

## Werk

**Titel:** Number 2

**Ort:** Wiesbaden

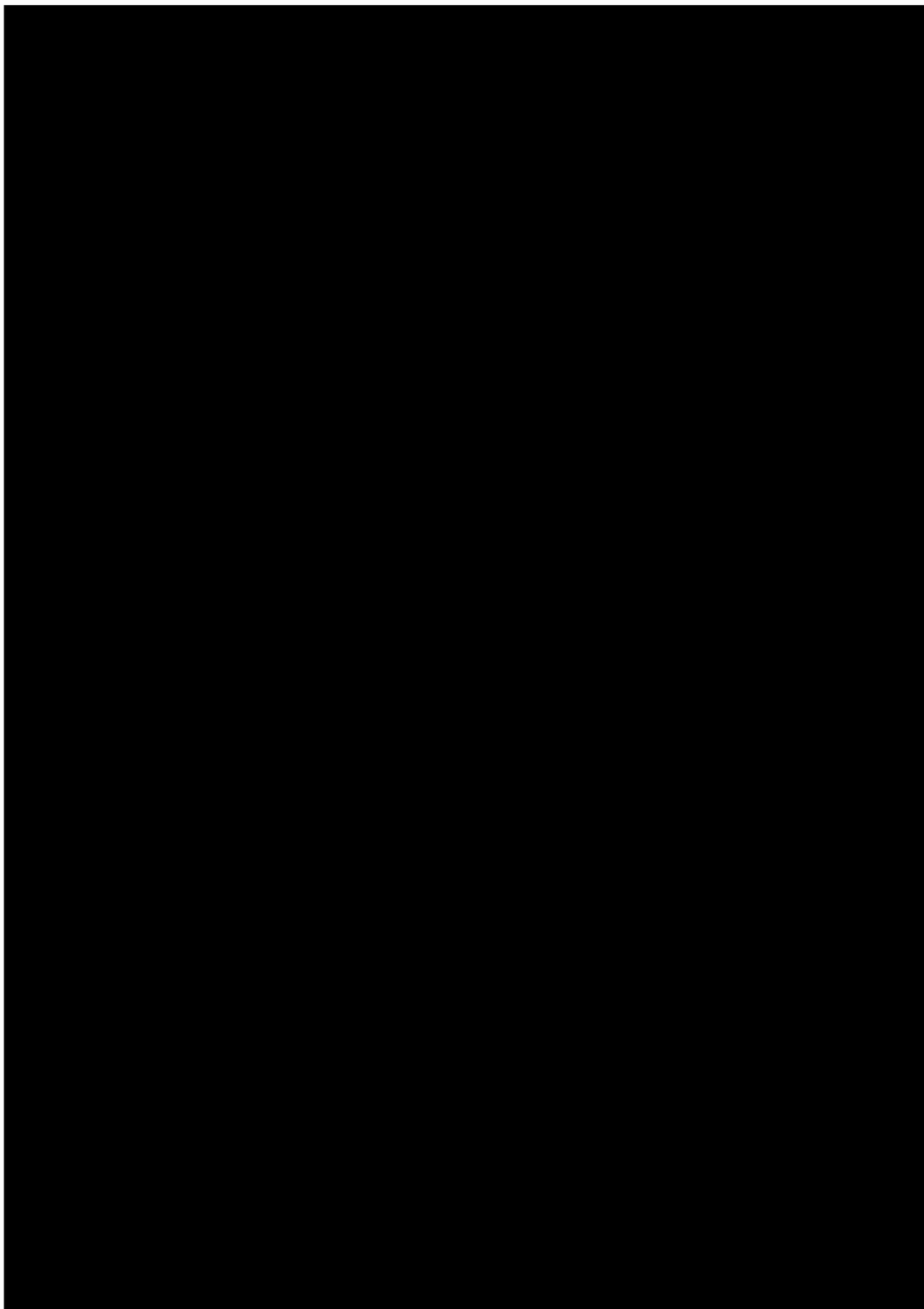
**Jahr:** 2018

**PURL:** [https://resolver.sub.uni-goettingen.de/purl?666048797\\_0022](https://resolver.sub.uni-goettingen.de/purl?666048797_0022) | LOG\_0021

## Kontakt/Contact

Digizeitschriften e.V.  
SUB Göttingen  
Platz der Göttinger Sieben 1  
37073 Göttingen

✉ [info@digizeitschriften.de](mailto:info@digizeitschriften.de)







# Turkic Languages

Edited by  
Lars Johanson

in cooperation with  
Hendrik Boeschoten, Bernt Brendemoen,  
Éva Á. Csató, Peter B. Golden, Tooru Hayasi, László Károly,  
Astrid Menz, Irina Nevskaya, Sumru A. Özsoy,  
Saule Tazhibayeva, Abdurishid Yakup

22 (2018) 2

Harrassowitz Verlag · Wiesbaden

The journal **TURKIC LANGUAGES** is devoted to linguistic Turcology. It addresses descriptive, comparative, synchronic, diachronic, theoretical and methodological problems of the study of Turkic languages including questions of genealogical, typological and areal relations, linguistic variation and language acquisition. The journal aims at presenting work of current interest on a variety of subjects and thus welcomes contributions on all aspects of Turkic linguistics. It contains articles, review articles, reviews, discussions, reports, and surveys of publications. It is published in one volume of two issues per year with approximately 300 pages.

Manuscripts for publication, books for review, and all correspondence concerning editorial matters should be sent to Prof. Dr. Dr. h.c. Lars Johanson, Turkic Languages, Department of Slavistics, Turcology and Circum-Baltic Studies, University of Mainz, 55099 Mainz, Germany. The email address [johanson@uni-mainz.de](mailto:johanson@uni-mainz.de) may also be used for communication.

Books will be reviewed as circumstances permit. No publication received can be returned.

Subscription orders can be placed with booksellers and agencies. For further information please contact: Harrassowitz Verlag, 65174 Wiesbaden, Germany; fax: 49-611-530999; e-mail: [verlag@harrassowitz.de](mailto:verlag@harrassowitz.de).

© Otto Harrassowitz GmbH & Co. KG, Wiesbaden 2018

This journal, including all of its parts, is protected by copyright. Any use beyond the limits of copyright law without the permission of the publisher is forbidden and subject to penalty. This applies particularly to reproductions, translations, microfilms and storage and processing in electronic systems.

Printing and binding by Hubert & Co., Göttingen

Printed on permanent/durable paper

Printed in Germany

[www.harrassowitz-verlag.de](http://www.harrassowitz-verlag.de)

ISSN 1431-4983

## Contents

### Turkic Languages, Volume 22, 2018, Number 2

Editorial note by Lars Johanson .....	149
---------------------------------------	-----

#### Obituary

István Vásáry: In memoriam András J. E. Bodrogligeti (1925–2017) .....	151
--	-----

#### Articles

Jaklin Kornfilt: NP versus DP: Which one fits Turkish nominal phrases better? .....	155
Hasmik Kirakosyan & Ani Sargsyan: The educational role of the Late Medieval Persian-Ottoman Turkish bilingual dictionaries. The codices of the Matenadaran .....	167
Henryk Jankowski: Krymchak language samples .....	175
Irina Nevskaya & Larisa Tybykova & Mikhail Vavulin & Olga Zaytseva & Evgeniy Vodyasov: 3D documentation of Old Turkic Altai runiform inscriptions and revised readings of the inscriptions Tuekta-V and Bichiktu-Boom-III .....	194
Elisabetta Ragagnin: A Turcological gem: The Tuhan language of Northern Mongolia .....	217
Gregory D. S. Anderson: The multi-functional converb <i>-GAŝ</i> and related forms in Sayan Turkic .....	230
Chris Lasse Däbritz: On <i>kim</i> , <i>kimne</i> :- and <i>kaŋa</i> :-. Three placeholder items in Dolgan .....	286



## Editorial note

*Turkic Languages, Volume 22, 2018, Number 2*

The present issue of *TURKIC LANGUAGES* starts with three papers on languages of the western part of the Turkic-speaking world.

