëk* in Burgenland dialect can be added to non-human nouns if the meaning is one of "possession." However, in this case, the meaning is not associative but an ordinary plural. Since the associative plural needs a focal element, non-human nouns with -*i͡ëk* cannot take an associative meaning. However, proper nouns with -*i͡ëk* have to be interpreted with the associative meaning. Finally, I have summarized the uses and meanings of -*i͡ëk* in the Burgenland dialect. The results of previous discussions clearly show that -*i͡ëk* is not only a plural possessive marker but also an associative plural marker. However, the forms are affected by certain criteria.

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
ADE	adessive
ALL	allative
ASSOC	associative
COND	conditional
DAT	dative
DEF	definite conjugation

IMP	imperative
INE	inessive
INF	infinitive
INST	instrumental
PERF	perfective
PST	past
POSS	possessive
PL	plural
SG	singular
SUB	sublative
TER	terminative

Acknowledgements

I wish to thank László Kelemen (Director of the Hungarian Media- and Information Center[7]) and László Gúthy (Minister of the Calvinist Church in Oberwart) for their assistance with my research.

This research was supported by the "Linguistic Dynamics Science Project 2"[8] of the Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies (TUFS).

References

Balogh, Judit 2000. A névragozás [Noun inflection]. In: Keszler, Borbála (ed.) *Magyar grammatika* [A grammar of Hungarian]. Budapest: Nemzeti Tankönyvkiadó. 183–208.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Gal, Susan 1979. *Language shift: Social determinants of linguistic change in bilingual Austria*. San Fransisco: Academic Press.

Imre, Samu 1971a. *A felsőőri nyelvjárás* [The Oberwart dialect], Nyelvtudományi értekezések 72. Budapest: Akadémiai kiadó.

Imre, Samu 1971b. *A mai magyar nyelvjárások rendszere* [The system of present-day Hungarian dialects]. Budapest: Akadémiai kiadó.

[7] Ungarisches Medien- und Informationszentrum (UMIZ) / Magyar Média és Informaciós Központ (http://www3.umiz.at/de/)

[8] The official name of the project is Project for Building an International Network of Collaborative Research on Endangered Linguistic Diversity
URL: http://lingdy.aacore.jp/en/index.html

Imre, Samu 1973. *Felsőőri tájszótár* [Oberwart dialect dictionary]. Budapest: Akadémiai Kiadó.

Kiss, Jenő (ed.) 2001. *Magyar dialektológia* [Hungarian dialectology]. Budapest: Osiris Kiadó.

Moravcsik, Edith 2003. A semantic analysis of associative plurals, Studies in *Language* 27: 469–503.

Romaine, Suzanne 1994. *Language in society: An introduction to sociolinguistics.* Oxford: Oxford University Press.

Silverstein, Michael 1976. Hierarchy of features and ergativity. In: Dixon, R. M. W. (ed.) *Grammatical categories in Australian languages.* Canberra: Australian Institute of Aboriginal Studies. 112–171.

Szolcsányi, Ákos 2007. A heterogén többes szám [The heterogeneous plural]. *Magyar Nyelvőr* 131/4: 467–470.

Russian impact on northern Khanty conditional sentences

Mária Sipos

Research Institute for Linguistics, Hungarian Academy of Sciences
Budapest

1. Introduction

As has been shown in numerous investigations of language contact, the borrowing of function words, specifically, conjunctions, may have syntactic consequences. At the same time, it is often emphasised that borrowing grammatical elements and following syntactic patterns do not necessarily correspond (Aikhenwald 2008: 15).

The present paper surveys the patterns conveying conditional content in Synya Khanty (Ob-Ugric, Uralic) texts, as well as the proportion of conditional sentences contaning the Russian conjunction *jesli* is discussed in the context of its present day role and occurrence of the Khanty marker of conditional sentence.

When talking about the impact of language contacts, following Matras & Sakel's terminology (2007b), MAT (matter) and PAT (pattern) borrowings, which denote the two basic ways of borrowing (Sakel 2007b: 15), are differentiated.

We speak of MAT-borrowing when morphological material and its phonological shape from one langugage is replicated to another language. PAT describes the case where only the patterns of the other language are replicated, i.e. the organization, distribution and mapping of grammatical or semantic meaning, while the form itself is not borrowed (Sakel 2007b: 15).

The relation of the two concepts has been investigated in numerous case studies (Aikhenwald 2008: 16, Grenoble 2000: 109–110), and it has also been emphasised that PAT borrowing is possible without MAT borrowing (Aikhenwald 2008: 15).

In Finno-Ugric linguistics, it is MAT borrowings that were generally collected, whereas for a long time the syntactic consequences of language contacts were much less often researched. Thus, it has been well known for decades that Finno-Ugric languages had borrowed a lot of conjunctions from Russian (see e.g. Maytinskaya 1983; Alvre 1983; Leinonen 2002). Maytinskaya (1983: 187) listed the Russian loan

conjunctions that were present in the majority of the Finno-Ugric languages of the Soviet Union: *i* 'and', *a* 'but', *il'i* 'or', *no* 'but', *jesli* 'if', *što* 'that'. Besides these, numerous Finno-Ugric languages also borrowed *khotja* 'although', *libo* 'or', *štoby* 'so that'. Alvre (1983) published a similar list of the borrowings of the Baltic Finnic languages, thus the lists in these two papers have been cited since then (Leinonen 2002: 254–255; Rießler 2007: 241).

In the investigated texts, there are more than thirty function words borrowed from Russian, among which conjunctions, several types of adverbs etc. can be found. As for conjunctions, the following appear: (subordinate) *što* 'that', *štoby* 'so that', *jesli* 'if'; coordinate: *i* 'and', *a* 'but', *il'i* 'or', *no* 'but'. As *kak budto* 'as if' only appears in one single sentence, we cannot exclude that it is a result of code switching. Compared to the frequency of the rest of conjunctions, the number of the occurence of *il'i* 'or' is greater by orders of magnitude due to the fact that the favourite expression of the speaker to display uncertainty is *il'i mŭj* 'or what', where it has a grammatical role other than linking clauses.

As has been mentioned, conditional *jesli* is also among the loan conjunctions in Khanty. In northern Khanty, where – as opposed to the southern and eastern dialect groups (Riese 1984: 101–113) – conditional relations are expressed with subordinate sentences without conjunctions, the intrusion of Russian *jesli* can indeed be foretold with much certainty. Conversely, in the chapter on northern Khanty Riese does not mention *jesli*. First, among the reasons must be the fact that the sources processed by Riese (1984) reflect the language of a period 2–8 decades earlier. Second, ideals and considerations behind the publication of linguistic material were quite different in the first half of the 20th century, therefore fieldworkers might have been attracted to "pure" Khanty language displaying no Russian impact. Third, folk genres, although not excluded, are less likely to use Russian borrowings than spoken Khanty.

At the same time, the conjunction in question does not occur in Éva Schmidt's Kazym Khanty texts collected in the 1990s and published in the 2000s (Khomlyak 2002). Among these texts there are not only folk tales or songs but also spontaneous texts, which, although undoubtedly having been told several times, still lack expressions characteristic of folk genres, furthermore, they exhibit a considerable number of Russian loans. As Éva Schmidt's intention was to produce an authentic written version of the speech production of the speakers influenced by neither grammatical nor dialectal expectations, it seems probable that in the northern dialects, or at least in the language of her speakers, the conditional conjunction of Russian origin had no special importance.

There are numerous classifications and categorisations of conditional sentences, e.g. on the basis of logic or linguistics. The aspect of reality of conditional content is also often discussed. In the literature, there is an abundance of terms referring to the two main types of conditional sentences i.e. **factual** vs. **counterfactual**, sentences of **open** vs. **rejected** condition, **realis** vs. **irrealis** (e.g. Riese 1984: 16). Furthermore, formal linguistic characteristics such as markedness, markers, order of clauses, the presence or lack of subordinators or correlatives, the use of tenses etc. are investigated. In the present paper my aim is to survey the markers of northern Khanty conditional sentences occuring in the speech production of a single speaker, in the context of Russian impact on the Khanty language. The tense of the sentences, the position of the Khanty conditional particle *ki*, as well as the ordering of clauses are outside the scope of this research. Considering the fact that the conditional sentence type expressing the unreal condition scarcely appears in texts, the focus of the present paper is the formation of sentences expressing real conditions in the northern Khanty dialects.

The paper aims at answering the following questions:

1. To what extent is the Russian conjunction *jesli* present in Khanty conditional sentences?
2. Is there any difference in the use of conditional sentences between the traditional northern Khanty texts and the spontaneous speech production of a present-day bilingual speaker?
3. Are there double marked sentences, i.e. ones containing conjunctions of both Russian and Khanty origin, in great number in the corpus?
4. What is the proportion of sentences (i) following the traditional Khanty pattern vs (ii) innovative constructions?

The paper investigates the conditional sentence patterns in the following steps. First, the Khanty language, the linguistic material, and the speaker are introduced (in Sections 1–3). Then the typical forms of conditional sentences in the Khanty and Russian languages are summarised (see Section 4). Section 5 describes the conditional sentences types appearing in the corpus. Results and conlusion are summarised in the final section (6).[1]

[1] I am also grateful to an anonymous reviewer, who greatly helped me to improve my paper with numerous valuable comments.

2. The Khanty language

Khanty is spoken in the Khanty–Mansi Autonomous District and the Yamalo-Nenets Autonomous District, and in the Tomsk Oblast, in Western Siberia, Russia. According to the 2010 Russian census, the ethnic population amounts to 30,900, while Khanty is spoken by only 9,580 speakers (Ethnologue[2]). The Khanty language has three dialect groups and a large number of subdialects which differ significantly from each other. Consequently, northern, southern (extinct), and eastern Khanty are often considered closely related yet separate languages. The variety to which the texts of the present investigation belong to is spoken by the Synya River, a western tributary of the Ob River, and is very close to the lects spoken in the following settlements along the western banks of the Ob River: Muzhi, Khantymuzhi, Vosyakhovo, Ust'-Voykar, Unselgort, and Shuryshkary.

An agglutinative language, Khanty employs SOV word order. On the basis of old folklore texts, it was considered to use nonfinite subordination as opposed to finite subordinate sentences, the latter being relatively new: especially finite subordinate sentences with conjunctions have begun to develop in recent times (Schmidt 2008: 49).

3. On the speaker, data, and corpus

The investigations are carried out on the basis of the Khanty text material that was collected by Ruttkay-Miklián as a result of fieldwork with a Synya Khanty speaker in the 2000s.

The speaker was born in 1946 in a village by the upper Synya, and she did not leave this region during her life. Similarly to her husband, she spoke the Synya subdialect of northern Khanty. Having been widowed, she raised her children alone, and moved to the regional centre, Ovgort, where she came into contact with a less traditional world. Her language is therefore not archaic but represents the knowledge of a Khanty living in a bilingual settlement, speaking her dialect very well, and having proficiency in Russian at the same time (for further details, see Ruttkay-Miklián 2008).

In the course of data collection, the speaker was asked to explain the meaning of given words, as well as to give examples in order to reveal each word's meaning, or to describe the use of the word. Furthermore, she had to produce every utterance

[2] http://www.ethnologue.com/country/RU/languages

intuitively, not being influenced by either linguistic or ethnographic expectations, i.e. neither the dictionary, nor the fieldworker. With this method, the whole northern Khanty material, especially the Synya Khanty entries, of Steinitz' dictionary (Steinitz 1966–1988) was processed by Ruttkay-Miklián, and this resulted in a 70-hour audio recording. The transcribed version of the audio recordings, which amounts to about 127,200 words, was not corrected or edited later either, so the linguistic material can be characterised as spontaneous speech consisting of texts of different length. It contains self-corrections, hesitations, fillers, contracted forms, ellipsis, repetitions, non-standard grammar etc., which might have importance when researching certain features of language use (Grenoble 2012: 102).[3]

4. Conditional sentences in the Khanty and Russian languages

4.1. Types of conditional sentences

Every language is able to express conditional content, i.e. condition–consequence relation of two events, facts, or factors etc.

The diversity of conditional sentences in the languages of the world originates in the different degrees of markedness, and the great number of the possible markers range from parataxis, in which the conditional relation of the unmarked clauses is suggested by the context, to the multiple marking of the conditional relations (through conjunctions, correlatives, tenses etc.) (Bakró-Nagy 2008; Veltmann 1994: 683). In this paper, a relatively broad definition is used, viz. in the conditional construction there is a statement (Lat. *apodosis*, Eng. *consequent*) whose realization depends on the fulfilment or the verity of the other part of the construction (Lat. *protasis*, Eng. *antecedent*) (Bakró-Nagy 2006: 1, Veltmann 1994: 683).

The texts being examined are definitely of descriptive character, the conditional sentences mainly present implications between facts, thus the use of rejected conditional sentences is not characteristic of them.

4.2. Conditional sentences in northern Khanty

In his monograph, Riese (1984: 102) characterises the conditional sentences of northern Khanty in the following way (confining the linguistic forms mentioned in the original to Synya Khanty elements). The sentences of open condition can be

[3] I am grateful to Eszter Ruttkay-Miklián for making her text collection, with its English and Hungarian translations, available for me.

expressed in three ways (in the clause containing the apodosis there is no correlative):

A. [protasis with the conditional particle *ki* 'if'] + [apodosis];
B. [protasis with neither a conditional particle nor conditional conjunction] + [apodosis];[4]
C. [protasis with the conjunction *xundi* / *xun* 'when', *xundi-ki* 'if'] + [apodosis].

The above types (as for C, subtypes are due to different conjunctions) are illustrated below.

In sentence (1), the protasis contains the conditional particle *ki* 'if':

(1) *ope-n-ŋ̣* *esəmjiŋk-ən* *ăt* *ki* *jeś-l-a,*
daughter-2SG-LOC milk-LOC NEG COND drink-PRS-PASS

ma *pošχ-em-a* *mij-i!*
1SG child-1SG-LAT give-IMP2SG

'If milk is not drunk by your daughter, give it to my child!'

(Steinitz 1975: 149, cited by Riese 1984: 103)

A paratactic conditional sentence can be seen in (2):

(2) *mŏlti* *pŏraj-ŋ̣* *pa* *jŏχət-l-ən* *sa* *ma*
some.kind.of time-LOC again come-PRS-2SG PTCL 1SG

χŏśe-m-a *śărγtə-ti,*
PPOS-1SG-LAT make.shaman.foretell-INF

ma *śi* *pŏraj-ŋ̣* *năŋen* *jastə-ti* *jasəŋ* *tăj-l-əm.*
1SG that time-LOC 2SG.LAT say-INF word have-PRS-1SG

'Wenn du irgendwann mal wieder zu mir kommst, mich schammanaisieren zu bitten, dann werde ich ein Wort mit dir zu sprechen haben' [If, sometime, you come to listen to my predictions, then I will have something to say to you]'

(Steinitz 1975: 73, cited by Riese 1984: 104)

The following conditional sentence[5] (3) contains the conjunction *χun* (*χǫn*):

[4] This structure is also called paratactic. In paratactic conditional sentences, the relationship of the clauses is inferred from the text.

[5] The sentence is from a text from Kazym Khanty, which is, similarly to Synya Khanty, belongs to the northern Khanty dialects. The word *χun* 'where' is thus spelt *χǫn* here, and *ʌ* is also characteristic of the Kazym dialect.

(3) *χǫn* *meńńe* *ăŋki-ja,* *aśi-ja* *weŋ-xǫ* *pŏrməs-ŏt*
when bride mother-LAT father-LAT bridegroom-man thing-PL

ăn *mǫstə-ʌ-ət,*
NEG be.liked-PRES-3PL

ewe-ʌ *ăn* *mă-ʌ-eʌ* *ʌŭweʌa.*
daughter-3DU NEG give-PRS-3PL.O 3SG.LAT

'Wenn die Gegenstände des Brätigams der Mutter un Vater der Braut nicht passen, geben sie ihm ihre Tochter nicht. [If the mother and father of the bride do not like the gifts of the bridegroom, they do not give him their daughter.]'

(Rédei 1968: 44; cited by Riese 1984: 106)

In sentence (4), there is a compound conjunction consisting of an interrogative / relative pronoun *χundi*, which is the equivalent of *χun* 'when' in the northernmost Khanty dialects and the conditional particle *ki* 'if':

(4) *χundi-ki* *ol-da* *ant* *raχ-l-əm,*
if be-INF NEG be.allowed-PRS-1SG

sem-em *χol* *pit-l,* *si* *man-l-əm.*
eye-1SG where fall-PRS.SG3 there go-PRS-1SG

'Ha pedig majd itt nem maradhatok tovább, amerre a szemem lát [tkp. esik], arra megyek. [If I cannot remain here any longer, I will go wherever I throw my glance]'

(Pápay 1910: 91; cited by Riese 1984: 106)

Among the above mentioned structures, the unmarked paratactic structure seems to be original, in which the order protasis + apodosis is, of course, dominant. It is also well known that the conditional particle *ke*/*ki* of the Khanty and Mansi languages was borrowed from Komi. The particle *ki* can occur in almost any position in the sentence except in the clause initial one, although the most frequent position for this particle in Khanty is the clause final position, or the one preceding the predicate. In the case of the order **protasis** + **apodosis**, the particle *ki* obligatorily appears in the sentence. In the northern Khanty dialects surveyed by Riese (1984), the sentences containing *ki* greatly outnumber the rest of the types of conditionals, i.e. they make the 89% of total (Riese 1984: 101–106).

4.3. Conditional sentences in Russian

Conditional in Russian is expressed with a complex sentence consisting of a conditional clause (protasis) and a subjunctive clause (apodosis).[6] The typical conjunction (occuring in clause initial position) is *если бы* or *если* depending on the type of conditional. If it expresses unreal condition, i.e. the action in the subjunctive clause cannot take place because the condition in the conditional clause cannot be realized, then it will begin with the conjunction *если бы*, followed by the verb in the past tense. The subjunctive clause has a verb in the past tense and the particle *бы* (Mitrevski) (5):

(5) *Если* *он* *разбуд-ит* *жен-у,* *она*
if he.M wake-3SG.PRS wife-ACC 3SG.F

рассерд-ит-ся.
get.angry-3SG.PRES-REFL

'If he wakes his wife, she will be angry.'

(Wade 2011: 333)

The sentence can begin with either clause (6) (7) (Mitrevski):

(6) *Если* *бы* *у* *меня* *был-и* *деньги,* *я*
if PTCL PREP 1SG be.PST-PL money 1SG

поеха-л *бы* *на* *юг.*
travel-PST.SG.M PTCL PREP south

'If I had money, I would go to the south.' (Mitrevski)

(7) *Я* *поеха-л* *бы* *на* *юг,*
1SG travel-PST.SG.M PTCL PREP south

если *бы* *у* *меня* *был-и* *деньги.*
if PTCL PREP 1SG be.PST-PL money

'I would go to the south if I had money.'

(Mitrevski)

Russian also has conditional expressions with conditional clauses that are realizable; these complex sentences do not use the particle *бы*. In these sentences,

[6] The formation of Russian conditional sentences is summarised on the basis of Wade 2011 and Mitrevski http://www.auburn.edu/~mitrege/russian/tutorials/0048.html**.**

the condition in the conditional clause upon which the situation in the main clause depends is possible and realizable (8) (9).

(8) *Я* *тебе* *позвон-ю,* *если* *ты* *буд-ешь* *дома.*
1SG 2SG.DAT call-PRS.1SG if 2SG be.FUT-2SG at.home

'I'll call you if you are at home.' (Mitrevski)

(9) *Если* *буд-ет* *хорош-ая* *погода,*
if be.FUT-3SG good-F weather

мы *поед-ем* *в* *парк.*
1PL go-1PL PREP park

'If the weather is nice, we'll go to the park.'

(Mitrevski)

As for the previous type of conditional clauses, i.e. those expressing unreal condition, it allows two kinds of interpretations (Wade 2011: 333) (10):

(10) *Я* *пошё-л* *бы,* *если* *бы* *меня* *пригласи-ли.*
1SG go-3SG.M PTCL if PTCL 1SG.ACC invite-PL.PST

'I would go if they invited me.'
'I would have gone if they had invited me.'

(Wade 2011: 333)

4.4. Borrowing foreign forms and patterns

Foreign forms and patterns make their way into the target language in several different ways (Aikhenwald 2008: 22–26). It can be a form simply transferred from one language into another one, there may be an enhancement of an already existing feature, extension by analogy, reinterpretation and reanalysis, areally induced grammaticalization, grammatical accommodation, or loan translation. Finally, lexical or grammatical parallelism is mentioned by Aikhenwald, which means that between typologically different languages it may happen that the pattern of the target language and that of the source language appear in one and the same clause / sentence (Aikhenwald 2008: 25). By means of a representative example in Tetun Dili (a Tetun based creole language in East Timor), which is in contact with Portuguese, Hayek (2008: 170) presents a stage named **lexical pairing** by him, in which "native and borrowed elements appear optionally together". The structure combining two patterns "allows a smooth transition from the older native structure to the newer, less complex one" (Hajek 2008: 170).

There are examples of similar mechanisms in Finno-Ugric languages as well (see Kaysina 2013, Tánczos 2013, Sipos 2014). Among these we can find cases where the conjunction of Russian origin appears in clause-initial position while the (original) conjunction of the target language, with a similar function, can also be found in the sentence, in a different position Jefremova 2013: 188; Kaysina 2013: 140; Tánczos 2013, Sipos 2014: 90–92). A similar phenomenon can be observed in the case of the conditional particle borrowed from Russian and the particle *ki* 'if' in Khanty.

In what follows, I will discuss the conditional structures and their diversity in the examined texts.

5. Conditional sentences in the corpus

Due to the descriptive and explanatory character of the corpus, the conditional sentences appearing in it in great number express general truths, facts about natural, psychological or social phenomena that always take place in similar ways, events with the *if-then* logical relation between them.

In the texts elicited from Ruttkay-Miklián's speaker, numerous types of sentences expressing conditional content can be observed. First, there is an abundance of examples of paratactic constructions, as well as sentences containing the Khanty particle ki. Second, there are clauses introduced by the conjunction of Russian origin. Furthermore, there are clauses containing the conjunction of Russian origin and the common Khanty conditional particle at the same time (*jesli… ki*). In addition to these, the same pattern can be found introduced by the following Khanty conjunctions: *χŏn* 'when'… *ki*, and *χŏta* 'where'… *ki*. Finally, there is a pattern in the corpus in which the conditional particle *ki*, which, normally, can never be found in clause initial position, shows up in both intitial and final positions in the clause at the same time (*ki*… *ki*). The sentence type described by Riese, in which the conjunction is *χŭn* 'when', does not occur in the corpus. I will now discuss these structures in detail.

5.1. Paratactic constructions

In paratactic constructions, the relation of the two clauses is merely logical, i.e. conditionality is not marked grammatically. This is why the typical order is protasis + apodosis, see (11)-(12):

(11) *ŏχ* *ăntəm* *in,* *wan,* *letŏt* *ŭ-ti* *śir-en* *ăntəm.*
money NEG now look food buy-INF way-2SG NEG

'If you haven't got money now, look, you can't buy food.'

(Ruttkay-Miklián 2010)

(12) *a* *pa* *śir-ən* *jasti-l-a,* *śit* *atəm* *jasəŋ.*
but different way-LOC say-PRS-SG3.PASS it bad word

'But if it is said in another way, it's a nasty word.'

(Ruttkay-Miklián 2010)

5.2. Sentences containing the conditional particle

In both language materials, i.e. in those investigated by Riese as well as in the texts produced by Ruttkay-Miklián's speaker, the most frequent sentence type is the one with a protasis containing the Khanty conditional particle *ki* (13):

(13) *nŏ,* *šŭw,* *ulten* *jastl,* *il* *pit-l* *ki,*
well fog generally let's.say down fall-PRS.3SG PTCL

tŏrəm *jăm-a* *ji-l,*
weather good-LAT become-PRS.3SG

nŏx *katləs-l* *ki,* *jert-a* *ji-l,* (...)
PREV thicken-PRS.3SG PTCL rain-LAT become-PRS.3SG

'Well, generally the fog, let's say, if it falls, the weather will be good, if it thickens, it will rain…'

(Ruttkay-Miklián 2010)

However, it should be noted that besides being the marker of conditional content in the protasis, the particle *ki* has a further function in Khanty, as it can also express uncertainty and low probability. Still, the two functions can easily be differentiated.

This modal use can be observed in the situations when the speaker was not sure whether she understood the word she had to explain, or whether she was able to give a sketch of its meaning or use, i.e. the wording expresses a kind of uncertainty. One of the numerous sentences of this type is (i):

(i) *śit* *moś* *jasəŋ* *ki.*
that tale word PTCL

'It might be a tale-word.'

(Ruttkay-Miklián 2010)

Similarly, in the following example, *ki* suggests uncertainty. It is suitable for illustrating the function in a question as the sentence expresses deduction ('it was cold when you arrived, mosquitos might have died'), but not condition + consequence (*'it was cold when you arrived, if there were no mosquitos') (ii):

(ii) *năŋ* *jŏχət-m-en*-[Ø] *ta* *iśki* *us,*
2SG arrive-PST.PRTL-2SG-[LOC] then cold be.PST.3SG

pelńa *xŏla-s* *ki.*
mosquito end-PST.3SG PTCL

'When you arrived it was cold, the mosquitos must have ended!'[7]

(Ruttkay-Miklián 2010)

In this function, it also appears in code-switching sentences, as in (iii)-(iv):

(iii) *lŏʟpi,* *śiməś* *karti,* *mŭj...* *swińec,*
lead this.kind.of iron/metal what lead(Ru)

kăk *năziwa-jet,* *swińec* *ki.*
how(Ru) call(Ru)-3SG(Ru) lead(Ru) PTCL

'Lead, a kind of metal, what… lead, what is it called, maybe lead.'

(Ruttkay-Miklián 2010)

(iv) *sŏwa* *tăm* *mŭj,*
gizzard this what

năwrena(!) *kăk* *eta* *počkaj-en* *iti* *ki,*
probably(Ru) as(Ru) this(Ru) kidney(Ru)-2SG in.the.way.of PTCL

'What is a gizzard, most likely it is like this kidney maybe, (…)'

(Ruttkay-Miklián 2010)

[7] The context of this sentence is the following: "The other day you went to the village, don't those people use net tents? Now the mosquitos are gone, I guess. Yes. They don't use them. In the summer it is full of mosquitos, they must use them. When you arrived it was cold, the mosquitos must have left! That's why they don't have net tents there." The conversation is undoubtedly about the possible causes of why the fieldworker is not familiar with net tents against mosquitos, i.e. the topic of this speculation is not the weather. This is why the interpretation, otherwise seeming possible, 'It must have been cold when you arrived, if there weren't any mosquitos' can be excluded. I am grateful to my anonymous reader for warning me about this ambiguity.

In order to survey the proportion of these two functions, the two types of its appearance have been counted in a section of the whole text, i.e. in the comments belonging to the dictionary entries beginning with *s*. It amounts to 19 full pages, the number of words is ca. 13,500, where *ki* functions as a conditional particle in 58 sentences, while in 24 cases it is used for expressing probability.

5.3. Conditional conjunction borrowed from Russian

In a further type of conditional sentences, illustrated in (14) below, a clause containing the conjunction of Russian origin can be seen, where apodosis precedes protasis. In case of this ordering, conditional relation must be grammatically marked (NB: this is the only example of this clause order):

(14) *χulm-a* *jŏχaʀsə-l.*
three-LAT ramify-PRS.3SG

ăntəm *kătn-a,* *χulm-a,* *jesli* *χuləm* *juš!*
NEG.PTCL two-LAT three-LAT if(Ru) three way

'[the road] goes into three directions. Not into two but three, if it's three roads.'

(Ruttkay-Miklián 2010)

5.4. The conditional conjunction and the conditional particle occuring together

5.4.1. The Russian conjunction *jesli* + Khanty particle *ki*

The conjunction *jesli* also appears in sentences containing the conditional particle *ki*. The sentence in (15) is not a prototypical conditional sentence but it illustrates the broad semantic area between conditionals and time clauses:

(15) *nŏ,* *mŏlti* *săχat,* *nŏ* *jesli* *pelt-s-en* *ki,*
well something for, well if(Ru) exchange-PST-2SG.O if

năŋ *jasti-l-ən:*
you say-PRS-2SG

ma, *ma* *śi* *săχat* *pelt-s-em* *tăme-m.*
I I it for change-PRS-1SG.O this.one-1SG

'Well, for something, well, if you exchanged it you say: I've exchanged this for that.'

(Ruttkay-Miklián 2010)

5.4.2. Khanty conjunction *χŏn* + Khanty particle *ki*

The particle *ki* can also occur within the same clause with Khanty conjunctions as well. As the semantics of conditional sentences and time clauses cannot always be easily differentiated as can be observed in numerous languages, the occurrence of *χŏn* 'when' in the protasis is not surprising, e.g. (16):

(16) *uχəl-en, lŭw χŏn lŏpas-en jem ki ăt*
sleigh-2SG it when pantry-2SG prohibition PTCL NEG

tăj-l, nŏ nŏχ χuχ-ti ki ăt raχ-əl,
have-PRS.3SG well up climb-INF PTCL NEG be.allowed-PRS.3SG

wante, śi oməs-l-en uχl-en.
look PTCL stand-PRS-2SG.O sleigh-2SG

'The sleigh, if the pantry is not forbidden, well, if you can't climb on it, look, you stand your sleigh there.'

(Ruttkay-Miklián 2010)

As Pápay's texts indicate (Riese 1984: 105), in the northernmost Khanty dialects there is a compound conjunction *χundi-gi* 'if'. It consists of the same elements, i.e. *xunti* in the Obdorsk subdialect is the equivalent of the Synya Khanty interrogative and relative pronoun *χŏn* 'when'; while *gi* is the Obdorsk equivalent of Synya Khanty *ki* 'if' with a voiced consonant. However, the syntactic environment, i.e. the positions of *ki*/*gi* in the clause, might have differed from present-day Synya Khanty patterns. The way of development in the case of the Obdorsk Khanty conjunction is out of the scope of the present paper.

As has already been mentioned, the borderline between time clauses and protases may not be semantically clear, consequently the occurence of *ki* in a clause containing the conjunction *χŏn* 'when' is not surprising (17):

(17) *nŏ, amp-ət χŏn jŏt-l-ət ki, śiti*
well dog-PL when play-PRS-3PL PTCL in.this.way

jast-l-a, atəm tŏrəm-a ji-l.
say-PRS-PASS.3SG bad weather-LAT become-PRS.3SG

'When dogs play, they say we'll have bad weather.'

(Ruttkay-Miklián 2010)

Although the next sentence (18) can be interpreted in two ways (a-b), it should also be classified in the transitional category described above:

(a) 'When a rope is cut, (then) you tie them, (and) – you say – "I bind them"';

(b) ‘When a rope is cut (and) you tie them, then – you say – “I bind them”’.

(18) *χŏn kel toχ-əl ki,*
when rope tear-PRS.3SG PTCL

jăχa jăr-l-en, jastl: oľ-ľ-em.
together tie-PRS-2SG.O he.says join-PRS-1SG.O

‘When the rope is cut you tie it together, you say, I am joining it up.’

(Ruttkay-Miklián 2010)

5.4.3. Khanty conjunction *xŏta* ‘where’ + Khanty particle *ki*

There is another conjunction cooccuring with the particle *ki*, which is *χŏta* ‘where’. There are three such examples in the corpus, and this low number makes it difficult to answer the question whether in these sentences *χŏta* ‘where’ should be considered a conditional conjunction, or whether we are dealing with relative clauses in which the particle is present in clause final position due to analogy. Both of the following examples (19)–(20) consist of loosely connected and fragmentary clauses, so it is difficult to categorize them.

(19) *ar-sir soχəl ul. χŏta năŋ ki woχ-l-ən,*
many-kind board to be-PRS.3SG where you PTCL cut-PRS-2SG

lŏpsaχ-a pa ji-l, pa soχəl.
flat-LAT PTCL become-PRS.3SG also board

‘There are several types of boards. If you just cut it with an ax like this, and if it is flat as well, that is also a board.’

(Ruttkay-Miklián 2010)

(20) *śit tăŋχa lŭw mŭj, sŭt, lŭw śiməś*
that.one perhaps it what whet-stone it that.kind

kew, atel kew, ăntəm ăl kew-šup, lŭw
stone, separate stone NEG simple stone-piece, it

śiməś kew, sŭt, χŏta keši lŏχət-ti
that.kind stone whet-stone where knife sharpen-PRS.PRTL

χŏr-pi ki, iľi lajəm lŏχət-l-ən, păsti-ja
shape-ADJ PTCL or axe sharpen-PRS-2SG sharp-LAT

ji-l.
become-PRS.3SG

‘That most likely is a whetstone, a kind of stone, a separate stone, not a simple piece of stone, but a kind of stone, sharpening stone, where/if it is of the form of a whetstone, or you sharpen an ax and it becomes sharp.’

(Ruttkay-Miklián 2010)

The following sentence (21) is, however, appropriate for syntactic evaluation.

(21) (Context: “Hard, hard, that’s usually fur, or mostly used when talking about fur. Or the ground is hard, sometimes you cannot dig into it and you say: the ground is hard.”)

χŏta	*jăm-a*	*šit-l*	*ki,*	*śit*	*mŭw-əl*	*lepət.*
where	good-LAT	be.possible-PRS.3SG	PTCL	that	earth-3SG	soft

‘Where it can be [dug] the ground is soft.’

(Ruttkay-Miklián 2010)

As Ruttkay-Miklián confirmed the clause beginning with *χŏta* could be a relative clause, only if the clause initial *śit* was *śita* ‘there’ (Ruttkay-Miklián personal communication 2015). In that case the original English translation of the sentence would be ‘Where it can be [dug], there the ground is soft’. In the given form, it seems to be a conditional sentence ‘If it can be dug the ground is soft’, which is in accordance with the context, i.e. the speaker had to make clear the word meaning ‘hard’. In any case, it should be noted that ki cooccurs with a conjunction other than *χŏn* or *jesli*, which is not mentioned as a potential source of conditional markers at Heine and Kuteva (2002: 329), so it needs further investigations.

5.4.4. Khanty *ki* ‘if’ as a conjunction + *ki* ‘if’ as a particle

In the protasis of the next sentence, a conjunction and a particle seem to be present at the same time. However, as opposed to the previous sentences, in clause initial position we have the particle *ki*, which is expected to occur in a position any other than this, and appears once more within the same clause in clause final position. This construction may be a mixture of the Russian and Khanty patterns. The primary marker of conditionality, not mentioning paratactic sentences here, is undoubtedly *ki*, which appears not only in one of its traditional positions but also clause-initially, which may be an influence of the Russian language abounding in conjunctions in general.

(22) *ki* *jăχa* *ar-šək* *sŏχ* *χŏśa* *ŏl-l* *ki,*
PTCL together many-COMP thing to lie-PRS.3SG PTCL

sŏχ *śi.*
stuff indeed

'When more things are lying together, that's stuff, indeed.'

(Ruttkay-Miklián 2010)

As this is the only example of this pattern, it might be a one-time construction, or a slip-of-the-tongue, which is not unusual in spontaneous speech. In Ruttkay-Miklián's opinion, this construction is hardly acceptable for Khanty speakers (Ruttkay-Miklián: personal communication, 2015).

5.5. The distribution of various conditional structures

The above examples (11)–(22) present various formations of conditional sentences, suggesting that, in addition to traditional patterns, sentences displaying MAT or PAT borrowing appear in great quantities. Clearly, in order to judge the actual importance of the innovative types, it is inevitable to know their ratio in the corpus.[8]

First, the number of occurrences of each recent type (11)–(22) concerning the whole corpus will be given (**5.5.1**). Due to the size of the corpus, the figures referring to the traditional types regarding the whole material will be estimated on the basis of one single file (**5.5.2**). Then (**5.5.3**), the proportions will be compared to those of Riese (1984: 102, 104), which were calculated on the basis of traditional texts.

5.5.1. How the individual sentence types are represented in the whole corpus

The following table displays the actual numbers of sentences presented above, on the basis of the whole corpus (Table 1) (uncertain i.e. fragmented or ill-formed sentences are included in the numbers in brackets):

Conjunctions / conjunction and particle	**Occurrence**
jesli…..	3
jesli … ki	2 (3)

[8] As Ruttkay-Miklián's texts are stored in distinct files distinguished by the initials of the entries processed in them, the ratio of the individual sentence types within the whole corpus can be estimated by a rate calculation.

χŏn … ki	5 (8)
χŏta …ki	2 (3)
ki… ki	1
Total	13 (18)

Table 1. Sentences not following traditional Khanty patterns

As can be seen in Table 1, *jesli* functioning as the only marker of the conditional is documented in only three sentences. As it is combined with the particle *ki* in three further sentences, *jesli* occurs in 6 sentences all in all.

5.5.2. The estimated proportion of traditional sentence types

Table 2 displays the size of the text (given in number of words) in the *s-* file as well as in the rest of the material.

	s- file	other files
Number of words	13,500	113,700

Table 2. Number of words in the texts to be compared

On the basis of the actual counts in the *s-* file, the approximate number of sentence types in the whole text can be estimated (Table 3):

Marking of conditional sentences	***s-* file (actual count)**	**other files (estimated)**	**total (estimated)**
Sentences containing *ki*	58	ca 487	ca 545
Paratactic sentences	15	ca 126	ca 141

Table 3. Calculated numbers of conditional sentence types

5.5.3. A comparison of old folklore texts to recent spontaneous speech production

Last but not least, the proportion of the two types of traditional formation can be compared to the proportions given by Riese (1984: 102, 104). Table 4 summarizes the percentage of the different types of conditional sentences in the texts commenting the dictionary entries beginning with *s-*.

	Dictionary entries beginning with s-		**Riese (1984)**
	Number	%	%
Particle *ki*	58	78	89
Parataxis	15	20	10
Other	1	2	1

Table 4. Proportion of the different types of conditional sentences

To sum up, both comparisons show that despite the relatively high number of innovative types, it is still the two traditional sentence types that appear in the great majority of the conditional sentences in question. In other words, from this point of view, Ruttkay-Miklián's speaker can still be qualified as a traditional speaker. Even if the problematic sentences in (15)–(17) as well as the uncertain data are included, the proportion of non-traditional ways of expressing conditional content is not higher than 2%.

6. Conclusion

Having surveyed the conditional sentences of the given texts, the questions listed in the introduction can be answered in the following way.

In the texts of a middle-aged bilingual speaker the number of occurrence of the conjunction *jesli* is much lower than expected considering the surveys reporting *jesli* to have been borrowed in Khanty by the 1980s.

The informant seems to be a traditional speaker from the perspective of the way she expresses conditional content (the investigation, adapted to the specialities of the corpus, was restricted to conditional sentences referring to general truths). On the one hand, it is because she produces a greater percentage of paratactic sentences than the texts in Riese's investigation (Riese 1984: 102, 104). On the other hand, the particle *ki*, which was the most typical marker of conditional sentences before the intense Russian influence, counts as the most typical one even today. As in many other languages, a double marked conditional sentence (i.e. the simultaneous use of a conjunction and a particle which cannot be in clause initial position) evolved due to favourable syntactic circumstances. However, there are only a small number of sentences of this type. This innovative construction also appears in clauses beginning with conjunctions other than *jesli*, or Khanty *xŏn* 'when'. However, on

the basis of the available examples, it is impossible to define the grammatical function of the *ki* particle in these sentences.

Concluding on the basis of the above data, the impact of Russian conditional sentences can be observed in both the domains of MAT and PAT borrowings. As for MAT borrowing, the Russian conjunction has appeared in Khanty sentences, although it occcurs quite rarely. PAT borrowing is also possible to detect in the sentences in which the clause of protasis contains or actually begins with a conjunction of Khanty origin, i.e. the changes in syntax cannot be said to be the consequence of borrowing *jesli* from Russian. Besides, the overwhelming majority of the conditional sentences produced by the speaker do follow the two traditional patterns, i.e. paratactic subordinate sentences and the ones in which conditional content is marked by a particle.