Jaklin Kornfilt raises the question of whether Turkish, according to principles in the framework of formal linguistics, is an “NP language” or a “DP language”. The question concerns the suggested typological divide between languages whose traditionally defined noun phrases are determiner phrases (DP) and languages whose noun phrases only extend to NP. It has been claimed that Turkish belongs to the latter type, and a number of relevant properties have been proposed. The author argues that, at best, the properties can be explained in other ways, and that, at worst, the NP hypothesis makes wrong predictions with respect to Turkish. The phenomena that are claimed to characterize Turkish as an NP language fail to establish such a characterization. Analyzing Turkish noun phrases as determiner phrases is at least as successful.

Hasmik Kirakosyan & Ani Sargsyan deal with medieval Persian-Ottoman bilingual dictionaries, in particular copies kept in the Mesrop Mashtots Institute of Ancient Manuscripts in Yerevan. Dictionaries of this kind, so-called *farhangs*, were compiled as educational tools and used in *medreses* to help with reading and understanding the patterns of Persian classical literature, and also for teaching the Arabic-Persian loan vocabulary and the parts of Persian grammar that were relevant for Ottoman Turkish. The manuals are important witnesses of the linguistic interrelations and cultural milieu of their time.

Henryk Jankowski presents a small corpus of spoken Krymchak recorded with one single informant. Written Krymchak is well-documented in modern and pre-modern texts representing a variety of styles. The spoken language, however, is very poorly documented. Even a small corpus, such as the one presented here, is thus of high value. Competent native speakers of Krymchak are, however, hardly available today. The linguistic competence of this informant, born in 1924 in Simferopol, was relatively limited. She could remember to some extent the language she heard and spoke in her childhood. According to the author, Krymchak as spoken by this informant is morphologically a NW Turkic language, but phonologically a mixture of SW and NW Turkic.

Four contributions deal with the northern-most parts of Turcia.

Irina Nevskaya, Larisa Tybykova, Mikhail Vavulin, Olga Zaytseva, and Evgeniy Vodyasov report on the results of field research in the Altai Mountains, where a number of Turkic runiform rock inscriptions were digitalized by means of three-dimensional technologies. The outcome has led to revised readings of two inscriptions on the basis of new analyses of their runiform signs. One of the texts sheds

light on the ancient religious terminology of the area. The recorded data will soon be freely accessible to all interested researchers.

Elisabetta Ragagnin presents brand new data on Tuhan, a small Sayan Turkic language spoken in the county of Tsagaan Üür in the East Khövsgöl region of Mongolia. It is moribund, only used by a handful of old speakers and semi-speakers. While displaying typical Sayan Turkic characteristics, it preserves a number of otherwise unknown archaic features. It shares certain isoglosses with Tofan, and others with Altay and Khakas. Since Tuhan appears to be highly important for comparative Turcology, the author quite rightly characterizes it as a precious “gem” among the South Siberian Turkic languages.

Gregory D. S. Anderson deals with the wide range of functions of the Sayan Turkic converb in {-GAš}, especially its roles in complex sentences and narrative structures. It appears to be in the process of being pushed into the finite system as an anterior marker. Its most salient and common feature is the propulsive function, the power to advance the discourse when expressing narratively equal events. In some taiga varieties, the converb is claimed to have been grammaticalized as a same-subject marker in a switch-reference system. Varieties such as Jungar Tuvan and Dukhan have a tail-head linkage system, where the converb is used to set off chunks of discourse by copying sequences of the finite predicative element of a sentence at the beginning of the following sentence. In varieties such as Tofan and Soyot, the use of {-GAš} has been extended into the domains of the {-p} converb in complex predicates and auxiliary verb constructions.

Chris Lasse Däbritz presents a corpus-based analysis of three so-called placeholders (lexical fillers) in the North-Siberian Turkic language Dolgan. These are lexical items that replace other lexical items in a clause. While searching for the right lexical item, the speaker typically fills the arising pause with a nominal or verbal placeholder. The items *kim* and *kimne:-* occupy the same syntactic slot as the substituted lexical item, whereas *kaña:-* functions differently. The Dolgan placeholders are relatively complex in form and function. They are investigated here with respect to their morphosyntactic properties, functional domains, and etymologies.

István Vásáry devotes an obituary to András J. E. Bodrogligeti (1925–2017).

The editors are pleased to announce that, starting from volume 21 (2017), *TURKIC LANGUAGES* is indexed in Web of Science’s Arts and Humanities Citation Index (AHCI).

*Lars Johanson*

## Obituary

### In memoriam András J. E. Bodrogligeti (1925–2017)

István Vásáry

Vásáry, István 2018. In memoriam András J. E. Bodrogligeti (1925–2017). *Turkic Languages* 22, 151–154.

*István Vásáry, Turkic and Central Asian Studies, Eötvös Loránd Tudományegyetem, Múzeum krt. 4/d, HU-1088 Budapest, Hungary. E-mail: vasaryi@gmail.com*



December 2017, Professor András J. E. Bodrogligeti passed away at the age of 92. Since 1972 he had been Professor of Turkic and Iranian Studies at the University of California, Los Angeles, and with his death one of the outstanding scholars of Turcology left us.

His long and fruitful life began in 1925 when he was born in Tiszaigar, Hungary. After the elementary school he moved to Eger, an old town of north-eastern Hungary, and a centre of a Catholic Archdiocese. (Eger was among the first ten dioceses founded in 1001–1009 by St. Stephen I, King of Hungary.) He must have imbibed the historical atmosphere of the town, which prides itself on having a castle that in



1552 resisted the Ottoman siege for a long time, halting the Turks for a few decades. In the wake of the romantic historical novel “Stars of Eger” (1901) written by the Hungarian author Géza Gárdonyi, the siege of Eger has for every Hungarian school-boy become part of the national mythology.

During 1940–1943 he attended the Archiepiscopal Lyceum in Eger, then during 1943–1947 the Archiepiscopal Teachers’ Training Institute. Between 1947 and 1951 he continued his studies in Budapest at the Péter Pázmány University (after the forceful communist takeover, in 1949 renamed Loránd Eötvös University—ELTE). He obtained his MA in Indo-European Studies with a specialisation in English and German. During 1953–1955 he earned a new BA from ELTE in Oriental Studies with a specialisation in Iranian, Turkic and Arabic, and in 1958 he graduated from the same university in Iranian and Turkic languages and literatures, with a thesis entitled “Sayf-i Sarayi’s Turkic Translation of Sa’di’s Gulistan”. Later, in 1963, he became a so-called “candidate of sciences” at the Hungarian Academy of Sciences, with his thesis “Preliminary Studies to the Exploration of Turkic and Iranian Linguistic Relations”. In 1971 he obtained the Doctoral Degree of the Hungarian Academy of Sciences with his thesis “The Persian Vocabulary of the Codex Cumanicus”. Needless to say he received all of his degrees with honours. His knowledge of and proficiency in languages were amazing; he was a real polyglot. In addition to his mother tongue Hungarian, he spoke and wrote in English, German, French, Italian, Russian, Persian and Turkic and knew a couple of other languages quite well.

In parallel with his academic work and achievements he occupied and made progress in different academic positions. During 1961–1964 he was assistant professor, and during 1965–1970 associate professor of Persian and Turkish at ELTE. Meanwhile, during 1964–1965 he taught courses in Modern Turkish Literature in Ankara, and during 1967–1968 at UCLA. His contact with UCLA became ever stronger because of his paternal friend Professor János Eckmann, who from 1961 was first a visiting then an acting associate professor at the Department of Near Eastern Languages at UCLA. In 1966 Eckmann became a full professor of Turkic Studies with a specialisation in Eastern Turkic language and literature. Bodrogligeti joined a common project on Chagatay language led by Professor Eckmann. In 1971, after Eckmann’s premature death, Bodrogligeti stood at a crossroads. The Hungarian authorities did not prolong his stay in the USA, and his choice was to stay. Within a year, in 1972 he was appointed Professor of Turkic and Iranian Studies at UCLA.