Abbreviations

ACC	accusative
ADJ	adjective
COMP	comparative
F	feminine
FUT	future
IMP	imperative
INF	infinitive
LAT	lative
LOC	locative
M	masculine
NEG	negative
O	objective conjugation
PASS	passive
PL	plural
PPOS	postposition
PREV	preverb
PRS	present
PRS.PRTL	present participle
PST	past
PST.PRTL	past participle
PTCL	particle
PURP	purpose

RU Russian
SG singular

References

Aikhenwald, Alexandra Y. 2008. Grammars in contact: A cross-linguistic perspective. In: Aikhenwald, Alexandra Y. and Dixon, R. M. (eds) *Grammars in contact*. Oxford: Oxford University Press.

Alvre, Paul 1983. Vene laenudest uurali keelte konjunktsioonides. In: Janhunen, Juha; Peräniitty, Anneli and Suhonen, Seppo (eds.) *Mémoires de la Société Finno-Ougrienne* 185: 25–50.

Bakró-Nagy, Marianne 2006. Grammatikai kölcsönzés és grammatikalizáció. A komi feltételes *kẹ* partikula a manysiban. (Grammatical borrowing and grammaticalization: The Komi conditional particle *kẹ* in Mansi]. In: Sipos, Mária (ed.) *Obi-ugorok a 21. században* [Ob-Ugric peoples in the 21st century] (CD). Budapest: MTA Nyelvtudományi Intézet

Bakró-Nagy, Marianne 2008. *Conditional constructions in Uralic.* http://www.univie. ac.at/urtypol/bakro.pdf [Accessed: March 2015]

Ethnologue = http://www.ethnologue.com/country/RU/languages [accessed: March 2015]

Grenoble, Lenore A. 2000. Morphosyntactic change: The impact of Russian on Evenki. In: Gilbers, D. G., Nerbonne, J.and Schaeken, J. (eds.) *Languages in contact*. Amsterdam: Rodopi. 105–120.

Grenoble, Lenore A. 2012. Areal typology and syntactic change. *Вестник ТГПУ* (TSPU Bulletin) 2012 (1): 101–104.

Hajek, John 2008. Language contact and convergence in East Timor: The case of Tetun Dili. In: Aikhenwald, Alexandra Y. and Dixon, R. M. (eds.) *Grammars in contact*. Oxford: Oxford University Press. 163–178.

Khomlyak 2002. = Хомляк, Л. Р. *Арем-моньщэм ел ки манл... Если моя сказка-песня дальше идёт...* [If my song, my tale goes on…] Ханты-Мансийск: Групп Полиграфист.

Jefremova, Tatyjana 2013. Mondattani kontaktusszerkezetek megítélése és használata városi marik beszédében. In: Csepregi, Márta; Kubínyi, Katalin, and Siivonen, Jari (eds.) *Grammatika és kontextus. Új szempontok az uráli nyelvek kutatásában* 3. [Grammar and context: New perspectives in research into the Uralic languages, vol. 3]. Budapest: Eötvös Loránd Tudományegyetem, Finnugor Tanszék. 181–193.

Kaysina, Inna 2013. The adoption of Russian conjunctions in Udmurt. *Journal of Estonian and Finno-Ugric Linguistics* 4 (2): 131–144.

Leinonen, Marja 2002. Influence of Russian on the syntax of Komi. *Finnisch-Ugrische Forschungen* 57: 195–358.

Majtinskaja 1983. = Майтинская, К. Е. Заимствованные служебные слова в финно-угорских языках. In: Janhunen, Juha; Peräniitty, Anneli and Suhonen, Seppo (eds.) Symposium saecularae Societatis Fenno-ugricae. *Mémoires de la Société Finno-Ougrienne* 185: 179–188.

Mitrevski, George: *Conditional mood.* http://www.auburn.edu/~mitrege/russian/tutorials/0048.html [Date of access: March 2016]

Pápay, József 1910. *Északi-osztják nyelvtanulmányok.* [Northern Khanty linguistic studies]. Budapest: Magyar Tudományos Akadémia.

Rédei, Károly 1968. *Nordostjakische texte (Kazym-Dialekt) mit Skizze der Grammatik*. Göttingen: VandenHoeck and Ruprecht.

Riese, Timothy 1984. *The conditional sentence in the Ugrian, Permian and Volgaic languages.* Studia Uralica 3. Wien: Verband der wissenschaftlichen Gesellschaften Österreichs (VWGÖ).

Rießler, Michael 2007. Grammatical borrowing in Kildin Saami. In: Matras, Yaron and Sakel, Jeanette (eds.) *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter. 31–74.

Ruttkay-Miklián, Eszter 2008. *The dialect spoken by Roza Makarovna.* http://hantisirn.nytud.hu/sites/default/files/Dialect%20of%20Roza%20Makarovna2.pdf. Date of access: March 2016.

Ruttkay-Miklián, Eszter 2010. *Synya Khanty dictionary by Roza Makarovna*. Manuscript.

Sakel, Jeanette 2007a. Language contact between Spanish and Mosetén. A study of grammatical integration. *International Journal of Bilingualism* 11 (1): 25–53.

Sakel, Jeanette 2007b. Types of loan: Matter and pattern. In: Matras, Yaron and Sakel, Jeanette (eds.) *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter. 15–31.

Schmidt Éva 2008. Északi osztják nyelvtani jegyzet (serkáli nyelvjárás) [Northern Khanty grammar (Sherkaly dialect)]. In: Fejes, László (ed.) *Serkáli osztják chrestomathia* [Sherkaly Khanty chrestomathy]. Schmidt Éva Könyvtár 3. Budapest: MTA Nyelvtudományi Intézet. 13–75.

Sipos Mária 2014. *Célhatározói mondatok idiolektális variánsai egy hanti–orosz kétnyelvű beszélő szövegeiben* [Intraspeaker variants of purpose clauses in the texts of a Khanty–Russian bilingual speaker]. *Nyelvtudományi Közlemények* 110: 79–97.

Steinitz, Wolfgang 1975. *Ostjakologische Arbeiten. Band I: Ostjakische Volksdichtung und Erzählungen aus zwei Dialekten. Texte.* Budapest: Akadémiai Kiadó.

Steinitz, Wolfgang 1966–1988. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie Verlag.

Tánczos Orsolya 2013. Hogy… hogy? Kettős kötőszók az udmurt mondatban [How come? Double conjunctions in Udmurt sentences]. In: Agyagási, Klára; Hegedűs, Attila and É. Kiss, Katalin (eds.) *Nyelvelmélet és kontaktológia 2* [Linguistic theory and language contact 2]. Piliscsaba: PPKE BTK Elméleti Nyelvészeti Tanszék – Magyar Nyelvészeti Tanszék. 95–112.

Thráinsson, Höskuldur 2013. Ideal speakers and other speakers: The case of dative and other cases. In: Fernández, Beatriz and Etxepare, Ricardo (eds.) *Variation in datives: A micro-comparative perspective.* Oxford Studies in Comparative Syntax. Oxford: Oxford University Press. 161–188.

Veltmann, Frank 1994. Conditionals. In: Asher, R. E. and Simpson, J. M. Y. (eds.) *The encyclopedia of language and linguistics*. Oxford and New York: Pergamon Press.

Wade, Terence 2011. *A comprehensive Russian grammar*. Oxford: Wiley and Blackwell.

A language without ‘get’?

Katalin Sipőcz

University of Szeged

The verb ‘get’ belongs to the most frequent “basic” verbs. Additionally this verb is a common source of grammaticalization in many languages. Thus, it may seem suprising that in the Mansi language there is no verb meaning ‘get’. The notion and situation of getting, together with the notion of giving, naturally can be expressed in Mansi, too, and in the expressions for the concept of getting the verb ‘give’ is used. The verb ‘give’ is probably the most frequent ditransitive verb in the languages of the world (Haspelmath 2013). Thus, this phenomenon is connected to the use of ditransitive constructions. In this paper I intend to describe and analyze such constructions in Mansi.

In the first part of my paper I provide a short description of the Mansi language (**1.1**) then I define the term ditransitivity and discuss the main typological aspects of ditransitivity (**1.2**). Next, I describe the ditransitive constructions of the Mansi language (**2.1** and **2.2**) and the main rules of their usage (**2.3**). Finally, I discuss the constructions expressing the event of ‘getting’ (**3**).

1.1. The Mansi language

The Mansi (or Vogul) language is an endangered Uralic language. Nowadays Mansi is spoken by fewer than 1,000 people, however, more than 12,000 people declare themselves to be of Mansi nationality (cf. Table 1).

In our time the only Mansi dialect that is still spoken is Northern Mansi, and this dialect also serves as the basis of the Mansi literary language. As a consequence of this situation, the term Mansi is usually used as referring to the Northern Mansi dialect. In this paper I use data from the Northern dialect so I also use the term Mansi referring to this variety. (Northern) Mansi people live in a few villages by the Lower Ob and its western tributaries, the Sosva and Sygva rivers in the Khanty-Mansi Autonomous District of the Tyumen Region of Russia, as well as by the Lozva River in the Ivdel Area of the Sverdlovsk Region. This dialect is currently

threatened by the process of language shift to Russian, almost all of its speakers are bilingual.

MANSI	LANGUAGE PROFICIENCY
12,269	938 (7.6%)

Table 1. Ethnically Mansi population and language proficiency according to data from the 2010 Russian census.
(http://www.perepis-2010.ru/results_of_the_census/results-inform.php)

1.2. Ditransitive constructions and ditransitive verbs

Ditransitive constructions are defined as argument structure required by the ditransitive verb, containing the verb itself, the agent (A), the recipient (R) and the theme (T) (Malchukov et al. 2010: 1). Compare:

English			Hungarian			
Mary gave	*John*	*a book.*	*Mari*	*könyvet adott*	*Jánosnak.*	'id.'
Mary told	*John*	*a story.*	*Mari*	*mesét mondott*	*Jánosnak.*	'id.'
A	R	T	A	T	R	

Ditransitive verbs are three-argument verbs which typically denote physical transfer (give, send, bring, etc.). If other verbs with similar semantic features are also used in similar constructions, they are included in the group of ditransitive verbs as well. For instance, such verbs are verbs of communication, as seen in the examples above. In Mansi the number of verbs occurring in ditransitive constructions is high. In addition to the transfer and mental transfer verbs the benefactive and instrumental verbs are also characterized by the same argument structure (Sipőcz 2015).

1.2.1. The main typological groups of the ditransitive constructions

The typological categorization of ditransitive constructions is based on the comparison of ditransitives with the categorization of monotransitive constructions. We differentiate between construction types taking into account whether the T or the R argument of the ditransitive verb occurs in the same position as the patient (P) of the monotransitive construction. On the basis of this, we can differentiate between three main construction types: (1) indirect object construction (IOC) / indirective alignment, in which marking of the P and T is the same, (2) secondary object construction (SOC) / secundative alignment, in which marking of the P and R is the

same, and (3) double object construction (DOC) / neutral alignment, in which both T and R are marked the same way as P (Malchukov et al. 2010: 2–8). Cf.:

(1) *Mari gave **money** to her son.* T = P (≠ R) Cf.: *Mari is counting **money**.*
T P

(2) *Mari supplied **the guests** with food.* R = P (≠T) Cf.: *Mari is expecting **guests**.*
R P

(3) *Mary gave **John a book**.* R = T = P Cf.: *Mary saw **John**.*
R T P

Further types which are logically also possible but can be disregarded due to their minimal occurrence are the so-called tripartitive (T ≠ R ≠ P) and horizontal (T = R ≠ P) constructions (Malchukov et al. 2010: 6). Finally, it must be mentioned that there are two further kinds of ditransitive constructions that are impossible to fit into the above classification. These types are not based on the comparison of monotransitive and ditransitive clauses, the indirective and secundative characters are however clearly distinguishable in their cases, too. These are the serial verb construction and the possessive construction (Malchukov et al. 2010: 11–15, Sipőcz 2015).

2.1. Mansi ditransitive constructions

Mansi is a language with more than one ditransitive construction. These constructions are the indirect object construction and secondary object construction.

(i) In Mansi, in the indirect object construction, the theme (T) of the ditransitive construction is the object, and the recipient (R) is encoded with the lative[1] suffix. The nominal object is in the nominative case and the (personal) pronominal object is in the accusative case.[2] The verb can be in the subjective (4) or objective (5) conjugation.

(4)	*am*	*ōs*	*χūrəm*	*sāt*	*sajt*	*naŋən*	*mīγ-əm*
	I	again	three	hundred	ruble	you-LAT	give-1SG

[1] The *-n* lative suffix has both lative and dative functions, furthermore, the agent of the passive construction is marked by this same suffix. I use the abbrevation LAT in glossing independently the function of the suffix.

[2] In contrast to the other Mansi dialects there is no accusative case in Northern Mansi. In each dialect the personals pronouns have a distinct accusative form.

akw‘ *ēt* *ūnl-ən-ən* *māγəs.*
one night sit-AN-SG for
‘I’ll give you 300 rubles for another night’s watching.’
(VNGy IV: 336)

(5) *akw’* *sup-ä* *kaťi-tä-n* *mi-s-tä*
one piece-3SG cat-3SG-LAT give-PST-SG.3SG
‘S/He gave one piece to his/her cat.’
(VNGy IV: 343)

(ii) In the secondary object construction, the R of the ditransitive construction is the syntactic object and the T is marked with the instrumental suffix. In this construction the verb is almost always in the objective conjugation.

(6) *Maṅ* *piγ-əm* *nē-γəl* *viγ-ləm.*
little son-1SG woman-INSTR take-SG.1SG
‘I will find a wife for my youngest son.’
(VNGY IV: 324)

2.2. Passivization of the ditransitive constructions

From the typological perspective it can be observed that the alignment of passivization often follows the general alignment of encoding. If a language uses a secundary construction, then most probably it will use a secundative alignment in passivization as well. In Mansi both the indirective and the secundative alignment can passivize. The passivization of the indirective construction leads to T-passivization (7), and passivization of the secundative construction always results in R-passivization (8), thus in other words the P-like arguments can be passivized:

T-passivization from an indirective construction:

(7) *jārm-ən* *ta-ke* *maj-wä-s-əm*
poverty-LAT that-PTCL give-PASS-PST-1SG
‘It is poverty that I was given to.’
[‘It is poverty that I was made to experience.’]
(VNGy IV: 330)

R-passivization from a secundative construction:

(8) *(tan) tōnt tax ōs akw Buran-əl mi-w-et.*
(they) then PTCL PTCL one Buran-ISTR give-PASS-3PL
‘Then they got [=were given] one more new Buran.’
(Dinislamova 2007: 11)

The R-passivization is crosslinguistically more common, and also in Mansi it is more frequent (Bíró and Sipőcz 2016).

2.3. Ditransitive alternation in the Mansi language

Several languages have more than one ditransitive constructions. This phenomenon is called alternation and is well-known from English (where it is often called dative shift), e.g.: *Mary gave a pen to John. / Mary gave John a pen*. In English the indirective and the neutral alignments alternate.

In the Mansi language we can see the alternation of the indirective and secundative types. This type of alternation is cross-linguistically more common than the alternation found in English (Malchukov et. al. 2010: 18). Regarding alternation the main task is to describe the rules of the choice between the different constructions. Typological studies mention several factors like the markedness of the arguments, the prominence differences between the T and R arguments, and the topicality of the arguments. There may be even semantic difference between the alternating constructions, etc. It is also common that several factors work together in a language (Malchukov et al. 2010: 20–21).

Mansi seems to belong to the type of languages in which the alternation is related to topicality. The alternation of the ditransitive constructions together with their passivization is part of a system in which the use of different conjugations (subjective or objective) and constructions is in connection with the information structure of the sentence (Skribnik 2001). According to this, the function of promoting the arguments to subject position by the means of passivization or to direct object position by the means of the alternation of the active constructions is to express the relative topicality of different noun phrases within a clause. As a result, T and R occur alternately in the subject or object position. The subject of the sentence is also the (primary) topic, whereas the object functions either as a secondary topic or as a focus. (By topic I mean a previously mentioned contextually or situationally given information, cf. Dalrymple and Nikolaeva 2014: 48–57.) The topicality of the object is marked by the objective conjugation of the verb.

In example (9) all arguments are new information except the A, consequently the predicate agrees only with the subject expressing the A (thus, the verb is in the subjective conjugation). Example (10) represents the case in which the A and the T are given participants, and the R is new information. Thus, the verb must agree with the A and the T. Consequently, the IOC is used where the T is the syntactical object, the verb agrees with it in the objective conjugation. And in example (11), besides the A the R is also a given participant, and the T is the new information. Consequently, the SOC is used, in which the R is the syntactic object which the predicate in the objective conjugation agrees with.

(9) *Pjotr Gavrivolič ānəmn jurt-ane jot ťit kassēta-γ ťēt-əs.*
P.G. I.LAT friend-PL.3SG with two cassette-DU send-PST.3SG
A = TOP [IOC + Subj. agreement]
'Pjotr Gavrivolich sent me two cassettes with his friends.'
(Dinislamova 2007: 5)

(10) *(tan) al-ne χul-anəl gosudarstwə-n miγ-anəl*
(they) kill-PTCP fish-3PL state-LAT give-SG.3PL
A + T = TOP [IOC + Obj. agreement]
'They give the fish they catch to the state.'
(Kálmán 1976: 136)

(11) *Nēnan am śopr-śonaχ-əl wāri-jaγəm.*
you(DU).ACC I silver-cup-INSTR do-DU.1SG
R + A = TOP [SOC]
'I'll make the two of you a silver cup.'
(Kálmán 1976: 70)

The following Mansi examples collected from a native speaker confirm the correlation between the information structure and the use of the different constructions (Sipőcz 2015). If T or R occurred as contrastive topics, the native speaker used the indirective construction in the case of T (12), and secundative construction in the case of R (13). Cf.:

(12) **T as contrastive topic:**

Wi-s-lum	*ńań*	*os*	*śakwit,*
buy-PST-SG.1SG	bread	and	milk

śakwit	*oma-m-n*	*mi-s-lum.*
Milk	mother-1SG-LAT	give-PST-SG.1SG

‘I bought bread and milk, I gave the milk to my mother.’

(data from informant)

(13) **R as contrastive topic:**

Uwśi-m	*tor-əl*	*mi-s-lum,*
sister-1SG	kerchief-INSTR	give-PST-SG.1SG

kaŋk-um	*sup-əl*	*mi-s-lum.*
brother-1SG	shirt- INSTR	give-PST-SG.1SG

‘I bought a kerchief for my (elder) sister and a shirt for my (elder) brother.’

(data from informant)

The following examples show the use of passivization in expressing the different information structural roles. The sentences (14) and (15) were recorded from a Mansi native speaker and they were uttered one after the other. The first sentence, (14), contains T-passivization, the word ‘dress’ is the topic – the dress was given to the informant as a present. In the next sentence, (15), the informant talks about herself as the recipient, someone who was given a present, so she uses R-passivization (Sipőcz 2015):

(14)	*Ti*	*mańśi sup*	*podruška-m-n*	*mujlupt-awe-s.*
	this	Mansi dress	girlfriend-1SG-LAT	present-PASS-PST.3SG
	T	**= TOP**	**[T-passivization]**	

‘This Mansi dress was given (to me) by a friend as a present.’

(data from informant)

(15)	*Tor-el*		*os*	*mujlupt-awe-s-əm.*
	kerchief-INSTR		also	present-PASS-PST-1SG
	(R)	**= TOP**		**[R-passivization]**

‘I was given a kerchief as well.’