Now the second period of his career began, and this second part was also successful and productive. He was an active member of the university faculty who won a number of major academic grants for his research, among others for the projects “Islam Among the Turks of Central Asia”; “The Complete Works of Ahmad Yasavi”; “The Chagatay Language”; “An Academic Reference Grammar of Modern Literary Uzbek”; and “The State of Islam in the Ferghana Valley after the First Decade of the Post-Soviet Era”. The grants were well used, and upon their conclusion valuable monographs and publications were produced. András Bodrogligeti was a

member of several scholarly associations (American Oriental Society, Middle East Studies Association, Society for Iranian Studies) and was an Honorary Member of *Türk Dil Kurumu* [Turkish Language Society], Ankara. Between 1979 and 1993 he was Director of the John D. Soper Central Asian Language Institute, and from 1994 onward he presided over the Association for Central Asian Studies, USA. During 1979–1993 he was co-editor of the *Ural-Altische Jahrbücher* (Bloomington), and during 1993–1998 of its continuation, the *Eurasian Studies Yearbook* (Los Angeles & Bloomington, Indiana).

András Bodrogligeti was one of the outstanding figures of international research into the Central Asian Turkic language and literature. He became a worthy successor to János Eckmann's legacy, and the works of these two scholars of Hungarian descent made UCLA one of the leading centres in the field of Central Asian Turcology. His scholarly work embraces a wide range of themes of Middle Turkic philology. Before dedicating himself totally to the Central Asian Turkic language and literature, he dealt with the two pre-Chaghatay 14th-century monuments in his first books: *A Fourteenth-Century Turkic Translation of Sa'dī's Gulistān*, Budapest, 1969; *The Persian Vocabulary of the Codex Cumanicus*, Budapest, 1971. Both works are laudable presentations of important monuments. His virtue of making reliable and precise philological and linguistic analyses, amply demonstrated in his subsequent books and articles, is already present in these works.

During his professorship at UCLA he devoted himself mainly to Central Asian Turcology. He published a number of hitherto unedited Chaghatay works in an exemplary way. Suffice it to mention a few of them: *Hāliş's Story of Ibrāhim. A Central Asian Islamic Work in Late Chaghatay Turkic. Edited with an introduction, a translation, and a glossary*, Leiden, 1975; "Muḥammad Shaybāni's Baḥru'l-Huda'. An Early Sixteenth Century Didactic Qasīda in Chaghatay", *UAJb* 54 (1982), pp. 1–56; "Bābur Shāh's Chaghatay Version of the *Risāla-i Vālidīya*: A Central Asian Turkic Treatise on How to Emulate the Prophet Muḥammad", *UAJb* 56 (1984), pp. 1–61.

But the crown of his lifelong effort to better understand and describe the Central Asian Turkic literary language and its closest continuation, Modern literary Uzbek, are two monumental monographs that will long be the most detailed and standard works on the topic: *A Grammar of Chaghatay* (Languages of the World / Materials 155.) Munich: LINCOM Europa, 2001; and *An Academic Reference Grammar of Modern Literary Uzbek*. (Lincom Studies in Asian Linguistics 50) Munich: LINCOM Europa, 2003, vols 1–2. As one of his reviewers, Stefan Georg, put it: the latter is a "comprehensive and useful grammar, which will remain the Uzbek grammar of choice for linguists, Turcologists and students of Central Asian languages and cultures" (*Language* 82/4, 2006, p. 938).

The writer of this obituary had the privilege, as a student and young scholar, to personally know and learn from András Bodrogligeti. He was a very resolute man who always knew what he wanted to do and study. He was also a dedicated teacher ready to share his enormous knowledge with all his students. I owe much of my classical Persian knowledge to his practised expertise. Though he left a rich legacy

behind, his sorrowful demise is an irreparable loss for Turcology and Turcologists, especially for those of us who personally knew this brilliant scholar.

# NP versus DP: Which one fits Turkish nominal phrases better?

**Jaklin Kornfilt**

Kornfilt, Jaklin 2018. NP versus DP: Which one fits Turkish nominal phrases better? *Turkic Languages* 22, 155–166.

An influential proposal offered and defended in a series of studies (e.g. Bošković 2008, 2012, 2013) and joint work with others, is Bošković's suggestion that there exists a typological divide between languages whose "traditional" noun phrases (*NPs*) are actually determiner phrases (*DPs*), such as English, and languages whose noun phrases extend only to *NP*. In a study related to this body of work, Bošković & Şener (2014) claim that Turkish is an *NP*-language, and that it therefore exhibits the properties which Bošković's system would ascribe to it. They further posit a structure of the *NP* from which (at least some of) the relevant properties of Turkish would follow. Kornfilt (2017) has addressed some of these properties, illustrating the lack of the claimed alignment between them and the assumed *DP*- versus *NP*-nature of English, German, and Turkish, respectively. The present paper aims to address the more interesting part of Bošković & Şener (2014), namely the part where certain additional (claimed) properties of the Turkish "traditional *NP*" are *made to follow* from an *NP*-structure (rather than simply positing a correlation between certain properties and a claimed status as an *NP*-language, as had been done in the parts of their paper which had been addressed in Kornfilt 2017). I show that, at best, the relevant properties can be explained in other ways, and, at worst, that the "*NP*-hypothesis" actually makes wrong predictions with respect to Turkish. My main objective here is not to insist that Turkish is a *DP*-language, but rather to show that the phenomena claimed by Bošković & Şener to characterize Turkish as an *NP*-language do not succeed in establishing such a characterization.

Keywords: Lexical projections, functional projections, *DP*, *NP*, definite article, indefinite article, determiner, *DP*-languages, *NP*-languages

*Jaklin Kornfilt, Syracuse University, Department of Languages, Literatures and Linguistics, 340 HBC, Syracuse, NY 13244-1160, U.S.A.*

## 1. Introduction

Since the early 1980s, theoretical approaches to syntax have tended to transpose functional morphological elements into syntactic heads of functional phrase-structure projections. Thus, clauses are tense phrases (*TPs*), and if they are introduced by

a complementizer, they are also complementizer phrases (*CPs*).<sup>1</sup> For Turkish in particular, Kornfilt (1984) proposed an agreement phrase (*AgrP*) for possessive noun phrases. Abney (1987) extended that proposal to all noun phrases, with suggested cross-linguistic validity, and with a focus on English. Instead of an *AgrP*, Abney proposed a determiner phrase (*DP*), with the determiner as the head of the phrase, taking a noun phrase (*NP*) as its complement. For Turkish, the *AgrP* would be translated into a *DP*, with the *Agr* head now in the position of the *D*-head. This would mean that an actual determiner, such as a demonstrative<sup>2</sup>, would not be a head, but rather a specifier associated with a head; this would follow from the fact that Turkish is, in general, head-final rather than head-initial.

Not all languages necessarily have determiners, however, and not all languages need have (morphological) tense or complementizers. What should our analysis be of clauses or noun phrases in such languages? Should we still posit *functional* projections (and thus propose that these projections have empty heads), or should we adopt a more concrete approach, analyzing such phrases as projections of the highest *lexical* head?

In a series of studies (e.g. Bošković 2008, 2012, 2013) and joint work with others, Bošković proposes and defends the interesting suggestion that there exists a typological divide between languages whose “traditional” noun phrases (*NPs*) are actually determiner phrases (*DPs*), such as English, and languages whose noun phrases extend only to *NP*. This would challenge, among other things, the proposal in Abney mentioned above, according to which all languages have *DPs*. Bošković further proposes that one immediate clue for the relevant type of a language in this respect would be whether it has articles; in other words, having demonstratives would not be relevant in this context. More interestingly, Bošković proposes additional properties which a language would or would not exhibit, depending on whether it is an “*NP*-” or a “*DP*-” language. For example, “*NP*-languages” are supposed to lack *Neg-Raising* (which “*DP*-languages” clearly have). They are also supposed to exhibit “pro-drop” without a local identifying type of morphology and also to allow long-distance scrambling—both of these being properties which *DP*-languages are supposed to lack.

In a study related to this body of work, Bošković & Şener (2014) claim that Turkish is an *NP*-language, and that it therefore exhibits the properties which Bošković’s system would ascribe to it. They further posit a structure of the *NP* from which (at least some of) the relevant properties of Turkish would follow.

1 Whether a clause not overtly introduced by a complementizer is or is not also a *CP* is a matter of controversy. This issue will not have any bearing on our concerns in this paper.