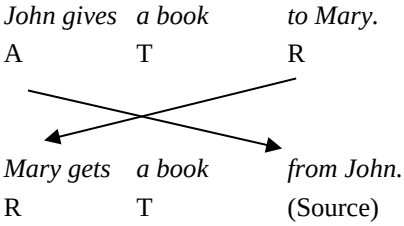
(data from informant)

3.1. Giving is getting

As has been already mentioned, the absence of the verb ‘get’ is an interesting feature of the Mansi lexicon. By this I mean the absence of the “basic” verb meaning ‘get’. Similarly to other languages Mansi also has several verbs for the concept of getting with different specialised meanings, like ‘obtain’, ‘receive’ etc. For instance, NM *patti* ‘take, obtain, get’, *wiɣ* ‘id.’ (Munkácsi and Kálmán 424, 725). “Modern” Russian-Mansi dictionaries contain the verb *wiɣ* ‘take’ as the Mansi equivalent of the Russian verb *получать* ‘get’. For instance, *wiŋkwe* / *wojiɣlaluŋkwe* ‘получить / получать’ (Rombandeeva, Rombandeeva and Kuzakova).[3] Dictionaries based on earlier collections do not mention this meaning of the verb *wiɣ*.

Crosslinguistically the verb ‘get’ belongs to the most frequent verbs[4] and, as a result, this verb – similarly to the verb ‘give’ – is a common basis for grammaticalization processes (Heine and Kuteva 143–149, 149–156).

The verbs ‘get’ and ‘give’ are considered a semantic pair (Primus 407). The events expressed by these verbs are represented by the same participants: the giver, the recipient and the given object. (As thematic roles, these are the Agent, the Recipient and the Theme.) In the case of the verb ‘give’ the giver is the subject of the verb, while in the case of the verb ‘get’ the recipient is the subject. Cf.:



In Mansi the ditansitive alternation is a device for differentiation between the events of giving and getting by the single verb ‘give’ putting the participants of the event into different grammatical roles.

[3] Probably, under the influence of Russian, modern Russian-Mansi dictionaries often list Mansi lexems the use of which is not typical or even questionable. (It can be seen also in the fact that these dictionaries often create the perfective/imperfective pairs for verbs, as we can see it in the example mentioned above. The perfective/imperfective opposition is characteristic of Russian verbs but not of Mansi.)

[4] According to the Hungarian National Corpus the verb *kap* ’get’ is among the 10 most frequently used Hungarian verbs. http://www.helyesiras.mta.hu/helyesiras/blog/show/tiz-leggyakoribb-tartalmas-szo-a-magyarban. (2016.08.03).

In the case of the active ditransitive constructions containing the verb *miγ* ‘give’, the event expressed by the verb is giving. The subject of the clause is the giver, who is at the same time the Agent of the event and usually the topic of the discourse. The event in which the giver is the Agent can be only the giving irrespectively of the fact whether the Theme or the Recepient is in the object position. Cf. 16–17:

(16) *Kāsəŋ xōtpa manasāwit wērm-əs, tasāwit oln ta mi-s.*
every person as.much able-PST.3SG same.much money PTCL give- PST.3SG
‘Everybody gave as much money as s/he could.’

(LS 2016/13: 13)

(17) *akw’ ēt ūnl-än-ən māγəs χūrəm-sat*
one night sit-AN-2SG for three-hundred

sajt-əl mīγ-ləm
ruble-INSTR give-SG.1SG
‘For watching for one night I give you 300 rubles.’

(VNGy IV: 334)

However, the passive constructions containing the verb ‘give’ allow the interpretation of both giving and getting. As has already been mentioned, there are two types of passive ditransitive constructions in Mansi, the T and R passivizations (**2.2**). In the case of T passivization the subject of the construction is the Theme argument and, regarding the connection of information structure and clause structure, the Theme is the topic of the discourse.[5] The event in which the Theme appears as the topical element can be either the giving or the getting. For instance, in Finnish both verbs can be used in the passive construction:[6]

[5] It is worth mentioning that T-passivization in more recent texts seems to have an emphasizing function, in addition to its topical use. Rather often this kind of passivization is used in order to place an extra emphasis on the Theme (Bíró and Sipőcz 2016). See for example:

Sverdlovski	*oblasť-it*	*mansi*	*mir-n*	*nemater*	*ńotmil*
Sverdlovsk	region-PL	Mansi	people-LAT	nothing	help
at	*majla-we.*				
NEG	give-PASS.3SG				

‘Since the Mansi people of the Sverdlovsk region are given no help at all.’

(Dinislamova 2007: 8)

[6] In Finnish only the T argument can be passivized.

(18) Finnish

Kirja	*anne-ttiin.*	/	*Kirja*	*saa-tiin.*
book	give-PASS.PST/		book	get-PASS.PST
‘The book was given.’		/	‘The book was received.’	

As far as the overt arguments of the structure are concerned, it would seem that if the A is present, the even should be interpreted as giving, whereas if the R is present, then as getting. A characteristic feature of Mansi passives is that the A is often present. The presence of the agent is, however, not typical of T-passive sentences, whereas the presence of the R is very typical (19). From this we could conclude that in T-passive sentences the main factor is getting. However, in my opinion it is unnecessary to separate the two meanings: the interpretation of the situation as getting or giving is dependent on the context and perspective.

(19)

vāt	*tal*	*kēr=tińśäŋ*	*naurəm-ən*	*maj-wə-s*
thirty	fathom	iron=tether	child-LAT	giv-PASS-3SG

‘The thirty fathoms long iron tether was given to the child.’
/‘The thirty fathoms long iron tether was received by the child.’
(VNGy II: 111)

In case of R-passivisation the subject of the construction (and, thus, the topic of the utterance) is the recipient. (Example 22 demonstrates well the topic nature of the R.) The event whose topic is R is getting and not giving. In other words, from the perspective of the R, the primary aspect of the event is getting. From this it follows that the verb ‘give’ in the secundative passive construction is used to express the notion ‘get’.[7] For instance:

(20)

kank-ä-n			*jäγ-ä*	*sēl-əm*		
elder.brother-3SG-LAT			father-3SG	gather-PTCP.PST		
puuŋ-nəl	*i*	*akw‘*	*ōln=pāl-əl*		*at*	*maj-wə-s.*
wealth-ABL	and	one	money=half-INSTR		NEG	give-PASS-PST[3SG]

‘He did not get even a half penny from his elder brother from the wealth gathered by his father.’ (VNGy IV: 326)

[7] Unlike in many other languages, traditionally in Mansi it is not characteristic that passive constructions are used when the agent of the verb is unknown or general. On the contrary, the Agent is often present in the clause (example 22). The use of the passive voice is motivated by the information structure. It is worth mentioning, however, that in more recent textes the use of the passive construction is not rare in the case of general agents.

(21) *Kit-it* *mesta-l* *Nižnewartowskij* *ūs-t* *ōl-ne*
two-DX place-INSTR Nizhnevartovsk town-LOC live-PTCP.PRS

xantə-t *maj-we-s-ət.*
Khanty-PL give-PASS-PST-3PL
‘Khanty people from Nizhnevartovsk got the second place.’
(LS 2016/13: 9)

(22) *Ruś-ət* *jornkol-t* *ōs* *ōl-s-ət,* *Raisa Iwanowna-n*
Russian-PL tent-LOC also to be-PST-3PL R.I.-LAT

tān *pussən* *ńāń-əl* *ōs* *maj-wē-s-ət.*
they all bread-INSTR also give-PASS-PST-3PL
‘There were Russians in the tent, too. All of them got bread from Raisa Ivanovna.’
(LS 2016/10. 9)

3.2. Taking as getting?

As has already been mentioned, modern Russian–Mansi dictionaries contain the verb *wiɣ* with the meaning ‘get’. The basic and most frequent meaning of this verb is ‘take, bring’ and earlier dictionaries do not mention this use of the verb, i.e. the meaning ’get’. It is noticeable, however, that in more recent texts we can see this kind of use of the verb *wiɣ*. Example (23) is from the same newspaper article as example (21), the sentences express the same situation, but the verbs are different.

(23) *Os* *xūrmit* *mesta* *ńeft’ejuganskij* *rajon-t*
and third place Nefteyugansk district-LOC

ōl-ne *xōtpa-t* *wi-s-ət.*
live-PTCP.PRS person-PL take-PST-3PL
‘And people from Nefteyugansk took/got the 3rd place.’
(LS 2016/13: 9)

I consider it important, however, that the verbs ‘get’ and ‘take’ differ regarding the thematic roles of their arguments. While the subject of the former verb is not an Agent but a Recipient, the subject of the latter verb is a typical Agent. This kind of alternation for expressing the situation of getting is common in other languages, too. Cf.: English *They got the first place.* (Subject = Recipient) / *They took the second place.* (Subject = Agent)

4. Summary

The absence of the verb meaning ‘get’ is an interesting lexicological feature of the Mansi language. The linguistic analysis of the event of getting makes it clear that the notion of getting is expressed by the verb ‘give’ used in the passive construction. The passivization of the verb ‘give’ belongs to the phenomenon of ditransitive alternation and passivization. In this paper I have argued that R-passivization of the verb ‘give’ is the main device to express the notion of getting.

Abbreviations

A	agent of a (di)transitive clause
ABL	ablative
ACC	accusative
AN	action nominal
DAT	dative
DOC	double object construction
DU	dual
INSTR	instrumental
IOC	indirect object construction
LAT	lative
LOC	locative
NEG	negative particle
NM	Northern Mansi
PASS	passive
PL	plural
PRS	present
PST	past
PTCL	particle
PTCP	participle
R	recipient
SG	singular
SOC	secundary object construction
T	theme
V	verb

References

Dalrymple, Mary and Nikolaeva, Irina. 2014. *Objects and information structure*. Cambridge University Press.

Dinislamova, Svetlana 2007. *Lavim ľoŋxanuw* [The way destined to us.] Khanty-Mansiysk: Poligrafist.

Haspelmath, Martin 2013. Ditransitive constructions: The verb ‘give’. In: Dryer, Matthew S. and Haspelmath, Martin (eds.) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
(Available online at http://wals.info/chapter/105, Accessed on 2016-08-22.)

Heine, Bernd and Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge University Press.

Kálmán, Béla 1976. *Wogulische Texte mit einem Glossar*. Budapest: Akadémiai Kiadó.

Malchukov, Andrej; Haspelmath, Martin and Comrie, Bernard 2010. Ditransitive constructions: a typological overview. In: Malchukov, Andrej; Haspelmath, Martin and Comrie, Bernard (eds.) *Studies in ditransitive constructions: A comparative Handbook.* Berlin: Mouton de Gruyter. 1–65.

LS = *Lujima Seripos*. Mansi newspaper. Khanty-Mansiysk.
(http://www.khanty-yasang.ru/)

Munkácsi, Bernát and Kálmán, Béla 1986. *Wogulisches Wörterbuch*. Budapest: Akadémiai Kiadó.

Primus, Beatrice 2015. Participant roles. In: Riemen, Nick (ed.) *The Routledge Handbook of Semantics*. London: Routledge. 403-418.

Rombandeeva, E. I. 2005. *Russko–mansiyskiy slovar*. Sankt-Peterburg: Mirall.

Rombandeeva, E. I. and Kuzakova, E. A. 1982. *Mansiysko–russkiy i russko–mansiyskiy slovar*. Leningrad: Prosveshenie.

Sipőcz, Katalin 2015. Ditransitivity in the Ob-Ugric languages. In: Harri, Mantila; Leinonen, Kaisa; Brunni, Sisko, Sivonen, Jari and Palviainen, Santeri (eds.) *Congressus Duodecimus Internationalis Fenno-Ugristarum.* Plenary Papers. Oulu: University of Oulu. 133–159.

Skribnik, Elena 2001. Pragmatic structuring in Northern Mansi. *CIFU* 9/6: 222–239.

VNGy = Munkácsi, Bernát 1892–1896. *Vogul népköltési gyűjtemény* I–IV. [Vogul folklore collection, vols. 1–4] Budapest: MTA.

The grammaticalization of Northern Mansi *mā* ‘earth, world, land, place’[1]

Bernadett Bíró
University of Szeged

1. Introduction

In this paper I discuss the grammaticalization of the word *mā* ‘earth, world, land, place’ into a nominalizer in Northern Mansi.

Mansi (or Vogul) is one of the most endangered languages of the Uralic (Finno-Ugric) language family. It is spoken by the river Ob and its tributaries in Western Siberia by less than 1,000 people. The only Mansi dialect that is still spoken today is Northern Mansi, and this dialect serves also as the basis of the Mansi literary language. The data used for this research are taken from written sources dated between the 1890’s and 2016.

2. The Northern Mansi *mā*

The word *mā* has several meanings in Northern Mansi: ‘earth, country, land, place, region, world, ground; part; field’, e.g.:

(1) *mā ēntəptanə mōjt* ‘tale of the girdling of the Earth’,
(2) *sēməl mā* ‘black soil’,
(3) *ūnlənə mā* ‘place of living’ (lit. ‘living place’),
(4) *ńāl mān ti pēlχati* ‘the arrow bores into the ground’,
(5) *χoti mā* ‘any region’ etc.

(cf. WW: 288–290)

[1] This research was funded by OTKA PD 116990 grant of the Hungarian Government. I also would like to thank Elena Skribnik for making her presentation and her papers available to me.

But in many cases it seems to have only a grammatical meaning, i.e. when combining with a participle it serves as a nominalizer[2], creating abstract nouns, e.g.:

(6) ***l'uś-nə*** ***mā-tä-nəl*** *pojt-s*
cry-PTCP.PRS NLZR-3SG-ABL stop/cease-PST[3SG]
's/he stopped crying'

(WW: 448)

There are several other nouns in Northern Mansi which can function also as nominalizers – similarly to *mā* –, and combining with adjectives or participles they can create concrete and abstract nouns. These nouns are the following: *äś* 'matter, thing, work', *ut* 'something, thing', *χar* 'something, thing, creature', *nak* 'joint, part, thing, place, space', *wārmal'* 'thing, work' (Riese 2001: 142–147). Cf.:

(7) *pəl* *wat-ne-äś*
berry pick-PTCP.PRS-NLZR
'berry-picking'

(8) *mas-n-ut*
dress-PTCP.PRS-NLZR
'clothes'

(9) *sāli* *janmalta-n* *wārmal'*
reindeer breed- PTCP.PRS NLZR
'reindeer-breeding'

Mā has not been mentioned in the literature as a nominalizer, although on the basis of both older and recent texts it seems to have this kind of function, too.

3. The grammaticalization of words meaning 'earth, land', 'area' and 'place'

Grammaticalization is the process when lexical forms develop into grammatical forms, and/or grammatical forms develop into even more grammatical ones (Heine and Kuteva 2002: 2). Grammaticalization consists of four interrelated steps:

(i) desemanticization (or "semantic bleaching") – the lexical form loses its meaning and semantic content gradually;

[2] Nominalizers are auxiliary nouns used for creating (concrete and abstract) nouns, and they are grammaticalized from participial relative clauses (cf. Skribnik 2008).

(ii) extension (or context generalization) – the given form starts to be used in new contexts;
(iii) decategorialization – the given form loses those morphosyntactic properties characteristic of lexical and other less grammaticalized forms;
(iv) erosion (or "phonetic reduction") – the given form loses its phonetic substance (Heine and Kuteva 2002: 2).

Concerning the sources of grammaticalization, the most frequent sources are lexical items with a considerably general meaning and also those items which occur frequently in the language. They are typically basic level terms (*back*, *hand*) or superordinate terms (*person*, *thing*). Body part terms, relational nouns and verbs meaning 'go, come, say, keep, take' typically tend to grammaticalize in most languages (Hopper and Traugott 1993: 41).

According to Heine and Kuteva (2002), the words *area* ('area', 'region'), *earth* ('earth', 'soil', 'land', 'ground') and place can often serve as a source of grammaticalization, too. It seems, however, that the result of the grammaticalization is usually not the same in other languages as the one found in Mansi. Both earth and place can commonly be grammaticalized into locative markers. earth can serve as a source of adverbs, prepositions or postpositions meaning 'below', 'under', 'down', 'beneath', e.g. Latvian *zeme* 'earth', 'ground' > *zem* 'under' (Heine and Kuteva 2002: 121–122). Place typically serves as the basis for prepositions or postpositions with the meaning 'at', 'toward' and 'to', e.g. Finnish *kohta* 'place' > *kohdalla* (kohta-ADESS) 'at' (postposition): *talon kohdalla* 'at the house' (Heine and Kuteva 2002: 240).

Although less commonly, but *area* 'area', 'region' can also be the source of locative markers, locative adverbials and postpositions meaning 'around', e.g. Imonda (Trans-New Guinea) *la* 'area' > 'around' (Heine and Kuteva 2002: 44).

In addition, place can often be the source of relational grammatical markers with the meaning 'instead of', and less frequently, the source of causal markers (conjunctions 'because' or 'therefore'). (For this latter case the examples involve one language family only.) E.g. Hungarian *hely* 'place' > *helyett* 'instead of' (postposition), Bambara (Niger-Congo) *yòrò* 'place' > *o yòrò kama* 'for this place' > *o yòrò kama* 'therefore' (conjunction). (Heine and Kuteva 2002: 239–240).

4. *Mā* as a nominalizer in Northern Mansi

The word *mā* as a nominalizer mostly combines with the present participle and creates action nominals (10) and – more rarely – result nouns (11–12).

(10) ***χōntlaχtə-nə*** ***mā-tä-nəl*** *ti* *pojt-əs.*
fight-PTCP.PRS NLZR-3SG-ABL PTCL stop-PST[3SG]
'S/He stopped fighting'

(VNGy II: 24)

(11) *manər-sir* ***pil-ne*** ***mā-n****?*
what-kind fear-PTCP.PRS NLZR-2SG

'What are you afraid of?'
(lit. 'what kind of fearing thing of yours' i.e. 'what kind of fear do you have?')

(Chernetsov Archives Nr. 42/10)

(12) *Tot* *aṅ* ***χōntl-ən*** ***mā-t***
there PTCL fight- PTCP.PRS NLZR-LOC

piγ-ēn *porsl-uw-es.*
son-3DU dirty-PASS-PST[3SG]

'There in the war their son dirtied.' [most probably a euphemism for 'died']

(LS. 2015/24: 12)

More rarely *mā* can also be combined with the past participle, cf. (13):

(13) *jaγ-ən* *opariś-ən* *ta* *untmit*
father-2SG grandfather-2SG PTCL sign[cut in the trees to show the way]

jal-um ***ma-te-t*** *sorumpat-s.*
walk.travel-PTCL.PST NLZR-3SG-LOC die-PST[3SG]

'The grandfather of your father died following that sign.'
(lit. 'in his walking that sign')

(Chernetsov Archaives, Nr. 44)

As has been mentioned before, *mā* as a nominalizer mostly creates action nominals. The two most frequent structures are the following:

a) present or past participle + *mā* + Px + LOC

(14) *naŋ* ***jäl-nə*** ***mā-n-t***
you travel-PTCP.PRS NLZR-2SG-LOC

matər	*ti*	*vār-s-ən!*
something	PTCL	do-PST-2SG

‘During your travelling you did something wrong!’

(VNGy I: 3)

This structure mostly expresses simultaneous action or event and the base verb of the participle is usually a motion verb (‘go’, ‘walk’, ‘travel’ etc.).

b) present or past participle + mā + Px + ABL + (jol)pojti ‘to stop, to cease’

(15)

am	***sāγra-ne***	***mā-m-nəl***
I	cut-PTCP.PRS	NLZR-1SG-ABL

jol-pōjt-ēγum,	*taw*	*χortal-i*
stop-1SG	(s)he	bark-3SG

‘I stop cutting [the tree with an axe], s/he [the dog] is barking.’

(Chrest. Vog.: 81)

In the more recent texts this construction appears typically without the possessive suffix:

(16)

Tuwəl	*tot*	*āγməŋ-əγ*	*jēmt-s-um,*
then	there	ill-TRANSL	become-PST-1SG

taji-māγəs	***χańištaχt-ən***	***mā-nəl***
therefore	study-PTCP.PRS	NLZR-ABL

jol-pojt-s-um,	*os*	*juw*	*ta*	*mina-s-um.*
stop-PST-1SG	and	home	PTCL	go-PST-1SG

‘Then I got ill there, therefore I gave up my studies and went home.’