2 I am not taking a stand here on whether a demonstrative is a regular determiner, positioned in the *D*-head of a *DP* in a language such as English, or whether it heads its own projection, i.e. a *DemP*. For Turkish, given the language’s head-final nature, a demonstrative would be the specifier of either a regular *DP* or of a *DemP*.

In one part of their paper, Bošković & Şener list some properties of the kind just mentioned, claiming that Turkish exhibits those reserved for *NP*-languages. They do not, however, discuss how these properties would follow from the claimed *NP*-nature of Turkish. Kornfilt (2017) addresses these properties, demonstrating the lack of the claimed alignment between these properties and the assumed *DP*- versus *NP*-nature of English, German, and Turkish, respectively. More specifically, the six phenomena addressed in that work are: *Neg*-Raising (i.e. the claim that *NP*-languages disallow clause-mate *NPI* (Negative Polarity Item) licensing under *Neg*-Raising, while *DP*-languages allow it); the claim that only *DP*-languages allow the majority superlative reading (in ambiguous utterances such as *At the party, most people drank beer*); the claim that *NP*-languages disallow transitive nominals with two lexical genitives; the claim that only *NP*-languages may allow (long-distance) scrambling; the claim that inverse scope is unavailable in *NP*-languages (“in some examples”, as phrased by Bošković & Şener 2014: 109); and the claim that “radical pro-drop” (i.e. “the productive pro-drop of subjects and objects in the absence of rich verbal agreement”; Bošković & Şener 2014:105) is possible only in *NP*-languages. The discussion addresses Turkish, as a representative of supposedly *NP*-languages, and, with respect to most of the claims just listed, also German, as a representative of *DP*-languages. The aim of Kornfilt (2017) was not to claim that Turkish is a *DP*-language, but rather to show that the typology based on a dichotomy of the world’s languages between *NP*-languages and *DP*-languages does not really work, in the sense that the phenomena and criteria proposed in the studies by Bošković, by Bošković & Şener, and in other related papers make, at least in part, wrong predictions concerning the languages they address. Thus, it was important to show that not only does Turkish pose challenges and problems to a characterization of the language as an *NP*-language, but that German, too, exhibits properties that are problematic for the claim that it is a *DP*-language. If such a typology makes sense at all, then other criteria should be found and tested with respect to both supposed types of languages, to see whether such criteria are sound.

My aim here is to address the more interesting part of Bošković & Şener (2014), namely the part where certain additional (claimed) properties of the Turkish “traditional *NP*” are *made to follow* from an *NP*-structure (rather than the authors simply positing a correlation between certain properties and a claimed status as an *NP*-language, as was done in the first part of their paper, and which was addressed in Kornfilt 2017), as opposed to a *DP*-structure. I shall show that, at best, the relevant properties can be explained in other ways, and, at worst, that the “*NP*-hypothesis” actually makes wrong predictions with respect to Turkish. Once again, my main objective here is not to insist that Turkish is a *DP*-language, but rather to demonstrate that the phenomena claimed by Bošković & Şener to characterize Turkish as an *NP*-language do not succeed in justifying such a characterization.<sup>3</sup>

3 Arslan-Kechriotis (2009) takes a more forceful stand in this respect, claiming that Turkish is a *DP*-language. The main phenomena on which her claim is based are as follows:

Before doing so, I would like to briefly discuss the issue of articles as one of the claimed central characteristics of “NP-languages”.

## 2. The issue of articles in the NP/DP typology

One question briefly mentioned in the introduction concerns determiners other than articles: Why shouldn’t other determiners (e.g. demonstratives) qualify as *D* (or, in a head-final language such as Turkish, as the specifier of a *DP*), as well? Yet another question I would like to raise is: What if a language has one type of article but not another?

Turkish has demonstratives, but those are not accepted as a genuine *D* by Bošković in general. Furthermore, Turkish has an indefinite article. However, Bošković & Şener (2014) analyze the morpheme in question, i.e. *bir*, as the numeral *one* in *all* its occurrences, and not as the indefinite article *a*.

I will mention here just the traditional reason for viewing certain usages of *bir* as an indefinite article, in contrast to its usages as a numeral; Kornfilt (2017) offers additional reasons and criticises Bošković & Şener’s view of this morpheme as a numeral exclusively.

The syntactic distribution of *bir* as an article differs from its distribution when it is a numeral; this is clear with respect to its placement with respect to adjectives; e.g.

- (1) *bir/üç*      *çürük*      *elma*  
       *one/three*    *rotten*      *apple*  
       ‘*one* rotten apple/three rotten apples’
- (2) *çürük*    *bir*      *elma*  
       *rotten*    *a*      *apple*  
       ‘*a* rotten apple’

These are the *unmarked* orders of the respective examples; numerals precede adjectival modifiers, while the indefinite article follows them. It is true that it is possible to override these unmarked correlations with intonation; e.g. in (1), de-accenting the numeral *bir* ‘one’ makes it likelier for it to be interpreted as the indefinite article, and in (2), stressing the article makes it likelier for it to be interpreted as the numeral. Nonetheless, the unmarked orders and intonation contours yield the meanings as stated above.<sup>4</sup> I therefore suggest that Turkish is not an article-less language; its

Turkish shows island-effects (with respect to *wh*-islands, complex *NPs*, and sentential subjects) in the application of scrambling—something that *NP*-languages are not supposed to exhibit; Turkish further exhibits, as she shows, instances of the Left Branch Condition, which is also supposed to be a property of *DP*-languages rather than of *NP*-languages.

4 The unmarked order within the Turkish *DP/AgrP* is as follows:

traditional characterization as a language that has an indefinite article<sup>5</sup> but not a definite article is correct.<sup>6</sup>

### 3. Bošković & Şener's proposed structure for Turkish "traditional NPs" and its implications for the DP/NP dichotomy

#### 3.1. Co-indexation or lack thereof between names and pronouns in "possessive phrases"

Bošković (2012) treats Serbo-Croatian (SC) as an NP-language (it has no articles), and treats SC possessors and demonstratives as NP-adjuncts. One of his arguments, noted in Despić (2011, 2013), is based on possible and impossible co-indexations of pronouns and names—possible in English (see 3) and not in SC (see 4):

- (3) a. His<sub>i</sub> latest movie really disappointed Tarantino<sub>i</sub>.  
 b. Tarantino<sub>i</sub>'s latest movie really disappointed him<sub>i</sub>. (Bošković & Şener 2014: 111, (23))
- (4) a. \*[<sub>NP</sub> Kusturicin<sub>i</sub> [<sub>NP</sub> najnoviji film]] ga<sub>i</sub> je zaista razočarao.  
 Kusturica's latest movie him is really disappointed  
 Intended: 'Kusturica<sub>i</sub>'s latest movie really disappointed him<sub>i</sub>.'  
 (Bošković & Şener 2014: 111, (24a)) (OK in English.)

Possessor > relative clause > demonstrative > numeral/quantifier > adjective (phrase) > indefinite article > head noun

- 5 Just like in the earlier discussion about the demonstratives in Turkish, defenders of the traditional view of *bir* as an indefinite article in the positions and usages sketched in the main body of the text will have to say that if Turkish nominal phrases that include *bir* are DPs (as I claim, agreeing with Arslan-Kechriotis (2009), who denies article-status to *bir*, but nonetheless assumes a DP-analysis of nominal phrases with *bir*), and if all projections in Turkish are head-final, then the D-head of such nominal phrases will have to be empty, and *bir* will have to be in the specifier position of such a DP (thus agreeing with the head, which would host features such as non-specificity) under specifier-head agreement. This is also the analysis assumed for such nominal phrases by Arslan-Kechriotis (2009).
- 6 Öztürk (2005) mentions a typological observation, based on statistics and on information in Crisma (1999) and Longobardi (2001): If a language has only one article, this will always be the definite article rather than the indefinite one. Kornfilt (2017) points out that this type of reasoning is circular. Even if more languages were to be discovered which have only an indefinite article and no definite article (such as Turkish), defenders of this typological bias would dismiss them, claiming that the "apparent" indefinite article cannot possibly be an article but rather has to be a numeral. Furthermore, as typological research is advancing, more languages are being discovered to have just this kind of situation, namely a distinct use of the numeral with the meaning "one" as an indefinite article.



- b. \*<sub>[NP Njegov<sub>i</sub> [<sub>NP najnoviji film</sub>]] je zaista razočarao Kusturicu<sub>i</sub>.</sub>  
 his latest movie is really disappointed Kusturica  
 Intended: ‘His<sub>i</sub> latest movie really disappointed Kusturica<sub>i</sub>.’ (Bošković & Şener 2014: 111, (24b)) (OK in English.)

The difference in grammaticality between the English and SC examples is explained by the posited different structure. In SC, the possessor would be an *NP*-adjunct; given that SC is claimed to lack *DP*, the possessor c-commands the pronoun in (4a) and the referential expression in (4b) from within the “traditional *NP*”, leading to a Condition B violation in (4a), and to a Condition C violation in (4b).<sup>7</sup>

Bošković & Şener state that Turkish patterns with SC rather than with English in these respects:

- (5) a. \*<sub>[ Özpetek<sub>i</sub>-in film-i] on<sub>i</sub>-u hayal kırıklığına uğrat-tı.</sub>  
 Özpetek-GEN movie-3SG he-ACC disappoint-PAST  
 Intended: ‘Özpetek<sub>i</sub>’s movie disappointed him<sub>i</sub>.’  
 b. \*<sub>[ On<sub>i</sub>-un film-i] Özpetek<sub>i</sub>-i hayal kırıklığına uğrat-tı.</sub>  
 he-GEN movie-3SG Özpetek<sub>i</sub>-ACC disappoint-PAST  
 Intended: ‘His movie disappointed Özpetek<sub>i</sub>.’ (Bošković & Şener 2014: 111, (25))

I claim that the facts in (5) have nothing to do with Binding Theory and with c-command, but rather with the particular pronouns used. In (5a), many speakers do get co-indexation. If the personal pronoun *o(n)* is replaced with the logophor *kendisi*, all speakers consulted got co-indexation.<sup>8</sup>

In (5b), we have the (soft) principle (cf. Chomsky 1981, Kornfilt 1987) *Avoid Pronoun at work*; the example is perfect with “small *pro*”, i.e. a silent pronominal element:

- (5) c. [<sub>pro<sub>i</sub></sub> (*kendi<sub>i</sub>*) film-i] Özpetek<sub>i</sub>-i hayal kırıklığına uğrat -tı.  
 pro own movie-3SG Özpetek<sub>i</sub>-ACC disappoint-PAST  
 ‘His<sub>i</sub> (own<sub>i</sub>) movie disappointed Özpetek<sub>i</sub>.’

7 Interestingly, Lyutikova & Pereltsvaig (2015) show that in Russian, also an articleless (and Slavic) language, examples such as those in (4) are well-formed (see their examples in (6), p. 293). Their claim is that most, if not all, of the phenomena used in Bošković & Şener to distinguish between *NP*- and *DP*-languages cannot be applied to Russian so as to characterize it as an *NP*-language.

8 For a detailed study of the distribution of the “regular” personal pronouns versus their logophoric counterparts, see Özsoy (1992). In that study, “small *pro*” is not included. Here, my main concern is the difference between any overt pronoun and “small *pro*”, rather than differences between the two types of overt pronouns.

(5c) is a version of (5b), whereby the overt personal pronoun is replaced with small *pro*, and it is perfectly fine.

Bošković & Şener do note this fact (but in a later section), claiming that possessors start very low within the NP (lower than adjectives, for example), but that they move to a higher, adjoined position if overt. If covert, i.e. if the possessor is *pro*, they remain in that low base position. Thus, *pro* is too low to c-command beyond the NP.

This analysis is less than convincing, and is in fact problematic, for a number of reasons:

1. Ascribing syntactic differences to phonology is an unwelcome direction to take. Such a direction should be avoided as much as possible, and certainly if it is not necessary. In particular, arguments in favor of positing silent, but syntactically present, elements are usually based on common syntax between the silent element and its overt counterpart.

2. Motivations for the obligatory nature of either raising overt possessors or the non-movement of silent possessors are lacking. In contrast, it is worth pointing out that in my proposal, *Avoid Pronoun* is not limited to possessors and applies in Null Subject languages in general.

3. “Light pronouns”, such as clitics, tend to move to positions higher than their “heavy” counterparts, i.e. full noun phrases. In Bošković & Şener’s analysis, the proposed positions are reversed, with the *pro*, an obviously super-light nominal, being crucially in a lower position than its heavy counterpart.<sup>9</sup>

I thus conclude that Bošković & Şener’s analysis of the binding phenomena is on the wrong track, and that those phenomena do not support the posited structure of Turkish nominal phrases as (necessarily) non-DP phrases, i.e. as phrases extending only to the NP-level. Instead, the account I propose, based on *Avoid Pronoun*, is more general, and it avoids the empirical problems that Bošković & Şener’s account faces with respect to differences between “light pronouns” and full nominal phrases. Note, however, that my account with respect to the binding facts just discussed does not differentiate between an NP- and a DP-analysis; it would hold with respect to either.

9 This argument might be seen as somewhat weakening the first argument in this paragraph, according to which phonologically based analyses should be excluded from syntax. Indeed, the phonological/phonetic and the syntactic aspects of language, and thus of linguistic analysis, do interface with each other, and phonological considerations cannot be excluded from syntactic accounts completely. However, there is a big difference between positing a distinction between “light” and “heavy” nominals (which is not only a phonological difference, but also a syntactic one), on the one hand, and stipulating a different syntactic position for *pro*, a phonologically empty pronoun, versus a phonologically overt nominal phrase, without any independent evidence—especially if such independent evidence, e.g. the positioning of light pronouns, goes directly against what is stipulated for the position of *pro*.

### 3.2 Word order within the “traditional NP” and its effects

Another syntactic property of nominal phrases which Bošković & Şener claim follows from the *NP*-analysis of such phrases is the order of subconstituents within nominal phrases. The authors note that either order of possessives and demonstratives is acceptable in Turkish “traditional *NPs*”, i.e. both *Poss* > *Dem*, and *Dem* > *Poss*. This, they claim, follows from their analysis. Given that both the possessor in a possessive nominal phrase, and the demonstrative are adjuncts to the *NP* (which they must be, given that Bošković & Şener deny the status of *DP* to the Turkish nominal phrase, and thus also deny positions to the possessor and to the demonstrative which are higher than the *NP*), they should be able to appear in either order, given that adjuncts typically can and generally do reorder:

- (6) *Can-in şu üç bisiklet-i*  
 Can-GEN that three bicycle-3SG  
 ‘those three bicycles of Can’s’ (Adapted from Bošković & Şener 2014: 112, (26a))
- (7) *şu Can-in üç bisiklet-i*  
 that Can-GEN three bicycle-3SG  
 ‘those three bicycles of Can’s’ (Adapted from Bošković & Şener 2014: 112, (27a))

Example (6) is indeed fine; this is the basic word order. Example (7) is also good, but is a marked order. Bošković & Şener do not mention this difference in markedness, which is problematic for them, given that, as I just mentioned, *Poss* and *Dem* are adjuncts for them, and so their order should be completely equivalent, which it is not. There are further differences, e.g. the demonstrative can be stressed in (6), but not in (7).

Bošković & Şener’s analysis is therefore untenable. Instead, I propose the following analysis: *Dem* is in the specifier of *DP*, and the possessor is in the specifier of an *AgrP* which dominates the *DP*—the same *AgrP* which I had originally proposed in Kornfilt (1984), as mentioned in the introduction.

Note that in (7), the *Dem* might indeed be adjoined, e.g. to *DP* or *PossP* (= *AgrP*). It is base-generated in *Spec, DP* in (51), and topicalizes, either by adjoining to *DP*, or by moving to an even higher *Spec, TopP*. But, in any event, treating the possessor and the demonstrative as equivalent adjuncts that can occur in either order is clearly incorrect.