(LS: 2015/24: 14)

If the finite verb of the sentence is *(jol)pojti* ‘to stop, to cease’, then almost always this construction is used. There is one example, though, where the lative case suffix is used instead of the ablative:

(17)

kantl-əm	***mā-tä-n***	*pojt-əs*
be.angry-PTCP.PST	NLZR-3SG-LAT	stop-PST[3SG]

‘S/He was not angry any more.’

(WW: 288)

In this construction there seem to be no restrictions concerning the base verb of the participle: motion verbs as well as any kind of verb can participate in it.

Examples found in my data show that mā can also serve as a nominalizer in Northern Mansi. It represents the third stage of the grammaticalization process, namely, decategorialization. *Mā* as a nominalizer behaves similarly to derivational suffixes, creating event and result nouns. It usually takes possessive suffixes and can also take case suffixes. The fact that in the given examples the case suffix and/or the possessive suffix is always attached to the element *mā* instead of the participle shows that this combination is treated as one unit. Participles can also function as action nominals independently (without any nominalizer element), there are hundreds of examples of this in Northern Mansi (cf. e.g. Bíró 2011, 2014). In this function, participles can combine with case suffixes, possessive suffixes (used for subject agreement, i.e. to refer to the subject of the base verb of the action nominal) and postpositions. If the participle/action nominal is combined with both a postposition and a possessive suffix then the latter is attached to the action nominal:

(18) *jūw* ***joχt-əm-ä*** ***jui-pālt*** *jol-χuj-əs.*
home come-AN-3SG after down-lie-PST[3SG]
'After s/he had come home, s/he lay down.'

(VNGy IV: 155)

Among the hundreds of examples there are only a few where the possessive suffix is attached to the postposition:

(19) *pīγkwə!* *am* *naŋən* ***rēχt-əm*** ***porä-m-t*** *uśt*
little.boy I you.ACC give.birth-AN time-1SG-LOC right.then

vorti *kit* *χapγä-lūpta* *kit* *pait-äγən* *ōl-s-eiγ;*
red two poplar-leaf two cheek-DU.2SG be-PST-3DU

'Little boy! When I gave birth to you, your cheeks were like two red poplar leaves.'

(VNGy I: 123)

This fact shows that the element *mā* as a nominalizer has gone further on the path of grammaticalization than the postpositions, and that it behaves like a derivational suffix.

5. The historical background of this process

The source of the grammaticalization process was most probably the meaning ‘place’. Presumably, the constructions containing similar expressions as *jaləm/jalnə mā* ‘walking/travelling place’ (i.e. ‘place for/of walking/travelling’) could give rise to the grammaticalization: ‘the place for/of walking/travelling’, that is ‘the place where somebody is/was walking/travelling’ can be easily interpreted as ‘while somebody is/was walking/travelling’ (i.e. ‘while somebody is/was away’). (The grammaticalization of spatial terms into temporal ones is a well-known process cross-linguistically – cf. Heine and Kuteva 2002: 6, among others.) Thus, in some of these examples the combination of the participle and the element mā allows not only the action nominal interpretation (‘during his travelling’) but also the “original”, lexical interpretation: ‘travelling place’ i.e. ‘the place where somebody is/was travelling’. See (14) again as (20):

(20) *naŋ* ***jäl-nə*** ***mā-n-t***
you travel-PTCP.PRS NLZR-2SG-LOC

matər *ti* *vār-s-ən!*
something PTCL do-PST-2SG

‘During your travelling you did something wrong!’

(VNGy I: 3)

Here the collector of the texts translated the participle + *mā* construction as an action nominal (cf. Hungarian “jártodban”, i.e. lit. ‘in your walking’) and there is no reason to question his competence although this sentence could also be translated as ‘You did something wrong at the place where you were travelling’ (‘at your travelling place’).[3]

Example (21) contains a quite similar expression: *tūjtχatəm mā* ‘hiding place’:

(21) *akwʻ* *mā-t* *toχ* ***tūjtχat-əm*** ***mā-m-t***
a place-LOC like.this hide-PTCP.PST NLZR-1SG-LOC

Lōpəχ-āγi-t *pūl-uŋkwə* *ti* *jōm-eγət.*

[3] It is noteworthy, however, that the use of the present participle instead of the past participle also supports the original translation (‘during your travelling’) since if the meaning ‘the place where you were travelling’ was intended, then rather the past participle (*jaləm*) would have been used.

Lōpəχ-girl-PL bathe-INF PTCL come-3PL

‘At a place as I am hiding like this, the girls from Lōpəχ come to bathe.’
(VNGy II: 186)

Here the expression ‘hiding place’ could also be interpreted literally (‘a place for/of hiding’, i.e. ‘to the place I’m hiding, the girls come to bathe’), however, here it is not only the original translation but also the presence of the adverb *toχ* ‘like this, so’ that contradicts this interpretation. Thus, in this sentence the item *mā* appears in two functions: at first as a lexical item meaning ‘place’ (*akw*’ *māt* ‘at a place’) and secondly as a grammatical item, as a nominalizer: *tūjtχatəm mā* ‘hiding’ (*tūjtχatəm māmt* ‘during my hiding’).

Thus, the grammaticalization of *mā* as a nominalizer (and probably even as a derivational suffix) supposedly has proceeded as follows:

(i) ‘the (concrete) place of the action’ (e.g. ‘travelling place’, noun) >
(ii) ‘time of the action’ (e.g. ‘during your travelling’ or ‘(while) travelling’, action nominal) >
(iii) ‘the name of the action’ (e.g. ‘travelling’, action nominal)/ ‘the result of the action’ (e.g. ‘trip’, result noun).

6. Similar grammaticalization processes in the same area

As has been mentioned before, the grammaticalization of words meaning ‘place’ as nominalizers does not seem to be common cross-linguistically, at least at first sight. After taking a closer look, however, we can see that very similar grammaticalization processes can be found in other Mansi dialects as well as in other languages of the Siberian and the neighbouring Mongolian area.

6.1. Eastern Mansi

The Eastern Mansi dialect was still spoken in the 1970’s along the river Konda, but it can be considered extinct today. A very similar grammaticalization process of the noun *mõõ* ‘earth, land, place’ (~ Northern Mansi *mā*) can be observed in this dialect (cf. Heikkonen 2013). The two most frequent structures containing *mõõ* are the following:

(i) action nominal(/past participle)[4] + *mõõ* + PX + LOC

This structure occurs almost only with the action nominal derived from the verb ‘to go’, e.g.

(22) *møn-nø-mõõ-m-t*
go-AN-NLZR-1SG-LOC
‘during my walking, as I walk(ed)’
(Heikkonen 2013: 15)

There are some examples also with other motion verbs, but in these *mõõ* is attached to the past participle instead of the action nominal, e.g.

(23) *jål-wojølp-ääm* *mõõ-tää-t*
down-fly-PTCP.PST NLZR-3SG-LOC
‘during his/her descending’
(Heikkonen 2013: 15)

Structures where *mõõ* is combined with the action nominal usually express simultaneous action while those containing the past participle generally express prior action. According to Heikkonen, this form has been grammaticalized and its function is to create adverbs (i.e. converbs) (Heikkonen 2013: 15).

This construction completely corresponds to the one found in Northern Mansi except that the non-finite verbal form appearing in the Northern Mansi construction is the (present or past) participle since there is no distinct form of the action nominal in Northern Mansi.[5] (Usually the participles are used as action nominals.) Heikkonen considers these Eastern Mansi forms (action nominal + *mõõ*) converbs (‘[while] travelling’) while I consider their Northern Mansi counterparts action nominals (‘during travelling’). Distinguishing between action nominals and converbs can be

[4] There are six non-finite verbal forms in Eastern Mansi (Kulonen refers to them as “nominaalimuodot”, i.e. “nominal verb forms”). They are the following: the infinitive, four participles (the present participle in *-p,* the past participle in *-m* and two other, more rarely used participles in *-i* and in *-s*) as well as the action nominal in *-n.* According to Kulonen, considering its function the action nominal is a verb form rather than a derived noun. (Kulonen 2007: 182–190).

[5] It is noteworthy, however, that the derivational suffixes appearing in these non-finite verbal forms are the same in both Mansi dialects: *-n* for the present participle and *-m* for the past participle in Northern Mansi, and *-n* for the action noimnal and *-m* for the past participle in Eastern Mansi.

problematic in some cases since converbs tend to originate – and in fact are continuously developing – from action nominals marked with a case suffix and used as adverbs (cf. Koptjevskaja-Tamm 1993: 44, Haspelmath 1995: 49, 1999: 114, Tikkanen 2001: 1121, among others). There are several non-finite verbal forms in many Uralic languages which historically constitute a transition between the transparent forms of action nominals marked by a case suffix and the completely opaque converbs or infinitives (Ylikoski 2003). In separating one from the other we can rely on the fact that "case inflection of action nominals is a living process and reflects their different syntactic and semantic uses" while "the cases of prototypical converbs are fossilized and are interpreted rather as a part of the whole converb marker" (Koptjevskaja-Tamm 2003: 44). Considering this, the Northern Mansi constructions (participle + *mā* + PX + LOC) can be regarded rather as action nominals for the following reasons:

(a) They are completely transparent.
(b) The possessive suffix – used for subject agreement, thus, consequently able to appear in different numbers and persons – precedes the case suffix, as it does in the case of non-derived nouns as well. This fact shows that it is the participle + *mā* construction is considered as one unit, a noun (i.e. an action nominal) and not that *mā* + PX + LOC is considered as a fossilized converb marker.
(c) (ii) action nominal(/past participle) + *mõõ* + ABL

Unlike in Northern Mansi, there is no connection between the use of this construction and the finite verb of the sentence. In Eastern Mansi the use of the ablative form of *mõõ* is not triggered by the finite verb *påns-* 'to stop, to cease' at all (Heikkonen 2013: 17). In these Eastern Mansi constructions the base verb of the action nominal (or the past participle) can be not only motion verbs but also other kinds of verbs, e.g.:

(24) *nee-tø* *roåwlaxt-øs* ***koj-øm-mõõ-tää-nøl***
woman-3SG wake.up-PST[3SG] lie-PTCP.PST-NLZR-3SG-ABL
'The woman woke up from her dreams.' [lit. 'from her lying']
(Heikkonen 2013: 17)

Although *mõõ* has been translated traditionally as 'place' in these examples, Heikkonen argues that on the basis of the context these forms could – and in some cases indeed should – be translated as action nominals or converbs (Heikkonen 2013: 17). Heikkonen claims that this form originally had the meaning 'the place of

the action' but then it grammaticalized as an action nominal meaning the action itself. The local case suffixes of the action nominal (especially the locative) then grammaticalized further into the function of the converb. According to Heikkonen, it is noteworthy that while *mõõ*-converbs containing the locative suffix appear in several persons, those containing the ablative occur in the 3rd persons only[6] (Heikkonen 2013: 18–19).

6.2. Surgut Khanty

In the Surgut dialect of Khanty (or Ostyak) – the language most closely related to Mansi and also a geographically neighbouring language – a similar use of the word meaning 'place' can be found. The word TAHI (*tăγi ~ taγi ~ tăχə ~ tăχi ~ tăχa*) 'place' combined with participles tends to be grammaticalized and cause the nominalization of the construction. TAHI can create nouns expressing the place, time, result and name of the action (action nominals) as well as other abstract nouns (cf. Csepregi 2008), e.g.

(25) *wŏʌ-tə* *tåγi*
be-PTCP.PRS NLZR
'life, living'

(Csepregi 2008: 129)

(26) *năm* *pŏn-tə* *taγi*
name put-PTCP.PRS NLZR
'giving a name'

(Csepregi 2008: 129)

(27) *əjnam* *tŏŋəmtə-tə* *taγi* *tŏj-əʌ*
every(thing) understand-PTCP.PRS NLZR have-3SG
'everything makes sense'[7]

(Csepregi 2008: 129)

These structures appear only in the Eastern Khanty dialects and Csepregi considers them to be a relatively new phenomenon. According to her, the broad

[6] As can be seen from (15) and (16), for example, this is not the case in Northern Mansi. Although the majority of the Northern Mansi examples containing *mā* + ablative appear also in the 3rd persons, there are examples in other persons as well.

[7] I would like to thank Márta Csepregi for her help in analyzing the Khanty sentence.

semantic structure of the word ‘place’ can cause its grammaticalization as a nominalizer (or even as a derivational suffix) (Csepregi 2008: 132).

6.3. Siberian Turkic languages

This kind of nominalizing technique is also very frequent in other, non-Uralic, languages of the Siberian area, i.e. in Siberian Turkic languages as well as in Mongolic Buryat. The most usual nominalizers are nouns meaning ‘man, person’, ‘thing’, ‘place’, ‘event, business’ (cf. Skribnik 2008, 2010: 569–570). According to Skribnik (2010: 571), there are four nominalizing techniques in the languages of Western and Central Siberia:

(i) using non-finite verbal forms,
(ii) using nominalizers,
(iii) using nominalizing suffixes with other verbal forms,
(iv) using the combination of question and demonstrative pronouns.

She states that the use of nominalizers is a technique predominant in the Ob-Ugric languages (Mansi and Khanty) as well as in Selkup (a Southern Samoyedic, Uralic language) while it is quite rare in the Northern Samoyedic languages. In Siberian Turkic languages and in Mongolic Buryat, however, it is one of the two most frequent nominalizing techniques (Skribnik 2010: 571–572). South Siberian Turkic languages, for example, use the following nominalizers:

kiži ‘man’,
čer ‘place’,
kerek ‘thing-to-do, business’,
and ‘things’ of pronominal origin:
Altai-kiži *neme* ‘thing’ < *neme* ‘what’,
Tuvan *čüve* ‘thing’ < *čüü* ‘what’,
Khakas *nime* ‘thing’ < *nime* ‘what’ (Skribnik 2014: 263).

Thus, a nominalizer with the meaning ‘place’ can also be found in the Siberian Turkic languages. It seems, however, that in these languages the nominalizer ‘place’ is not used for action nominalization, but rather for creating locative nouns (expressing the place of the action), e.g. Tofan (Sayan Turkic) *emned=ir čer* (Ort zum Heilen) ‘Krankenhouse’, *ńemnen=ir čer* (Ort zum Essen) ‘Kantine’ (Skribnik 2010: 580) and also (28):

Čalqandu (Altay)

(28) *Qïs* *par=ɣan* *t'er=de* *t'at=tan*
girl give=PRT place=LOC stay=PRT.US
'A girl must live where she was married (given) into'
(example and glossing Skribnik 2008)

Using nominalizers is a special feature of this area, other Turkic languages do not apply them (Skribnik 2014: 263). Skribnik mentions that in Mansi, Khanty and Selkup these nominalizers often develop into derivational suffixes, e.g. Mansi *tēnut* 'food' < *tē-ne ut* 'eating thing'; Selkup *apsodimį* 'food' < *ap-sodi mį* 'thing to eat'. This phenomenon can also be observed in South Siberian Turkic languages with the Uralic substrate, e.g. Tofan *tïn-ar čüme* 'air' < 'thing to breathe' (Skribnik 2014: 268–269).[8]

7. Conclusions

The word *mā* 'earth, land, place' has been grammaticalized as a nominalizer in Northern Mansi. It has undergone the third stage of the grammaticalization process, i.e. decategorialization. It behaves similarly to derivational suffixes, combining with participles it creates action nominals and – more rarely – result nouns. *Mā* as a nominalizer usually takes possessive suffixes (for subject agreement, although in the newer texts this is less typical) and it also can take case suffixes (usually the locative and the ablative suffix). It is a productive nominalizer, it appears both in older and newer texts, although it is not a very frequent nominalizer. The reason for this is undoubtedly the fact that there are other, more common nominalizers (cf. 2) as well as that in most cases participles – without any nominalizing element – are used as action nominals (cf. 4).

A quite similar grammaticalization process of the word meaning 'place' into a nominalizer can be observed in other languages of the Siberian (Surgut Khanty, Siberian Turkic languages) and the neighbouring Mongolic area (Buryat). The identification of the possible areal influences, however, requires further investigations.

[8] Whereas in Mongolic languages as well as South Siberian Turkic languages in contact with Mongolic (e.g. Shor, Khakas, Tuvan) these nominalizer constructions "are used as predicate nominals for purposes of focussing (the scheme 'I did it' > 'I am the person who did it'), which leads to grammaticalization of their NRs [nominalizers] as assertive particles" (Skribnik 2008).

Abbreviations

ABL	ablative
ACC	accusative
AN	action nominal
DU	dual
INF	infinitive
LAT	lative
LOC	locative
NLZR	nominalizer
PASS	passive
PL	plural
PST	past
PTCL	particle
PTCP	participle
PTCP.PST	past participle
PTCP.PRS	present participle
PX	possessive suffix
SG	singular
TRANSL	translative
US	habitual

References

Bíró, Bernadett 2011. Cselekvésnevek a manysiban – Az északi manysi cselekvésnevek valenciája. [Action nominals in Mansi: The valency of Northern Mansi action nominals.] *Folia Uralica Debreceniensia* 18: 3–33.

Bíró, Bernadett 2014. *Cselekvésneves szerkezetek a manysiban*. [Action nominal constructions in Mansi]. Szeged: University of Szeged PhD-dissertation. http://doktori.bibl.u-szeged.hu/2304/

Csepregi, Márta 2008. Az osztják TAHI ‘hely’ szó helye a mondatban [The place of the Ostyak word TAHI ‘place’ in the sentence]. In: Bereczki, András; Csepregi, Márta and Klima, László (eds.) *Ünnepi írások Havas Ferenc tiszteletére. Urálisztikai Tanulmányok* 18. 125–135.

Chernetsov Archives (hand-written texts collected by V. Chernetsov in 1933–1934.) Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest.

Chrest.Vog.: Kálmán Béla 1989. *Chrestomathia Vogulica* (3rd edition). Budapest: Tankönyvkiadó.

Haspelmath, Martin 1995. The converb as a cross-linguistically valid category. In: Haspelmath, Martin and König, Ekkehard (eds.) *Converbs in cross-linguistic perspective: Structure and meaning of adverbial verb forms – adverbial participles, gerunds. Empirical Approaches to Language Typology* 13. Berlin and New York: Mouton de Gruyter. 1–56.

Heikkonen, Petri-Tapio 2013. *Itämansin temporaaliset infinittirakenteet* [Temporal non-finite constructions in Eastern Mansi]. Helsinki: University of Helsinki MA thesis.
https://www.academia.edu/15358520/It%C3%A4mansin_temporaaliset_infinittirakenteet_Temporal_non-finite_constructions_in_Eastern_Mansi_Masters_thesis_

Heine, Bernd and Tania Kuteva 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Hopper, Paul J. and Elizabeth Closs Traugott 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Koptjevskaja-Tamm, Maria 1993. *Nominalizations*. London and New-York: Routledge.

Kulonen, Ulla-Maija 2007. *Itämansin kielioppi ja tekstejä* [Eastern Mansi grammar and texts]. Apunevoja suomalaisugrilaisten kielten opintoja varten XV; Helsinki: Suomalais-Ugrilainen Seura.

LS: *Luima Seripos*. Mansi language newspaper. Khanty-Mansiysk. (http://www.khanty-yasang.ru/)

Riese, Timothy 2001. *Historische Nominalderivation des Wogulischen*. (Studia Uralica 10) Wiesbaden: Harrasowitz Verlag.

Skribnik, Elena 2008. South Siberian Turkic constructions with nominalizers and their further grammaticalization. (poster presentation) The 12th Biennial Rice University Symposium on Language. The Genesis of Syntactic Complexity: An Interdisciplinary Symposium. Poster Session. Houston, March 28, 2008.

Skribnik, Elena 2010. *Die Nominalisatoren 'Mensch' und 'Ding' in den uralischen und türkischen Sprachen Südsibiriens: I, Lexikon und Wortbildung*. Finnisch-Ugrische Mitteilungen 32/33. Hamburg: Helmut Buske Verlag. 567–586.