### 3.3. Ellipsis within the “traditional NP”

Bošković & Şener analyze numerals and adjectives as *NP*-specifiers (necessitating the assumption of multiple *NP*-specifiers, given that numerals and adjectives can co-occur), and possessors as *NP*-adjuncts:

- (8) a.  $\begin{array}{c} \text{NP} \\ / \quad \backslash \\ \text{Num} \quad \text{N}' \end{array}$     b.  $\begin{array}{c} \text{NP} \\ / \quad \backslash \\ \text{Adj} \quad \text{N}' \end{array}$     c.  $\begin{array}{c} \text{NP} \\ / \quad \backslash \\ \text{Poss} \quad \text{NP} \end{array}$

They further claim that this analysis offers a “simple unified account” of a number of ellipsis facts, under the assumption that “only phrases (not segments or bar-level categories) can be elided”:

- (9) \**Pelin her gün [beş elma] ye-r, Can-sa [iki elma] ye-r.*  
P. (NOM) every day five apple eat-AOR C-‘however’ two eat-AOR  
‘Pelin eats five apples every day, while John eats two.’ (Bošković & Şener: p. 118, (41))
- (10) \**Pelin [eski kitap] sat-tı, Suzan-sa [yeni kitap] sat-tı.*  
P. (NOM) old book sell-PAST Suzan-‘however’ new sell-PAST  
‘Pelin sold old books, while Susan sold new ones.’ (Bošković & Şener 2014: 118, (42a))
- (11) \**[Pamuk-un kitab-ın]-ı oku-du-m, ama [Oe-nin kitab-ın]-ı oku-ma -dı-m.*  
P.-GEN book-3SG-ACC read-PAST-1SG but O.-GEN  
read-NEG-PAST-1SG  
‘I read Pamuk’s book, but I didn’t read Oe’s.’

However, there is an equally simple and at least as general way to account for these ill-formed attempts at applying ellipsis within the *NP/DP*, as proposed in von Heusinger & Kornfilt (2017) to explain ill-formed instances of ellipsis within partitive phrases:

- (12) Condition: An “overt nominal head” requirement for DPs:  
Nominal phrases need to have at least one nominal head (i.e. a head with the [+N] feature) which has to be filled overtly. (von Heusinger & Kornfilt 2017: (24))

This is proposed not only for possessive phrases, but also for *DPs* (i.e. “traditional *NPs*” in Bošković & Şener’s terminology) in general. Some relevant examples, taken from von Heusinger & Kornfilt (2017), follow:

- (13) a. \**[Meyve-ler-den üç elma] ye -dı -m.*  
fruit-PL-ABL three apple eat-PAST-1SG  
Intended: ‘I ate three (apples/pieces of fruit/items) of the fruit.’

Eliding the nominal head (here, *elma* ‘apple’) leads to ill-formedness, due to a violation of the condition in (12). The example can be salvaged in a variety of ways, e.g.

by using a classifier with the feature [+N], i.e. *tane* ‘item, entity’, or by using a “dummy”, default 3SG agreement element with a clitic-like function:

- (13) b. [*Meyve-ler-den üç tane*] *ye-di-m.*  
 fruit-PL-ABL three item eat-PAST-1SG  
 ‘I ate three items/pieces of the fruit.’
- c. [*Meyve-ler-den üç-ün-ü*] *ye-di-m.*  
 fruit-PL-ABL three-3SG-ACC eat-PAST-1SG  
 ‘I ate three (entities) of the fruit.’

Other issues brought up in Bošković & Şener (2014) concerning the structure of the “traditional *NP*” in Turkish are likewise amenable to other, and at least as general, analyses and conditions, such as *Avoid Pronoun*, and the condition in (12), as we have seen.

#### 4. Conclusions and further questions

The discussion in this paper has shown that the proposal to regard the Turkish nominal phrase as only as low as an *NP*, rather than as high as a *DP*, faces empirical problems, and that analyzing such phrases as *DPs* is at least as successful in capturing their syntactic properties, and in fact is more successful with respect to word order within the nominal phrase.

If one were to remain within an approach which categorizes languages as to “*NP*–” versus “*DP*–languages”, Turkish would probably best be viewed as a *DP*–language which has one article. However, the cross-linguistic approach based on a *NP*–*DP* typology is not attractive anyway, given that such a typology faces problems not only due to the supposed *NP*–nature of Turkish, but also because clear-cut *DP*–languages such as German and English challenge the correlations posited in such a typology.<sup>10</sup> Following the conclusions in Kornfilt (2017), I therefore propose abandoning the kind of typological endeavor suggested by Bošković & Şener, and examining each language individually as to whether it does or does not have *DPs* (and it may turn out that all do), independently from whether or not it has articles.<sup>11</sup> In any event, the most attractive aspect of Abney’s *DP*–hypothesis for nominal phrases is its conceptual character which allows us to posit parallels between nominal and clausal domains. The burden of proof for “small nominal phrases” which

10 The reader is referred to Kornfilt (2017) for examples and discussion of problems for the *NP*–*DP* typology posed by *DP*–languages such as German and English.

11 Lyutikova & Pereltsvaig (2015), which has come to my attention too late to be included in this discussion in any detail, reject the “Bošković typology”, as well. While allowing for some *NPs* without a higher *DP* projection in some constructions, in some languages, they reject a “parameterized *DP* Hypothesis”. Giusti & Iovino (2016) arrive at similar conclusions, based on Latin.

only extend to the highest lexical projection, without involving functional projections, thus lies with advocates of such smaller phrases, especially if an *entire* language is to be characterized as an “NP-language”.

### Acknowledgments

I thank Akgül Baylav, Demir Dinç, Alp Otman and Ayşe Yazgan for Turkish judgments (as well as to Deniz Tat and Güliz Güneş for further confirmation), and Artemis Alexiadou, Josef Bayer, Anna Cardinaletti, Guglielmo Cinque, Peter Cole, Giuliana Giusti, Martin Haspelmath, Klaus von Heusinger, Bill McClure, Tim McKinnon, Volker Struckmeier, Florian Schaefer, and Sasha Vovin for discussion. I am also grateful to two anonymous reviewers for their questions, comments, and additional references. I further thank Lars Johanson for having been such an important guiding light in Turkological studies, and Éva Á. Csató for her encouragement. The usual disclaimers apply.

### Abbreviations

[+N] positive value for the nominal feature	GEN Genitive
1SG first person singular	NEG Negative (element), negation
3SG third person singular	NOM Nominative
ABL Ablative	NP Noun phrase
ACC Accusative	PL Plural
AgrP Agreement phrase	Poss Possessive
AOR Aorist	PossP Possessive phrase
CP Complementizer phrase	pro A phonologically unrealized pronoun
Dem Demonstrative	Spec Specifier
DemP Demonstrative phrase	TopP
DP Determiner phrase	TP Tense Phrase

### References

- Abney, S. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT.
- Arslan-Kechriotis, Z. C. 2009. *Determiner phrase and case in Turkish: A Minimalist account*. Saarbrücken: VDM Publishing House.
- Bošković, Z. 2008. What will you have, DP or NP? In: Elfner, E. & M. Walkow (eds.) *Proceedings of the Northeastern Linguistics Society (NELS) 37*. Amherst, MA: GSLA Publications. 101–114.
- Bošković, Z. 2012. On NPs and Clauses. In: Grewendorf, G. & Zimmermann, T. E. (eds.) *Discourse and grammar: From sentence types to lexical categories*. Berlin: Mouton de Gruyter. 179–242.
- Bošković, Z. 2013. Phases beyond clauses. In: Schürcks, L. & Giannakidou, A. & Etxebarria, U. & Kosta, P. (eds.) *The nominal structure in Slavic and beyond*. Berlin: Walter de Gruyter. 75–128.