Skribnik, Elena 2014. South Siberian Turkic languages in linguistic contact: Altay-kiži nominalizer constructions as a test case. In: Besters-Dilger, Juliane; Dermarkar, Cynthia; Pfänder, Stefan and Rabus, Achim (eds.) *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity.* Berlin/Boston: De Gruyter. 258–274.

Tikkanen, Bertil 2001. Converbs. In: Haspelmath, Martin; König, Ekkehard; Oesterreicher, Wulf and Raible, Wolfgang (eds.) *Language typology and language universals: An international handbook.* Volume 2. Handbücher zur Sprach- und Kommunikationswissenschaft 20.2. Berlin: Walter de Gruyter, 1112–1123.

VNGY: Munkácsi Bernát 1892–1896. *Vogul népköltési gyűjtemény* I–IV [A collection of Vogul folklore texts]. Budapest: MTA.

Wog.Volk.: Kannisto, Artturi; Liimola, Matti and Eiras, Vuokko 1982. *Wogulische Volksdichtung. Wörterverzeichnis zu den Bänden* I-VI. Helsinki: Suomalais-ugrilainen Seura. MSFOu 180.

WW: Munkácsi, Bernát and Kálmán, Béla 1986. *Wogulisches Wörterbuch*. Budapest: Akadémiai Kiadó.

Ylikoski, Jussi 2003. Defining non-finites: Action nominals, converbs and infinitives. SKY *Journal of Linguistics* 16: 185–237.

Compounding in Aral–Caspian Kipchak languages

Bence Grezsa

University of Szeged

Research on compounding as an instrument of word formation is a rather new field in Turcology. This type of word formation might be used in various situations, for instance, it can perform the function of reduplication or suffixation. Therefore, compounding should be analysed from the aspect of structural, semantic and syntactic characteristics in Turkic languages. The present study provides an overview of compounding in the Aral–Caspian Kipchak languages applying the latest approaches of linguistics. The corpus of data used is collected from various dictionaries and grammars, including written materials, mirroring spoken languages.

1. Introduction

The aim of this study is to provide classification possibilities of the compounds in the Aral–Caspian Kipchak languages including Kazakh, Kirghiz, Karakalpak and Noghay. Another goal is to overview the characteristics of the compounds in these languages and to discuss some controversial questions on the subject, principally focusing on the structural and the semantic aspects. The topic of this study is specifically relevant. Although many works deal with compounding from the aspect of general linguistics, the number of the papers about compounding in Turkic languages is very small except for isolated examples. Consequently, I intend to provide a classification of compounds in the above mentioned Kipchak languages applying some of the latest methods of linguistics. I have chosen as the model of theoretical background the classification of the Morbo/Comp[1] project based on research by Bisetto and Scalise. The data in this article was collected from various dictionaries of these languages (see in References) and online written sources. My

[1] The Project of the Department of Foreign Languages in Bologna, Italy, with a data base of compounds, which aims to classify compounds primarily in Indo–European languages. For further information, see http://www.morbocomp.sslmit.unibo.it/

whole material contains approximately 500 items, but because of the space limitations of this article, I provide only some illustrations.

2. Research on compounds

The issue of compounds in general linguistics, as mentioned above, is a well-researched subject: numerous articles can be found about the role of compounds in word formation and a number of systems were suggested for their classification. I will mention only those sources which are relevant in this case. The first undertaking to classify compounds is Bloomfield (1933), who divided them into two groups, namely, exocentric and endocentric compounds and distinguished two more subcategories in the endocentric group, subordinative and coordinative compounds.[2] Spencer (1991) does not distinguish subgroups in his classification, but he has three categories, subordinative, coordinative and appositional compounds, disregarding the exocentric and endocentric classification. Fabb has categorized compounds according to headedness in three groups: those with no head, one head and two heads. Haspelmath (2002: 85−98) has created in his classification the new category of affix compounds and has differentiated five groups, meanwhile Bauer (2001: 695-707) and Booij (2005: 75−95) differentiate four in their works. The main groups of these three categorizations are exocentric, endocentric, coordinative, possessive and appositional compounds (Bisetto and Scalise 2005: 321–325).

A problem of these classifications is that the categories overlap. Moreover, some compounds cannot be classified, because not every attribute of compounds has been taken into consideration. However, the proposal by Bisetto and Scalise (2005: 326–330) attempt an unambiguous, clear and simple classification on the basis of syntactic constructions. According to them, every compound has either exocentric or endocentric attributes marking the presence or the absence of the head, providing the base of their classification. Besides exocentricity/endocentricity, compounds might be classified into three groups, namely, subordinate, attributive and coordinate compounds (Bisetto and Scalise 2005: 321–328). This grouping classifies compounds logically and reasonably. Because of this consideration, I have chosen to apply this method in the case of the Aral–Caspian Kipchak branch of Turkic.

[2] For the definition of the concepts, see Bisetto and Scalise (2005).

3. Compounds in Turkic languages

In Turcology, research on word formation is confined mostly to suffixes, with sourceson Turkic compounds limited to only one.

Research covers every aspect of compounds in Turkish: their structure, semantics, headeness and stress. Dede's research is the first to be emphasized, who analysed in her 1978 dissertation the semantic and syntactic properties of Turkish nominal compounds. Göksel discussed in a number of studies the compounding system of Turkish, first categorizing them on the basis of the findings of the MorboComp Project. Bağrıaçık and Ralli (2014) describe nominal–nominal concatenations in Turkish compounding. As for the Kipchak languages, few studies discuss the topic of compounding and are of relevance for the present paper. Krejci and Glass (2015) partially touch on compounds in Kazakh in their paper about the adjective/noun distinction in Kazakh. Van Hofwegen (2014) focuses on nominal compounds in the Kazakh language, discussing the remarks of Göksel and Haznedar (2007) regarding Turkish compounds, which will be also referred to below regarding the classification possibilities of the Aral–Caspian languages and their characteristics. The studies concentrate generally on nominal compounding in Turkic languages and leave other categories like verbal, adverbial and pronominal compounds out of consideration, even though such constructions can be found in the Aral–Caspian branch. Besides, it would be important to examine the influence of other non-Turkic languages, which probably had considerable effect on compounding as well. In our case, two languages must be taken in account, Russian and Persian.

4. Compounds in Kipchak languages

Similarly to all agglitunative languages, Kipchak languages have many bound morphemes with many allomorphic variants. Words are formed usually through suffixation and there is a lot of variation for creating new words and creating notions (Johanson 1998: 34–38). Thus, word formation in Kipchak languages is a very productive process, but not the only one. Generally, in nominal word formation there are numerous compounds. The compounds are often formed by joining two nouns, like the Kazakh *kün žarïġï* 'sunlight' (*kün* 'day, sun' + *žarïq* 'light' *+(s)I* possessive suffix) or Noghay *yïl šaġï* 'season' (*yïl* 'year' + *šaq* 'time, period, age' *+(s)I* possessive suffix). The second element of these nominal constructions contains generally the third person possessive suffix, which is *+(s)I* in every Aral–Caspian

Kipchak language. For this suffix, I use henceforth Göksel's term linking element. However, the linking element cannot be found in every case in the same type of nominal construction. This issue will be discussed below. Another sizeable group of compounds is composed of adjective plus noun or adjective plus adjective elements.[3] Furthermore, verbal compounds (Noun + Verb or Verb + Verb) are also found in this Kipchak group, as an outstanding part of the system. Tables 1, 2, 3, and 4 summarize the compounds of the Aral–Caspian languages taking in account endo– and exocentricity and Bisetto and Scalise's classification:

Subordinate		**Attributive**		**Coordinate**	
endocentric	**exocentric**	**endocentric**	**exocentric**	**endocentric**	**exocentric**
azïq–tülik dükeni 'grocery' < *azïq–tülik* 'food, feeding' + *düken* 'shop, store' +*(s)I* {Poss.Sg3};	*kelisimsöz* 'contract' < *kelisim* 'agreement, accord' + *söz* 'word';	*aġaš üy* 'timber house' < *aġaš* 'tree, wood, timber' + *üy* 'house, home, building';	*aq qandïlïq* 'leukemia' < *aq* 'white' + *qan* 'blood' +*DI* {NN/Adj.} +*LIK* {NN};	*meken–žay* 'residency' < *meken* 'place' + *žay* 'residence, accommodation';	*aldaqašan* 'long ago' < *alda* 'before, forth, ahead' + *qašan* 'when, as';
kitap söresi 'bookshelf'< *kitap* 'book' + *sore* 'shelf' +*(s)I* {Poss.Sg3};	*qolšatïr* 'umbrella' < *qol* 'arm, hand' + *šatïr* 'tent, roof';	*tüski tamaq* 'lunch' < *tüski* 'meridion, meridional' + *tamaq* 'food, throat'	*aqqaynar* 'champagne' < *aq* 'white' + *qayna–* 'boil, blaze' *–(A)r* {CV};	*it–qus* 'predators' < *it* 'dog' + *qus* 'bird';	*alïp–satar* 'tradesman'< *al–* 'to buy, take' *–Ip* {CV} + *sat-* 'to sell' *-(A)r* {CV};
tuwġan žeri 'birthplace' < *tuw–* 'to be born' *–GAn+* {CV} + *žer* 'place' +*(s)I* {Poss.Sg3};	*šaŋsorġïš* 'vacuum cleaner' < *šaŋ* 'dust' + *sor–* 'to absorb, suck' *-GIš* {CV};	*laqap at* 'nickname' < *laqap* 'alias, shortcut' + *at* 'name';	*alqïzïl* 'purple' < *al* 'bright tone' + *qïzïl* 'red';	*dos–dušpan* 'everybody' < *dos* 'friend' + *dušpan* 'enemy';	*ämir–qudiret* 'power' < *ämir* 'command, order' + *qudiret* 'strength, power'.

Table 1. Compounds in Kazakh

[3] According to Krejci and Glass (2015: 1−12), the noun/adjective distinction in Kazakh is not clear, but in compounding parts of speech like nouns and adjectives play an impontant role in compounding. Because of this I analyse them separately.

Subordinate		Attributive		Coordinate	
endocentric	**exocentric**	**endocentric**	**exocentric**	**endocentric**	**exocentric**
abet ubagï ‘lunchtime’ < *abet* ‘lunch’ + *ubaq* ‘time’ *+(s)I* {Poss.Sg3};	*čaŋ sorguč* ‘wacuum cleaner’ < *čaŋ* ‘dust’ + *sor-* ‘to absorb, suck’ *–GXč* {CV};	*deŋiz baš* ‘conceited’ < *deŋiz* ‘sea, lake’ + *baš* ‘head’;	*altïn kemer* ‘gold belt’ < *altïn* ‘gold’ + *kemer* ‘belt’;	*adïr–čïbïr* ‘hills, hilly region’ < *adïr* ‘hill’ + *čïbïr* ‘mountain range’;	*bāz–bāz* ‘sometimes’ < *bāz* ‘some, a little’;
at ǰalï ‘horse mane’ < *at* ‘horse’ + *ǰal* ‘mane’ *+(s)X* {Poss.Sg3};	*ǰer ǰüzü* ‘surface’ < *ǰer* ‘place’ + *ǰüz* ‘face’ *+(s)X* {Poss.Sg3};	*aq/tunuq sū* ‘vodka’ < *aq* ‘white’/ *tunuq* ‘transparent’ + *sū* ‘water’;	*ker sarï* ‘pale–face, white face’ < *ker* ‘brown, chestnut’ + *sari* ‘yellow’;	*ališ–beriš* ‘shopping’ < *al–* ‘to buy, take’ *–Iš* {VN} + *ber-* ‘to give’ *–Iš* {VN};	*boz–boz* ‘dun’ < *boz* ‘grey’;
ata meken ‘homeland’ < *ata* ‘father, dad’ + *meken* ‘place’;	*tiš ǰūġuč* ‘toothbrush’ < *tiš* ‘tooth’ + *ǰū–* ‘to wash’ *–GXč* {CV};	*čay qašïq* ‘teaspoon’ < *čay* ‘tea’ + *qašïq* ‘spoon’;	*qara altïn* ‘rock–oil’ < *qara* ‘black’ + *altïn* ‘gold’;	*bātïr qïz* ‘heroine’ < *bātïr* ‘hero’ + *qïz* ‘girl’;	*aqe–üke* ‘brothers’ < *aqe* ‘father, brother’ + *üke* ‘younger brother’.

Table 2. Compounds in Kirghiz

Subordinate		Attributive		Coordinate	
endocentric	**exocentric**	**endocentric**	**exocentric**	**endocentric**	**exocentric**
ata qonïs ‘homeland’ < *ata* ‘father’ + *qonïs* ‘stop, station’;	*bawïr et* ‘diaphragm’ < *bawïr* ‘liver’ + *et* ‘meat’;	*aqïlï az* ‘stupid, fool’ < *aqïl* ‘mind, logic’ *+(s)I* {Poss.Sg3} + *az* ‘few, little’;	*qara kök* ‘dark blue’ < *qara* ‘black’ + *kök* ‘blue’;	*aġa–ini* ‘brothers’ < *aġa* ‘brother, elder brother’ + *ini* ‘younger brother’;	*bolar–bolmas* ‘hardly’ < *bol–* ‘to be’ *–Ar* {Aor.Sg3} + *bol–* ‘to be’ *–mAs* {Neg.Aor.Sg 3};
köz žasï ‘tear’ < *köz* ‘eye’ + *žas*	*orïnbasar* ‘vicarious, supply’ <	*eki qabat* ‘pregnant’ < *eki* ‘two’ +	*qara may* ‘lubricant’ < *qara* ‘black’	*aġayïn–tuwġan* ‘relatives’ <	*demalïs* ‘rest’ < *dem* ‘respiration’

‘tear’ *+(s)I* {Poss.Sg3};	*orïn* ‘place, ground’ + *bas–* ‘to press, push’ *–(A)r* {CV};	*qabat* ‘layer, floor’;	*+ may* ‘grease, fat’;	*aġayïn* ‘relative’ + *tuw–* ‘to be born’ *–GAn* {CV};	*+ al–* ‘to buy, take’ *– Is* {CV};
teŋiz žaġasï ‘seaside’ < *teŋiz* ‘sea, lake’ + *žaġa* ‘side’ *+(s)I* {Poss.Sg3};	*qïliš balïq* ‘swordfish’ < *qïliš* ‘sword’ + *balïq* ‘fish’;	*er žürek* ‘brave’ < *er* ‘valiant, man’ + *žürek* ‘heart’;	*qurġaq žer* ‘mainland’ < *qurġaq* ‘dry’ + *žer* ‘place, ground’;	*kelim–ketim* ‘guests’ < *kel-* ‘to come’ *–Im* {VN} + *ket–* ‘to go’ *–Im* {VN};	*bala–šaġa* ‘family’ < *bala* ‘child’ + *šaġa* ‘group, relative’.

Table 3. Compounds in Karakalpak

Subordinate		**Attributive**		**Coordinate**	
endocentric	**exocentric**	**endocentric**	**exocentric**	**endocentric**	**exocentric**
ana tili ‘mother tounge’ < *ana* ‘mother’ + *til* ‘language’ *+(s)I* {Poss.Sg3};	*balta sap* ‘axe haft’ < *balta* ‘axe’ + *sap* ‘stem, shaft’;	*aq köŋili* ‘honest’ < *aq* ‘white’ + *köŋil* ‘mood, mind’ *+DI* {NN/Adj.};	*aq kök* ‘light blue’ < *aq* ‘white’ + *kök* ‘blue’;	*aġalï–inili* ‘brothers’ < *aġa* ‘brother, elder brother’ *+DI* {NN/Adj.} + *ini* ‘younger brother’ *+DI* {NN/Adj.};	*baqa–šanaq* ‘small shell’ < *baqa* ‘frog’ + *šanaq* ‘cup, bowl’;
at azbarï ‘stable’ < *at* ‘horse’ + *azbar* ‘yard’ *+(s)I* {Poss.Sg3};	*awïz ašuw* ‘Iftar, evening meal in fasting’ < *awïz* ‘mouth’ + *aš–* ‘to open’ *–uw* {VN};	*beti qalïn* ‘shameless’ < *bet* ‘face’ *+(s)I* {Poss.Sg3} + *qalïn* ‘thick, fat’;	*bos söz* ‘silliness’ < *bos* ‘empty’ + *söz* ‘word’;	*bolsa bolar* ‘maybe, possible’ < *bol–* ‘to be’ *– sA* {Cond.Sg3} + *bol–* ‘to be’ *– Ar* {Aor.Sg3};	*ata–ana* ‘parents’ < *ata* ‘father’ + *ana* ‘mother’;
ay yarïġï ‘moonlight’ < *ay* ‘moon’ + *yarïq* ‘light’ *+(s)I* {Poss.Sg3};	*tün ortasï* ‘midnight’ < *tün* ‘night’ + *orta* ‘middle’ *+(s)I* {Poss.Sg3};	*azarlï awïz* ‘foulmouthed’ < *azar* ‘quarrel, jaw’ *+DI* {NN} + *awïz* ‘mouth’;	*kiyiz etik* ‘felt boots’ < *kiyiz* ‘felt’ + *etik* ‘boots’;	*at–mat* ‘horsekinds’ < *at* ‘horse’ + *m-* {Red} + *at* ‘horse’;	*ömir–ömirge* ‘forever’ < *ömir* ‘life’ + *ömir* ‘life’ *+GA* {Dat}.

Table 4. Compounds in Noghay

5. Observations on compounding in the Aral–Caspian Kipchak branch

According to the structural characteristics, compounds can be formed in three ways: from two constituents with the third person possessive marker *+(s)I* suffix,[4] from two bare constituents, and through reduplication. In this section the most cardinal question is in which case the linking element occurs in compounds. It might undoubtedly appear only in subordinate and attributive structures, mainly in the endocentric subgroups. However, taking a closer look, we find more regularities in the structures with linking elements. In the case of Kazakh, I accept partly van Hofwegen's argument (2014: 1−21), that the presence or the lack of the linking element depends on the characteristic of the non-head noun in noun–noun compounds. If the non-head position contains a noun which can fulfilan adjectival function (and is, thus "neutral") as well, the linking element is used, unlike in "non-neutral" nouns, where no linking element is ever used. In my opinion, only the second statement is completely right in the Kazakh language, because the usage of the linking element in the "neutral" noun–noun constructions is optional, as Table 5 demonstrates.

'neutral' noun–noun constructions	**'non-neutral' noun–noun constructions**
balara ~ balarasï 'bee' < *bal* 'honey' + *ara* 'fly, bee' *+(s)I* {Poss.Sg3};	*qolžazba* 'handwriting' < *qol* 'arm, hand' + *žaz–* 'to write' *–MA* {VN};
tuwġan kün ~ tuwġan küni 'birthday' < *tuw–* 'to be born' *–GAn* {CV} + *kün* 'day, sun' *+(s)I* {Poss.Sg3};	*külsawït* 'ashtray' < *kül* 'ash' + *sawït* 'vessel, vase, jar';
tuwġan žeri ~ tuwġan žer 'birthplace' < *tuw–* 'to be born' *–GAn* {CV} + *žer* 'place' *+(s)I* {Poss.Sg3};	*äwežay* 'airport' < *äwe* 'air, sky' + *žay* 'residence, accomodation';

Table 5. Kazakh compounds

In Kirghiz, Karakalpak and Noghay there is no exact rule for the usage of the linking element. In my opinion, all of the subordinate and attributive endocentric

[4] Hereinafter I adopt the concept Linking Element following Göksel and Haznder's proposal (2007).

constructions oirginally disposed of the linking element, and its optional usage or disappearence is a new tendency in these languages (See Table 6).