- Bošković, Z. & Şener, S. 2014. The Turkish NP. In: Hofherr, P. C. & Zribi-Hertz, A. (eds.) *Crosslinguistic studies on noun phrase structure and reference*. Leiden: Brill. 102–140.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Crisma, P. 1999. Nominals without the article in the Germanic languages. *Rivista di Grammatica Generativa* 24, 105–125.
- Despić, M. 2011. Syntax in the absence of determiner phrase. Doctoral dissertation, University of Connecticut.
- Despić, M. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44, 239–270.
- Giusti, G. & Iovino, R. 2016. Latin as a Split-DP language. *Studia Linguistica* 70: 3, 221–249.
- von Heusinger, K. & Kornfilt, J. 2017. Partitivity and case marking in Turkish and related languages. *Glossa* 2(1): 20. Falco, M. & Zamparelli, R. (eds.) *A special issue on partitivity*, 1–40. DOI: <https://doi.org/10.5334/gjgl.112>.
- Kornfilt, J. 1984. Case marking, agreement, and empty categoris in Turkish. Unpublished doctoral dissertation, Harvard University.
- Kornfilt, J. 1987. Beyond binding conditions: The case of Turkish. In: Boeschoten, H. & Verhoeven, L. (eds.) *Studies on Modern Turkish*. Tilburg: Tilburg University Press. 105–120.
- Kornfilt, J. 2017. DP versus NP: A cross-linguistic typology? In: Vovin, A. & McClure, W. (eds.) *Studies in historical and synchronic Altaic: Festschrift presented to John B. Whitman*. Leiden: Brill. 138–158.
- Keskin, C. 2009. Subject agreement dependency of accusative case in Turkish, or Jump-starting grammatical machinery. Doctoral dissertation. Utrecht University.
- Longobardi, G. 2001. The structure of DPs: Some principles, parameters and problems. In: Baltin, M. & Collins, C. (eds.) *The handbook of contemporary syntactic theory*. Oxford: Blackwell. 562–601.
- Ljutikova, E. & Pereltsvaig, A. 2015. The Tatar DP. *The Canadian Journal of Linguistics* 60: 3, 289–325.
- Özsoy, A. S. 1992. Türkçe’de eşgönderge örüntüsü [Coreference patterns in Turkish]. *Dilbilim Araştırmaları*. Ankara: Hitit Yayınevi. 77–85.
- Öztürk, B. 2005. *Case, referentiality and phrase structure*. Amsterdam & Philadelphia: John Benjamins.

# The educational role of the Late Medieval Persian-Ottoman Turkish bilingual dictionaries. The codices of the Matenadaran

Hasmik Kirakosyan & Ani Sargsyan

Kirakosyan, Hasmik & Sargsyan, Ani 2018. The educational role of the Late Medieval Persian-Ottoman Turkish bilingual dictionaries. The codices of the Matenadaran. *Turkic Languages* 22, 167–174.

Medieval Persian-Ottoman Turkish bilingual dictionaries (*farhangs*) were created as cultural and educational tools and were educational, instructive, and translating types of works. These dictionaries were special tools in the medieval educational system with which not only the Persian language was taught, but also the Persian and Arabic loan vocabulary in Turkish. Primarily, the medieval bilingual dictionaries aimed to help users read and understand the patterns of Persian classical literature, but they also assisted in teaching Persian grammatical categories that are relevant for the grammar of Ottoman Turkish. The examination of the medieval bilingual dictionaries kept in the Matenadaran support this theory.

Keywords: Persian-Ottoman Turkish dictionaries, *farhang*, Matenadaran

Assoc. Prof. Hasmik Kirakosyan, Institute of Ancient Manuscripts-Matenadaran, Yerevan State University, Department of Iranian Studies. 1 Alex Manoogian, 0025, Yerevan, Armenia. E-Mail: [hasmik.kirakosian@ysu.am](mailto:hasmik.kirakosian@ysu.am)

Res. Ass. Ani Sargsyan, Asien-Afrika - Institut, University of Hamburg, Edmund-Siemers-Allee 1, Flügel Ost 20146. Hamburg, Germany. E-mail: [ani.sargsyan@uni-hamburg.de](mailto:ani.sargsyan@uni-hamburg.de)

## 1. Introduction

The Persian-Ottoman Turkish bilingual dictionaries (*farhangs*) were compiled by men of letters to be used as cultural and educational tools for instruction, translation and explanation as early as the end of the 13th century (Öz 2016: 49–50; Yavuzarslan 2009: 12–14). These works were special tools in the medieval educational system for teaching the Persian language and through it Persian poetry. The later (16–17th c.) glossaries, in particular the *Daqā'iq al-ḥaqā'iq* (16th c.) (see below) and especially those works that incorporated grammatical instruction, began the process of creation of Turkish grammatical terms and contributed to the emergence of scientific language.<sup>1</sup>

1 See the use and improvement of grammatical terms in Turkish in the material of the glossary *Daqā'iq al-ḥaqā'iq* (16th c.) in an article in preparation.



Knowledge of Persian was the essential and one may even say, the necessary condition of Turkish medieval instruction and literacy. Old Anatolian Turkish passed to the level of a literary language under the influence of Persian. Beginning with the period of the Samanid dynasty, New Persian became a language of regional communication, a *lingua franca* within a certain territory of Islamic culture. It served as special means of cultural communication up until the 19th century. The influence of Persian language and poetry on Turkish language and literature chronologically coincides with the period of the Persian Rebirth, i.e. the 11th century, while the 13th century is known as the period of cultural “Iranization” in the history of the Turkic states of Asia Minor. Ottoman Turkish was so saturated with Persian that Fragner defined it as “a daughter language of modern Persian in its structural aspect”.<sup>2</sup>

In the 16th century, Ottoman Turkish was the administrative language of the Ottoman Empire, a means of literary and conceptual expression, but Persian language and poetry remained in their foremost positions.<sup>3</sup> The process of teaching Persian continued and rapidly developed, as did the propagation and appreciation of Persian poetry, with which Persian-Ottoman Turkish lexicography was strongly connected. In the 16th–17th centuries, within the territory of Ottoman Empire members of the Turkish population did not master Persian, or knew it only on the level of spoken language, and thus were ignorant about the literary language and did not understand literary works. In order to study Persian literature beginners had to learn Persian, which was also considered to be desirable for serving in the Ottoman court.

Thus, *farhangs*, Persian-Ottoman Turkish bilingual dictionaries, were needed to understand Persian poetry and to correctly speak and write Persian and Ottoman Turkish. That was why *farhangs* were used in medieval schools (*medrese*), where they also served as manuals for creative, educated and scientific persons. Beginners especially needed to master the lexical and grammatical nuances of Persian to correctly understand Persian poetry and use it in their works. A huge Persian vocabulary and Persian methods of word composition were introduced into Ottoman Turkish, thereby requiring the development of philological works included within Turkish bilingual dictionaries.

Persian-Ottoman Turkish dictionaries were mainly produced in the Ottoman Empire. Between the early period, when the first “*Şiḥāḥ ul-‘Ajam*” dictionary<sup>4</sup> was composed in 1279, and the publication of the Persian-Turkish Dictionary (*Farhang-e fārsī-torkī*) by İbrahim Olgun and Cemşid Draḥşan in 1967, two-third of such dictionaries were written in Turkey (Sāme‘ī 1996: 395). The bilingual dictionaries of the Matenadaran presented in this article were written on the territory of Modern Turkey.

2 For this and further information see Fragner (2006: 39–48).

3 See also Csató et al. (2016: 1).

4 On the first Persian-Ottoman Turkish dictionary, with a section devoted to Persian word composition in Arabic and Turkish, see Turan (2014: 431–444).

## 2. The methodology of writing bilingual dictionaries

Medieval Persian-Ottoman Turkish dictionaries were written according to the principles of Persian lexicography and dictionary-writing, which were already formulated and had rich traditions (Baevskii 2007: 151). Medieval Persian dictionaries had methodological peculiarities; they served not only as explanatory and encyclopedic tools for translation, but also as dictionaries of proper names, synonyms, antonyms, and metrical versification. The demand for these works in medieval society was largely driven by their use as directories, educational manuals, and commentaries on literary works. *Farhangs* were, with rare exceptions, not limited to the definitions of word, but also cited instances of their use in literary works. They are considered to be unique literary anthologies.<sup>5</sup>

In bilingual educational dictionaries it was also usual to categorize words by their part of speech: noun (*asāmī*) and infinitive (*maṣādīr*). Dictionaries which included both explanatory lists of words and grammatical explanations were called universal (*jāme'*).