	Kirghiz	**Karakalpak**	**Noghay**
‘shoes’	*but kiyim* < *but* ‘shoes’ + *kiyim* ‘dress’;	*ayaq kiyim* < *ayaq* ‘leg, foot’ + *kiyim* ‘dress’;	*ayaq kiyimi* < *ayaq* ‘leg, foot’ + *kiyim* ‘dress’;
‘railway’	*temir jol* < *temir* ‘iron’ + *jol* ‘road, way, path’;	*temir žolï* < *temir* ‘iron’ + *žol* ‘road, way, path’;	*temir yol* < *temir* ‘iron’ + *yol* ‘road, way, path’;
‘apple tree’	*alma jïğačï* < *alma* ‘apple’ + *jïğač* ‘tree’ *+(s)X* {Poss.Sg3};	*alma aġašï* < *alma* ‘apple’ + *aġaš* ‘tree’ *+(s)I* {Poss.Sg3};	*alma terek* < *alma* ‘apple’ + *terek* ‘tree’ *+(s)I* {Poss.Sg3};

Table 6. Kirghiz, Karakalpak and Noghay compounds with(out) linking elements

The examples presented below demonstrate that in the Aral–Caspian Kipchak languages the linking element is used only in noun–noun or adjective–noun constructions, and only in subordinate and attributive compounds. Nevertheless, the linking element can be almost unexceptionally detected, when the originally Russian adjective plus noun compounds are translated word for word, as is presented in Table 7.

	Russian	**Kazakh**	**Kirghiz**	**Karakalpak**	**Noghay**
‘nervous system’	*nervnaja sistema*	*nerv sistemasï*	*nerv sistemasï*	*nerv sistemasï*	*nervlar sistemasï*
‘diabetes’	*saharnyj diabet*	*qant diabeti* < *qant* ’sugar’ + *diabet* ’diabetes, diabetic’ *+(s)I* {Poss.Sg3};	*qant diabeti* < *qant* ’sugar’ + *diabet* ’diabetes, diabetic’ *+(s)I* {Poss.Sg3};	*qant keseli* < *qant* ’sugar’ + *kesel* ’disease’ *+(s)I* {Poss.Sg3};	*seker diabeti* < *seker* ’sugar’ + *diabet* ’diabetes, diabetic’ *+(s)I* {Poss.Sg3}.

Table 7. Loan translations of Russian compounds in Aral–Caspian Kipchak languages

6. Reduplication

As was seen in Tables 1, 2, 3, and 4, coordinative compounds are a very special group in the Aral–Caspian Kipchak languages from the point of view of word formation. A great majority of coordinative compounds are formed through reduplication.[5] Even though onlytotal and partial reduplication (for more on the types of reduplication, see Wiltshire and Marantz 2000: 557–562) belong to compounding, they should be analysed withbinomes.[6] The reason for this method is that the meaning of these compound categories are very similar to each other. These phenomena are very productive in the Kipchak languages: they can form collective nouns, nouns with special meanings on the basis of the two consituents, and they can express intensification as well. Table 8 summarizes the various forms of coordinative compounds.

Total reduplication	**Partial reduplication**	**Synonym compounds**	**Hyponym compounds**
Kazakh *bara–bara* ‘continually, more and more’ < *bar*–‘to go’ – *A* {CV}; *dara–dara* ‘singly’ < *dara* ‘only, just’; *žal–žal* ‘stack, salient, avalanche’ < *žal* ‘mane, swell’;	**Kazakh** *nan–pan* ‘bread and other bakery products’ < *nan* ‘bread’; *tars–turs* ‘clattering noise’ < *tars* ‘manner, way, method’; *ühilep–ahïlap* ‘complaining and suffering’ < *ühile-* ‘to huff’ *–Ip* {CV};	**Kazakh** *ämir–qudiret* ‘power, strength’ < *ämir* ‘command, order, permit’ + *qudiret* ‘strength, power’; *dabïr–dübir* ‘shouting’ < *dabïr* ‘noise’ + *dübir* ‘noise’; *žer–düniye* ‘surface, the whole world’ < *žer* ‘place’ + *düniye* ‘world’;	**Kazakh** *as–su* ‘food’ < *as* ‘food’ + *su* ‘water’; *äke–šeše* ‘parents’ < *äke* ‘father’ + *šeše* ‘mother’; *dos–dŭšpan* ‘everybody’ < *dos* ‘friend’ + *dušpan* ‘enemy’;

[5] According to the definition “the term reduplication is applied to a type of word formation (in the broad sense, including both derivation and inflection) in which the phonological form of an affix is determined in whole or in part by the phonological form of the base to which it attaches” (Wiltshire and Marantz 2000: 557).

[6] Binomes (or twin words) can be divided into two subgroups on the basis of the constituents and the meaning of the compound: synonym compounds (hendiadys) and hyponym compounds (Johanson 1998: 50).

Kirghiz	**Kirghiz**	**Kirghiz**	**Kirghiz**
bāz–bāz ‘sometimes’ < *bāz* ‘some, a little’; *boz–boz* ‘dark brown’ < *boz* ‘grey’; *ǰeke–ǰeke* ‘singly, severally’ < *ǰeke* ‘individual, private’;	*kitep mitep ~ kitep sitep* ‘books’ < *kitep* ‘book’; *kök–sök* ‘vegetables’ < *kök* ‘blue, vegetable’; *mayda–čayda* ‘fiddle–faddle’ < *mayda* ‘small’;	*aġa–ini* ‘brothers’ < *aġa* ‘brother, elder brother’ + *ini* ‘younger brother’; *aqe–üke/aqe–ükö* ‘sisters’ < *aqe* ‘mother’ + *üke/ükö* ‘sister’; *köl–dayra* ‘lakes and seas, big lake’ < *köl* ‘lake’ + *dayra* ‘sea, lake’;	*alïš–berïš* ‘shopping, trade’ < *al–* ‘to buy, take’ *–Iš*+ {VN} + *ber-* ‘to give’ *–Iš* {VN}; *azïq–tülük* ‘food–stuff’ < *azïq* ‘food, feeding’ + *tülük* ‘food’; *keldi–ketti* ‘visit’ < *kel–* ‘to come’ *–DI* {Praet.Sg. 3} + *ket–* ‘to go’ *–DI* {Praet.Sg3};
Karakalpak	**Karakalpak**	**Karakalpak**	**Karakalpak**
mezgil–mezgil ‘sometimes, once in a while’ < *mezgil* ‘time, season’; *sonday–sonday* ‘either way, anyway’ < *sonday* ‘like that’; *töbe–töbe* ‘hilly area’ < *töbe* ‘hill’;	*adam–padam* ‘people, troops’ < *adam* ‘human, man’; *etik–petik* ‘boots and other footwears’ < *etik* ‘boots’; *sadaqa–padaqa* ‘burial feast’ < *sadaqa* ‘victim, commemoration’;	*ot–žem* ‘forage, feed’ < *ot* ‘grass’ + *žem* ‘food’; *qural–žaraq* ‘weaponry, armour’ < *qural* ‘weapon’ + *žaraq* ‘weapon’; *üy–žay* ‘flat, residence’ < *üy* ‘house’ + *žay* ‘residence’;	*ata–baba/ata–ana* ‘grandparents’ < *ata* ‘father’ + *ana* ‘mother’; *barïs–kelis* ‘behaviour, attitude’ < *bar–* ‘to go’ *–Is* {VN} + *kel–* ‘to come’ *–Is* {VN}; *kelim–ketim* ‘guests’ < *kel–* ‘to come’ *–Im* {CV} + *ket–* ‘to go’ *–Im* {CV};
Noghay	**Noghay**	**Noghay**	**Noghay**
ömir–ömirge ‘forever’ < *ömir* ‘life’ +*GA* {Dative}; *üzik–üzik* ‘staccato, jerky’ < *üzik* ‘snatch, wiff’; *zaman–zamanda* ‘sometimes’ < *zamam* ‘time’ +*DA* {Loc}.	*birem–sirem* ‘one by one’ < *birem* ‘once’; *köylek–möylek* ‘all kind of shirts’ < *köylek* ‘shirt’; *qasqïr–masqïr* ‘wolves’ < *qasqïr* ‘wolf’.	*bäle–qaza/qaza bale* ‘misfortune, trouble’ < *bäle* ‘misfortune’ + *qaza* ‘misfortune’; *xabar–xäter* ‘news’ < *xabar* ‘ news’ + *xäter* ‘news’; *yaw–dušpan* ‘enemies’ < *yaw* ‘enemy’ + *dušpan* ‘enemy’.	*bügün–erten* ‘fast, tight’ < *bügün* ‘today’ + *erten* ‘tomorrow’; *mezgilsiz–mekansïz* ‘unsuitable, inadequate’ < *mezgil* ‘season’ +*sIz* {NN/Adj.} + *mekan* ‘place’ +*sIz* {NN/Adj.}; *yetim–yesir* ‘orphans’ < *yetim* ‘orphan’ + *yesir* ‘oprhan’.

Table 8. Types of coordinative compounds in Aral–Caspian Kipchak languages

Taking a stock of the examples, some characteristics can be observed about their function and the usage. Compounds with total reduplication are used to represent idiomatic expressions or adverbs. They can establish collective nouns replacing suffixes, like the abstracness suffix *+LIK* (Johanson 1998: 36) and the suffixes *+KIl* expressing shades of colours (Erdal 1991: 98–99). Similarly to all Turkic languages, Kipchak languages form echo words by partial reduplication.[7] These compounds are translated as 'a thing etcetera', 'a thing and the like', and 'something and similar things'. In this case, the partial reduplicated word has an initial labial *m–*/*b–*/*p–* consonant (Johanson 1998: 50). However, it might be sometimes initial *s–*, or only vocal changes in the reduplicated form.

Synonym compounds express essentialy the plural form or represent a new concept with a meaning very close to the constituents. The most interesting group of coordinate compounds is that of hyponym compounds. Generally, it displays collectivity with antonyms, but it can semantically the plural form or word a new meaning, which is deduced from the basic meaning of the two constituents.

7. Verbs in compounding

In Turkic languages, there are two possibilities to form verbs: by suffixation (a synthetic method) or compounding (an analytical method). Analytical verb formation is very productive in most Turkic languages (Johanson 1998: 42). Verbs which serve as constituents of compounds can form compounds[8] with different parts of speech in the Aral–Caspian Kipchak languages: [Verb + Verb]Verb[9], [Verb + Verb]Noun, [Verb + Verb]Adverb, [Noun/Adjective + Verb]Verb. Verb + verb constructions which create nouns or adverbs are very rare.

The [Noun/Adjective + Verb]Verb compounds generally create idiomatic expressions or verbs with the meaning 'to do something'. In this case, the second constituent has the meaning 'to do'. (See Table 9.)

[7] About the process of reduplication, see Göksel–Kerslake (2005: 90–93). This system is very close to Aral–Caspian Kipchak reduplication.

[8]On verb formation in Aral–Caspian Kipchak, see Kirchner (1998a: 325–325), Csató and Karakoç (1998: 338–339), and Kirchner (1998b: 349–351).

[9] In this way of marking the elements in the brackets denote the word class of the constituents, while the third component (subscript) provides the class of the formed compound.

	Idiomatic expressions	**[Noun/Adjective + Verb]$_{Verb}$**
Kazakh	*aŋ al–* ‘to hunt, catch’ < *aŋ* ‘hunting’ + *al-* ‘to buy, take’; *wäde ber–* ‘promise’ < *wade* ‘to swear, promise’ + *ber–* ‘to give’;	*žumïs iste–* ‘to work’ < *žumïs* ‘work, labour’ + *iste–* ‘to do’; *sayaχat qïl–* ‘to wander’ < *sayaχat* ‘voyage’ + *qïl–* ‘to do’;
Kirghiz	*tamaq ič–* ‘to meal’ < *tamaq* ‘food’ + *ič-* ‘to drink’; *tameki tart–* ‘to smoke’ < *tameki* ‘tobacco’ + *tart–* ‘to pull’;	*ada qïl–* ‘to finish’ < *ada* ‘end’ + *qïl–* ‘to do’; *operaciya ǰasa–* ‘to operate’ < *operaciya* ‘operation’ + *ǰasa–* ‘to do’;
Karakalpak	*aytïp öt–* ‘to mention’ < *ayt–* ‘to say’ *–Ip* {CV} + *öt–* ‘to say’; *dem al–* ‘to have a rest’ < *dem* ‘respiration’ + *al-* ‘to buy, take’;	*duwa et–* ‘to pray’ < *duwa* ‘pray’+ *et–* ‘to do’; *buyrïq qïl–* ‘to act, to dispose’ < *buyrïq* ‘command, order’ + *qïl–* ‘to do’;
Noghay	*bala tap–* ‘to give birth’ < *bala* ‘child’ + *tap–* ‘to find’; *ötirik söyle-* ‘to lie’ < *ötirik* ‘lie’ + *söyle–* ‘to say’;	*habar et–* ‘to inform, to post’ < *habar* ‘news’ + *et–* ‘to do’; *süret yasa–* ‘to paint’ < *süret* ‘picture, painting’ + *yasa–* ‘to do’.

Table 9. Verbal compounds in Aral–Caspian Kipchak languages

The Aral–Caspian Kipchak languages have a special verbal compound category, when the [Verb + Verb]Verb constructions form a new meaning, which originally the two constituents did not have. Their construction is the same asthat of “auxiliary compounds”: the first contituent is conjoined to the second by a converbial form, but they can be contracted.

	Basic meaning	**Word-for-word translation**
Kazakh	*alïp bar– > apar–* ‘to carry’;	‘to take and go’
Kirghiz	*alïp ket– > apket–* ‘to carry away’;	‘to take and go’
Karakalpak	*alïp kel– > äkel-* ‘to bring’	‘to take and come’
Noghay	*alïp ber– > äper–* ‘to deliver, to put in’	‘to take and give’

Table 10. Verbal compounds with a new meaning

The most interesting group of the Aral–Caspian compounds is that of the [Verb + Verb]Noun/Adverb structures. These compounds are formed unexceptionally by the conjoining of two finite verbal forms, which express idiomatically a noun or an adverb.

Kazakh	*bolar–bolmas* 'a little bit, a bit' < *bol–* 'to be' *-Ar* {Aor.Sg3} + *bol–* 'to be' +*mAs* {Aor.Neg.Sg3}
Kirghiz	*keldi–ketti* 'visit, observation' < *kel–* 'to come'*-DI* {Past.Sg3} + *ket–*'to go' *-DI* {Past.Sg3}
Karakalpak	*bolar–bolmas* 'hardly, barely' < *bol–* 'to be' *-Ar* {Aor.Sg3} + *bol–* 'to be' +*mAs* {Aor.Neg.Sg3}
Noghay	*bolsa bolar* 'possibly, maybe' < *bol–* 'to be'*-sA* {Cond.Sg3} + *bol–* 'to be' *-Ar* {Aor.Sg3}

Table 11. [Verb + Verb]$_{\text{Noun/Adverb}}$ compound structures

On the basis of the MorboComp classification of compounds (Bisetto and Scalise 2005:321–328), the verbal constructions must be considered as coordinatives from the semantic point of view.

8. Headedness

As was highlighted above, the headedness (or more exactly, the presence or the absence of the head) is one of the criteria for the classification of compounds. Göksel and Haznedar (2007) discuss some characteristics of the headedness of the Turkish, which shows a lot of similarities with the Aral–Caspian Kipchak languages. On the basis of the collected corpus and the presented examples in this study, the compounds can be divided into three classes in these languages: one–headed compounds, double–headed compounds, and headless compounds.[10] The one-headed compounds are represented in the endocentric class of subordinate and attributive groups, and they are typically right-headed. Nevertheless, there are left-headed structures as well, although they constitute an unusual phenomenon in the Kipchak languages. Left-headedness occurs in the case of the izafet structures, which remains as the heritage of the former Chagathay literature languages.[11] This

[10] All of the exocentric compounds are considered to be headless.

[11] The Aral–Caspian Kipchak people used earlier as written languages the Chaghatay tradition (Boeschoten and Vandamme 1998: 167–169).

construcion was copied from Persian (Boeschoten and Vandamme 1998: 174–175).[12] Izafet structures are not used in word formation. These compounds are idiomatic, like the name of the Quran: Kazakh *qurani kärim*; Kirghiz *qurani qarim*; Karakalpak *quranï kärim* or Noghay *quranï kerim*.

Another special one-headed construction can be detected in the Aral–Caspian Kipchak languages when the head relation changes between the constituents. So, in this sense, the head might be optional. This kind of constructions occurs only in the attributive group. They can optionally substitute adenominal suffix, which can form adjectives (Baskakov 1958: 810; Baskakov 1963: 511; Kirchner 1998a: 322; Kirchner 1998b: 347).

Kazakh	**Kirghiz**	**Karakalpak**	**Noghay**
‘generous, propitious’	‘crazy, insane’	‘honest, true’	‘pregnant’
qolï ašïq < *qol* ‘arm’ *+(s)I* {Poss.Sg3} + *ašïq* ‘free, open, clear’;	*bašï del* < *baš* ‘head’ *+(s)X* {Poss.Sg3} + *del* ‘fool’;	*niyeti χaq* < *niyet* ‘intention’ *+(s)I* {Poss.Sg3} + *χaq* ‘true, correct’;	*ayağï awïr* < *ayaq* ‘leg’*+(s)I* {Poss.Sg3} + *awïr* ‘heavy’;
ašïq qoldï < *ašïq* ‘free, open, clear’ + *qol* ‘arm’ *+LI* {NN/Adj.};	*del baštū* < *del* ‘fool’ + *baš* ‘head’ *+LŪ* {NN/Adj.};	*χaq niyetli* < *χaq* ‘true, correct’ + *niyet* ‘intention’ *+LI* {NN/Adj.};	*awïr ayaqlï* < *awïr* ‘heavy’ + *ayaq* ‘leg’ *+LI* {NN/Adj.}.

Table 12. Compounds with ‘optional head’ in Aral–Caspian Kipchak

From the perspective of the semantic field, these optionally headed compounds appear only as such attributive constructions, which designate internal and external properties, so as a part of speech they must be considered to be adjectives. However, the linking element can be changed depending on which person it should mark. Therefore, the linking element is a part of the possessive paradigm in the optionally headed constructions. This can be expressed by nominal inflection as well, when the right-headed construction is used with the adjectival suffix. (See the example in Kazakh in Table 13.)

[12] In this case the Persian *–i* linking element is attached to the first constituent of the construction, making it the head.

(*Meniŋ*) *niyetim aq* → *niyet +(I)m* {Poss.Sg1}	cf. (*Men*) *aq niyettimin* ‘I am generous’
(*Bizdiŋ*) *niyetimiz aq* → *niyet +(I)mIz* {Poss.Pl1}	cf. (*Biz*) *aq niyettimiz* ‘We are generous’
(*Olardïŋ*) *niyeti aq* ‘their intention is good’ → *niyet +(s)I* {Poss.Pl3}	cf. (*Olar*) *aq niyetti*[13] ‘They are generous’

Table 13. Parts of the paradigm of the possessive and personal markers

The topic of the double-headed (or two headed) compounds has already been partially touched upon in connection with reduplication. Categorically, the double-headed contructions constitute a group of endocentric coordinate compounds. From the semantic point of view, they create collective nouns (see Tables 1, 2, 3, and 4) or new words, which are related to the basic meaning of the two constituents (generally binomes). And as has been mentioned above, morphologically, they can replace suffixes, like the abstractness suffix *+LIK* and the plural marker *+LAr* (Johanson 1998: 36; 38).

Kazakh	Kirghiz	Karakalpak	Noghay
as–su ~ *astïq* ‘nutrition, food’ < *as* ’food’ + *su* ’water’	*adïr–čïbïr* ~ *adïrdūluq* ‘hills, hilly region’ < *adïr* ‘hill’ + *čïbïr* ‘mountain, mountain range’; *adïr* ‘hill’ *+DX* {NN/Adj.} *+LXK* {NN}	*aġayïn–tuwġan* ~ *aġayïnlar*; *tuwġanlar* ‘relatives’ < *aġayïn* ‘relative’ + *tuw–* ‘to be born’ *–GAn+* {CV}; *aġayïn* ’relative’ *+LAr* {Plur}; *tuw–* ‘to be born’ *–GAn+* {CV} *+LAr* {Plur}	*aġalï–inili* ~ *qardašlar* ‘brothers’ < *aġa* ‘brother, elder brother’ + *DI* {NN/Adj.} + *ini* ‘younger brother’ *+DI* {NN/Adj.}; *qardaš* ’brother’ *+LAr* {Plur}

Table 14. Two-headed compounds

[13] There is no difference in the paradigm of the possessive and the personal markers between the singular and the plural third person forms (Kirchner 1998a: 324–326).

9. Suffixation or compounding?

The special function of compounds in word formation is that they can replace suffixes. This function has already come under consideration many times. However, there are some special constituents in compounds which originally have an autonomous usage, but as a part of a compound they behave as suffixes. The most conspicuous in this case is that the boundary between suffixation and compounding is not clear at all. A list of these words is given with a short explanation in the Aral–Caspian Kipchak languages:

(1) Kazakh *χana* 'residence, adress, room'; Kirghiz *qana* 'place, room'; Karakalpak *χana* 'place, house'; Noghay – ← Iranian: cf. Persian *khāna* 'house, dwelling, tent' (Steingass 1996: 444).

The word is undoubtedly of Persian origin. Its usage is very frequent except in Noghay. It is found as a lexical item as well, but in compounds it appears like a suffix which forms places, institutions and all kinds of buildings which are connected to a special activity. In Noghay, these sort of words are mostly expressed by Russian borrowings.

	Kazakh	**Kirghiz**	**Karakalpak**	**Noghay**
'hospital'	*awrwχana* < *awrw* 'ill, disease'	*ōruqana* < *ōru* 'ill, disease'	*keselχana* < *kesel* 'ill, disease'	*gospital'*
'pharmacy'	*däriχana* < *däri* 'medicine'	*darïqana* < *darï* 'medicine'	*däriχana* < *däri* 'medicine'	*apteka*
'dormitory'	*žataqχana* < *žat–* 'to lie' *–AK* {VN}	*ǰataqana* < *ǰat–* 'to lie' *–XK* {VN}	*žataqχana* < *žat–* 'to lie' *–AK* {VN}	*obščežitie*
'lavatory, toilet'	*äžetχana*/*däretχana* < *äžet* 'need'/*däret*'stool'	*aǰatqana*/*dāratqana* < *aǰat* 'need'/*dārat* 'stool'	*häžžetχana*/ *däretχana* < *häžžet* 'need'/*däret* 'stool'	*äžetqana* < *äžet* 'need'

Table 15. Semi-affixes in Aral–Caspian Kipchak

(2) Kazakh *qora* 'court, courtyard'; Kirghiz *qorō* 'court, courtyard'; Karakalpak *qora* 'court, courtyard'; Noghay – ~ Old Turkic *qorïg* 'an enclosure, enclosed area' (Clauson 1972: 652b).

The usage of this word as a suffix is very similar to the former example. It is found only in Kazakh and Karakalpak as a method of word formation. It forms only words which are related to agriculture and animal husbandry and include an enclosed place. Kirghiz has the semi-affix *–qana* in all but one of the forms, meanwhile Noghay uses other suffixes or the word *avla* ‘court, courtyard’.

	Kazakh	**Kirghiz**	**Karakalpak**	**Noghay**
‘animal farm’	*malqora* < *mal* ‘animal, wealth’ + *qora* ‘court, courtyard’	*malqana* < *mal* ‘animal, wealth’ + *qana* ‘place, room’	*malqora* < *mal* ‘animal, wealth’ + *qora* ‘court, courtyard’	*mal avla* < *mal* ‘animal, wealth’ + *avla* ‘court, courtyard’
‘stable’	*atqora* < *at* ‘horse’ + *qora* ‘court, courtyard’	*atqana* < *at* ‘horse’ + *qana* ‘place, room’	*atχana* < *at* ‘horse’ + *χana* ‘place, room’	*atlïq* < *at* ‘horse’ +*LIK* {NN}

Table 16. Semi-affixes of Turkic origin

Kazakh *ögiz qora* ‘stable for oxen’ cf. **Kirghiz** *qoy qorō* ‘manger’
šošqa qora ‘pigpen’ cf. **Kirghiz** *čočqoqana* ‘pigpen’
qoyanqora ‘rabbithutch’

Karakalpak *ǰemqora* ‘manger’
otqora ‘loft’
qoyanqora ‘rabbithutch’

(3) Kazakh *nama* ‘letter, writing’; Kirghiz *nāma* ‘holy book, scripture’; Karakalpak – ; Noghay – ← Iranian: cf. Persian *nāma* ‘a writing, letter, epistle’ (Steingass 1998: 1380).

The originally Persian word is used as an affix actively only in Kazakh. In Karakalpak, there are only few examples, but it is not detected in Kirghiz and Noghay. Generally as a suffix it forms words related to written documents and papers (Kazakh *azanama* ‘obituary’; *ǧarïšnama* ‘cosmogony’; *tariχnama* ‘historiography’).

	Kazakh	**Kirghiz**	**Karakalpak**	**Noghay**
‘yearbook’	*žïlnama* < *žïl* ‘year’	*jïlbayan* < *jïl* ‘year’ + *bayan* ‘short story, story’	*žïlnama* < *žïl* ‘year’	*letopis’* ← Russian
‘contract’	*šartnama* < *šart* ‘condition’	*kelišim/kontrakt*	*šartnama* < *šart* ‘condition’	*kontrakt* ← Russian

Table 17. The usage of Persian *nāma* in Aral–Caspian languages

(4) Kazakh *qumar* ‘desire, request, will’; Kirghiz *qumar* ‘passion, desire’; Karakalpak *qumar* ‘desire, passion’; Noghay *qumar* ‘habit, request’ ← Iranian: cf. Persian *khumār* ‘the effect of love, of drowsiness, of drinking’ (Steingass 1998: 474).

This word was borrowed from Persian in all of the Aral–Caspian languages, but it is plays a role in word formation only in the Kazakh language. Additionally, it can replace three adjectival suffixes, namely, the intensifying *+GOy*; *+šIl* and *+šAŋ* (Balakaev, Baskakov and Kenesbaev 1962: 140; 185; 203) and nominal suffix *+qor* (Balakaev, Baskakov and Kenesbaev 1962: 140):

Kazakh

‘rapacious’: *aqšaqumar* < *aq* ‘white’ + *qumar* ‘desire, request, will’;
aqšašïl < *aq* ‘white’ *+šIl* {NN/Adj.};

‘suitor’: *arïzqumar* < *arïz* ‘wish, desire’ + *qumar* ‘desire, request, will’;
arïzqoy < *arïz* ‘wish, desire’ *+GOy* {NN/Adj.};

‘a person who likes jokes’:
äzilqumar < *äzil* ‘joke, jest’ + *qumar* ‘desire, request, will’;
äzilšil < *äzil* ‘joke, jest’ *+šIl* {NN/Adj.};
äzilqoy < *äzil* ‘joke, jest’ *+GOy* {NN/Adj.};
äzilšeŋ < *äzil* ‘joke, jest’ *+šAŋ* {NN/Adj.};

‘verbose’: *äŋimequmar* < *äŋime* ‘conversation, story’ + *qumar* ‘desire, request’;
äŋimeqoy < *äŋime* ‘conversation, story’ *+GOy* {NN/Adj.};

‘vainness, vanity’:
mansapqumar < *mansap* ‘place, job, career’ + *qumar* ‘desire, request’;
mansapqor < *mansap* ‘place, job, career’ *+qor* {NN/Adj.};
mansapšïl < *mansap* ‘place, job, career’ *+šIl* {NN/Adj.}.

Kazakh and Karakalpak show almost the system from the perspective of the semi-affixes, and Kirghiz is partially similar too, but Noghay has a totally different system, which might be explained by the spatial distance from the other three languages.

10. Conclusion

Through the analysis of the compounds in the Aral–Caspian Kipchak languages, it becomes evident that this type of word formation is as productive as suffixation. These languages represent a very wide and varied system regarding the structural and semantic characteristics of compounds. Therefore, it is almost impossible to suggest a classification which could not separate well the compounds into groups without overlaps. From the semantic point of view, the compounds can replace in a lot of cases nominal and adjectival suffixes. For further research, it would be useful to go into further detail regarding the topic of the common characteristics of compounding and suffixation.

Abbreviations

AOR	Aorist
COND	Conditional
CV	Converb
DAT	Dative
LOC	Locative
NN	Denominal noun suffix
NN/ADJ.	Denominal noun suffix forming adjectives
PAST	Past tense, third person singular
PL	Plural
POSS	Possessive, third person singular
RED	Reduplication
VN	Deverbal noun suffix

References

Bağrıaçık, Metinand Ralli, Anagela 2014. NN *–sI* concatenations in Turkish: Construct state nominals and phrasal compounds. *MIT Working Papers in Linguistics* 67: 13–24.

Balakaev, Maulen B., Baskakov, Nikolaj A. and Kenesbaev, Smet K. 1962. *Sovremennyj kazaxskij jazyk*. Almaty: Akademii nauk Kazahskij SSR.

Baskakov, Nikolaj A. 1940. *Nogajskij jazyk i ego dialekty: grammatika, teksty i slovar'*. Moskva: Akademii nauk SSSR.

Baskakov, Nikolaj A. 1958. *Karakalpaksko–russkij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx u nacional'nyx slovarej.

Baskakov, Nikolaj A. 1963. *Nogajsko–russkij slovar'*. Moskva: Gosudarstvennoe izdatel'stvo inostrannyx u nacional'nyx slovarej.

Baskakov, Nikolaj A. 1967. *Russko–karakalpakskij slovar'*. Moskva: Sovetskaja ėnciklopedija.

Bauer, Laurie 2001. Compounding. In: Haspelmath, Martin et al. (eds.) *Language typology and language universals* I. Berlin: Walter de Gruyter. 695-707.

Bektaev, Kaldybaev B. 1999. *Ülken kazakša–orysša orysša–kazakša sözdik = Bol'šoj kazahsko–russkij russko–kazaxskij slovar'*. Almaty: Kazyna.

Bisetto, Antonietta& Scalise, Sergio 2005. The classification of compounds. *Lingue e Linguaggio* 4 (2): 319–332.

Boeschoten, Hendrikand Vandamme, Mark 1998. Chaghatay. In: Johanson, Lars and Csató, Éva (eds.) *The Turkic languages*. London & New York: Routledge. 166–178.

Booij, Geert, Lehmann, Christian et al. (eds.) 2000. *Morphology: An international handbook on infelction and word–formation* I–II. Berlin & New York: Walter De Gruyter.

Booij, Geert 2005. *The grammar of words*. Oxford: Oxford University Press.

Bloomfield, Leonard 1933. *Language*. New York: Holt.

Bybee, Joan 2000. Verb. In: Booij, Geert, Lehmann, Christian et al. (eds). *An international handbook on infelction and word–formation* I. Berlin–New York: Walter de Gruyter. 794–808.

Clauson, Gerard 1972. *An etymological dictionary of pre–thirteenth–century Turkish*. Oxford: Clarendon Press.

Csató, Éva & Karakoç, Birsel 1998. Noghay. In: Johanson, Lars and Csató, Éva (eds.) *The Turkic languages*. London & New York: Routledge. 333–343.

Dede, Müşerref 1978. *A syntactic and semantic analysis of Turkish nominal compounds*. Michigan: University of Michigan.

Erdal, Marcel 1991. *Old Turkic word formation* I–II. Wiesbaden: Harrassowitz.

Fabb, Nigel 1998. Compounding. In: Spencer, Andrew and Zwicky, Arnold M. (eds.) *Handbook of morphology*. Oxford. 66–83.

Göksel, Aslıand and Kerslake, Celia 2005. *Turkish: A comprehensive grammar*. London: Routledge.

Haspelmath, Martin 2002. *Understanding morphology*. London: Arnold.

Imart, Guy 1981. *Le kirghiz (turk d'Asie Centrale soviétique). Description d'une langue de littérisation récente*. Aix–en–Provence: Publications de l'Université de Provence.

Johanson, Lars 1998. The structure of Turkic. In: Johanson, Lars and Csató, Éva (eds.) *The Turkic languages*. London & New York: Routledge. 30–66.

Johanson, Lars and Csató, Éva 1998. *The Turkic languages*. London & New York: Routledge.

Judahin, Konstantin K. 1965. *Kirgizsko–russkij slovar'*. Moskva: Sovetskaja ėnciklopedija.

Koç, Kenan 2003. *Kazak Türkçesi Türkiye Türkçesi Sözlüğü*. Ankara: Akçağ.

Kirchner, Mark 1998a. Kazakhs and Karakalpaks. In: Johanson, Lars and Csató, Éva (eds.) *The Turkic languages*. London & New York: Routledge. 318–332.

Kirchner, Mark 1998. Kirghiz. In: Johanson, Lars and Csató, Éva (eds.) *The Turkic languages*. London & New York: Routledge. 344–356.

Schönig, Claus 2007. Some notes on Modern Kipchak Turkic. *Ural–Altaische Jahrbücher* 21. 170–202.

Spencer, Andrew 1991. *Morphological theory*. Oxford: Basil Blackwell.

Steingass, Francis J. 1996. *A comprehensive Persian–English dictionary*. (Rev. ed.) Delhi: Munshiram Manoharlal.

Wiltshire, C. & Marantz, A. 2000. Reduplication. In: Booij, Geert, Lehmann, Christianet al. (eds). *Morphology: An international handbook on inflection and word–formation* I. Berlin & New York: Walter de Gruyter. 557–567.

Online sources

Göksel, Aslı & Haznedar, Belma 2007. *Remarks on compounding in Turkish*. (MorboComp Project, University of Bologna). http://componet.sslmit.unibo.it/download/remarks/TR.pdf. Date of access: 04.15.2017

Hofwegen, von Janike 2014.*The role of –sI marking in Kazakh compound nouns: Implications for the relationship between case–licensing and agreement.* Stanford. http://web.stanford.edu/~jmvanhof/index_files/Kazakh_final_paper.pdf. Date of access: 04.15.2017.

Krejci, Bonnie & Glass, Leila 2015. *The noun/adjective distinction in Kazakh*. Stanford. http://web.stanford.edu/~lelia/krejci_glass.pdf. Date of access: 04.15.2017.

Studia uralo-altaica

No. 1 Róna-Tas, A. – Fodor, S. *Epigraphica Bulgarica*. (1973)

No. 2 *Die erste sölkupische Fibel aus dem Jahre 1879*. Eingeleitet von P. Hajdú. (1973)

No. 3 Новицкий, Гр. *Краткое описание о народе остяцком (1715)*. (1973)

No. 4 Paasonen, H. *Tschuwaschisches Wörterverzeichnis*. Eingeleitet von A. Róna-Tas. (1974)

No. 5 A. Molnár, F. *On the History of Word-Final Vowels in the Permian Languages*. (1974)

No. 6 Hajdú, P. *Samojedologische Schriften*. (1975)

No. 7 *N. Witsens Berichte über die uralischen Völker.* Aus dem Niederländischen ins Deutsche übersetzt von T. Mikola. (1975)

No. 8 Ph. Johann von Strahlenberg *Das Nord und Ostliche Theil von Europa und Asia*. With an Introduction by I. R. Kruger. (1975)

No. 9 Kiss, J. *Studien zur Wortbildung und Etymologie der finnisch-ugrischen Sprachen*. (1976)

No. 10 U. Kőhalmi, K. *Chrestomathia Sibirica. Auswahl aus der Volksdichtung der sibirischen Urvölker.* (1977)

No. 11 Mikola, T. *Materialien zur wotjakischen Etymologie*. (1977)

No. 12 Попова, Я. Н. *Ненецко-русский словарь. Лесное наречие*. (1978)

No. 13 Tardy, L. *Beyond the Ottoman Empire*. (1978)

No. 14 Clauson, G. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish, Index*. Vol. 1, With a Preface by A. Róna-Tas. (1981)

No. 15 Clauson, G. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish, Index*. Vol. 2. (1982)

No. 16 Honti, L. *Nordostjakisches Wörterverzeichnis*. (1982)

No. 17 *Studies in Chuvash Etymology*. 1. Edited by A. Róna-Tas. (1982)

No. 18 Domokos, P. *Handbuch der uralischen Literaturen*. (1982)

No. 19 M. Korchmáros, V. *Definiteness as Semantic Content and its Realization in Grammatical Form*. (1983)

No. 20 Tardy, L. *Sklavenhandel in der Tartarei*. (1983)

No. 21 Berta, Á. *Die russischen Lehnwörter in der Mundart der getauften Tataren*. [Tatarische etymologische Studien I.] (1983)

No. 22 Helimski, E. *The Language of the First Selkup Books*. (1983)

No. 23 Pusztay, J. *Die Pur-Mundart des Waldjurakischen. Grammatikalischer Abriss aufgrund der Materialien von T. V. Lehtisalo*. (1984)

No. 24 Radics, K. *Typology and Historical Linguistics. Affixed Person-Marking Paradigms*. (1985)

No. 25 Róna-Tas, A. *Language and History. Contributions to Comparative Altaistics*. (1986)

No. 26 Keresztes, L. *Geschichte des mordwinischen Konsonantismus* 2. *Etymologisches Belegmaterial*. (1986)

No. 27 Keresztes, L. *Geschichte des mordwinischen Konsonantismus* 1. (1987)

No. 28 *Papers on Derivation in Uralic. Szegeder und Turkuer Beiträge zur uralischen Derivation.* Edited by J. Vastag (1987).

No. 29 Schulze, B. *Der Wortparallelismus als ein Stilmittel der (nord-) ostjakischen Volksdichtung*. (1988)

No. 30 *Tatarische etymologische Studien* II. Edited by Árpád Berta (1988)

No. 31 Berta, Á. *Lautgeschichte der tatarischen Dialekte*. (1989)

No. 32 Zimonyi, I. *The Origins of the Volga Bulgars*. (1990)

No. 33 Róna-Tas, A. *An Introduction to Turkology*. (1991)

No. 34 Bereczki, G. *Grundzüge der tscheremissischen Sprachgeschichte* 2. (1992)

No. 35 Bereczki, G. *Grundzüge der tscheremissischen Sprachgeschichte* 1. (1994)

No. 36 Mikola, T. *Morphologisches Wörterbuch des Enzischen*. (1995)

No. 37 Doerfer, G. *Formen der älteren türkischen Lyrik*. (1996)

No. 38 Kakuk, Zs. *Mischärtatarische Texte mit Wörterverzeichnis*. (1996)

No. 39 *Historical and linguistic interaction between Inner-Asia and Europe, Proceedings of the 39th Permanent International Altaistic Conference [PIAC], Szeged, Hungary: June 16–21, 1996*. Edited by Árpád Berta. (1997)

No. 40 Rybatzki, V. *Die Toñuquq-Inschrift*. (1997)

No. 41 Helimski, E. *Die matorische Sprache. Wörterverzeichnis − Grundzüge der Grammatik − Sprachgeschichte*, unter Mitarbeit von Beáta Nagy. (1997)

No. 42 Мecарош, Э. *Словообразовательные суффиксы глагола в эрзянском языке*. (1999)

No. 43 Wagner-Nagy, B. *Die Wortbildung im Nganasanischen*. (2001)

No. 44 Ivanics, M.–Usmanov, M. A. *Das Buch der Dschingis-Legende. Däftär-i Čingiz-namä* 1. (2002)

No. 45 Mikola, T. *Studien zur Geschichte der samojedischen Sprachen*, Aus dem Nachlass herausgegeben von Beáta Wagner-Nagy. (2004)

No. 46 Körtvély, E. *Verb conjugation in Tundra Nenets*. (2005)

No. 47 *Turcology in Turkey. Selected papers.* Edited by László Károly, papers selected by Nurettin Demir and Emine Yılmaz. (2007)

No. 48 Rassadin, V. I. *Soyotica*. Edited by Béla Kempf. (2010)

No. 49 *The Szeged Conference. Proceedings of the 15th International Conference on* Turkish *Linguistics held on August 20-22, 2010 in Szeged*. Edited by Kincses-Nagy, É. & Biacsi, M. (2012)

No. 50 Kakuk, Zs. Krimtatarisches Wörterverzeichnis. Aufgrund der Sammlung von Ignác Kúnos zusammengestellt (2012)


UNIVERSITAS SCIENTIARUM SZEGEDIENSIS
SZEGEDI TUDOMÁNYEGYETEM
SCIENTIA VIRTUS
LIBERTAS