The inclusion of grammatical material in *farhangs* is supporting evidence of their educational significance.<sup>6</sup>

In Persian-Ottoman Turkish *farhangs*, such materials were included in a different way: in separate sections or in word-articles. In the grammatical parts, authors also discussed problems of the spelling and pronunciation of entries. In the 16th century, the Persian or Arabic vocabulary already introduced in Ottoman Turkish was adapted to the Turkish pronunciation, i.e. underwent palatalization (Stein 2006: 144–145), as can be observed in the glossary of *Daqā'iq al-ḥaqā'iq* where the author explains the pronunciation of the words.

The prefaces of the medieval dictionaries contain words of praise addressed to God, *farhang* writing reasons and stimulus, the person who ordered the work or proposed its writing and why, as well as a list of *farhangs* that were used as sources in writing the work. In some cases, issues concerning the contents of the work are noted.

In the late Middle Ages, dictionaries were written in verse (*niṣāb*) and in prose. The dictionaries in verse had two aims: educational and poetic. They were mainly used in *medreses* to enable students to learn foreign words by heart and master versification and rhyming schemes.

5 In medieval *farhangs* it was traditional to base word explanations on examples borrowed from literary works. For instance, in his dictionary *Mi'yār-i jamālī*, Šams-i Faxrī gives examples of the words from his own verses; the author of *Šihāḥ ul-furs* takes examples from his father's works. The most frequently cited Persian classical authors in dictionaries are Neẓāmī, Rudakī, Ferdowsī, 'Unṣurī, Ḥāfeẓ, Sa'adī, etc. See Baevskii (2007: 127–134).

6 The inclusion of grammatical material in Persian *farhangs* begins in the 11th century. The first example is the *Farhang-i ja'farī* (1040/1630–1631).

Dictionaries in verse were poems called *maṣnavī* and they were divided in *qīṭa*'s and *baḥrs* (verse measures). The appearance of these dictionaries was conditioned by the goal of propagating Persian classical literature, of which the versified dictionaries were devoted to providing samples, commentaries, and explanations. The medieval book format did not allow for inscribing such explanations and commentaries in the books themselves and therefore there were separate volumes of dictionaries. Such Persian-Ottoman Turkish dictionaries, which explained or taught in verse the vocabulary of any Persian literary work, and showed the metrical versification rules, were also appreciated as philological works and had commentaries which are known as *ṣerḥ*. Alongside their educational significance, these works had great literary and aesthetic value, and were usually included in the same manuscript miscellanies (*meḡmū'a*) as literary works.

The short colophons of the copies of dictionaries are full of important biographical information about the scribe, the composer (sometimes about the renovator also), the circumstances of copying, renovating or buying the manuscript, as well as data about the time of its manuscript copying. Many copies of Persian-Turkish dictionaries lack such rich colophons, which may however be explained by their educational and teaching character. Often dictionaries were copied by learned persons, or teachers at *medreses* who wanted to have a copy of their own to use for teaching, and who did not think it necessary to leave their names in the colophons. This is why there are now numerous copies of medieval bilingual dictionaries kept in diverse manuscript depositories.

### 3. The bilingual Persian-Turkish dictionaries of the Matenadaran

The Persian-Turkish bilingual *farhangs* kept in the Yerevan Mesrop Mashtots Institute of Ancient Manuscripts (below MM) are copies from the 16th–18th centuries. They are witnesses of the cultural milieu of that time.

In terms of structure, these are works in prose and verse that are included in different volumes and miscellanies. Notes, additions and scribbling in the margins by the owners point to their practical significance, which at first was educational. Some of these manuscripts, especially those in verse, which are important from a literary and aesthetic viewpoint, were copied with literary ornaments, i.e. miniatures, vignettes and leather bindings. Most of these probably also had “commercial” value.

#### 3.1. Dictionaries in prose

3.1.1 The Miscellany N 133 from the collection of Arabic Script manuscripts at the MM is a copy of a well-known Persian-Turkish dictionary (written in 947/1540–1541) called *Luyat-i Ni'metullāh* (Dictionary of Ni'metullāh) by Ni'metullāh ibn Aḥmad b. Ġāzī al-Rūmī (date of death 969/1561–1562) (Lagarde 1884: 53–55,

Öz 2010: 156–160).<sup>7</sup> The colophon states that the scribe was Jabbar ibn Pīr Aḥmad al-Rāz Rūmī and the copy was made in 1105/1693. The colophon is at f.237 and the manuscript has 238 pages (Kostikyan 2017: 69). It is a dictionary in prose, written in one column with the script *nasx*. The preface occupies ff.3v–4v with the above-mentioned structure, and mentions the dictionaries *Uqnūm-i 'Ajam* (from which the copy was imitated), *Şihāh ul- 'Ajam*, and that by Luṭfullāh Ḥalīmī and Ḥasan ibn Ḥuseyn ibn 'Omād Qaraḥiṣārī (f.4r).

The dictionary *Luyat-i Ni'metullāh* includes the genres *asāmī* and *maṣādīr* of medieval dictionary writing, i.e. *jāme'*. The lemmata part is divided according to parts of speech intermediating by grammatical section. The part devoted to morphology, occupying ff.20r–26v and called *Qā'ede-i zabān-i fārsī* (Grammar of the Persian Language), contains explanations devoted to correct speech, word composition and morphology, based on examples in verse borrowed from literature. The words are ordered according to the following sequence of vowels: long *ā*, short *a*, *o*, *e*; divided into subheadings.

3.1.2 The Miscellany N 196 from the collection of Arabic Script manuscripts at the MM includes two works by Kemāl Pāšāzāde (d.1533): *Risāle-i yā-yi fī lisān al-fārsī* (Treatise about the Persian *ya*) in Turkish and the glossary *Daqā'iq al-ḥaqā'iq* (Details of the Truth) (Kostikyan 2017: 87–88).<sup>8</sup> The first work is not a dictionary, but is devoted to examining the word-composing *-ī* morpheme; in fact it demonstrates the importance of teaching and studying Persian in the 16th century.

The glossary *Daqā'iq al-ḥaqā'iq* is included in the miscellany on ff.27r–174v. The f.174r contains the colophon: “ ٩٧٢ سنه ”, which says that these philological works were copied in 1564/5 but does not mention the name of the scribe. This fact also supports an educational purpose of the dictionary because, in order to have their own working copies, teachers at the *medrese* frequently copied the manuscripts without mentioning their names and used these copies in the educational process.

The *Daqā'iq al-ḥaqā'iq* is a glossary of Persian synonyms and homonyms with Turkish explanations, and is based on *beyts* (double verse) and *qīṭa*'s cited from the Persian classical authors Neẓāmī, Ferdowsī, Ḥāfez, Sa'adī, and others (Lagarde 1884: 37–38, Öz 2010: 148–153). Its genre corresponds with that of Armenian dictionaries called *Barḵ ' govasanakank'* (Words of Praise), which examine laudatory words and expressions quoted from oratorical, poetical and highly artistic works. The *Daqā'iq al-ḥaqā'iq* consists of word articles in which related compounds and word-forming suffixes and prefixes are examined.

In separate word articles the lemmas are given as, for instance, Persian *bozorg-dašt* and *dorūd* 'greetings', *nīk o nīkū* 'good, kind', *angešt* and *axgar* and *ātaš* and

7 For the oldest copies, see: Catalogue des manuscrits persans, 1989, Persan n°195, 196, 197, 200, 205, 206; Suppl: turc 265 à 272, 561, 568 et 1347; Flügel (1977, S. 132, n° 128; S. 137 n° 134–135), etc.

8 Among copies of the dictionary, see Flügel (1977: 130–132, n°127, n° 128); Rieu (1966: 514ab (Add.7887)), Dmitriev (2002, N 846, 229), etc.

*āzar* ‘spark, fire’, *deh* and *dehqān* and *kand* and *rustā* ‘village, town’, *xāme* and *kelk* ‘pen’, *abraš* ‘spotty horse’ and *šabdīz* ‘black horse’, comprising nearly 400 words in total.

The character and structure of this work are supporting evidence that it had the role of an educational manual and was used in the *medreses* and by aspiring poets who studied Persian poetry and its nuances as part of their literary development as well as Persian lexicography in Turkish.

The glossary was not published, despite its great number of manuscripts, but the material it contains is important for studying sources and history in the field of Persian and Turkish dictionary writing (the scientific study of the glossary, based on the material in this copy, is being done).

### 3.2 Dictionaries in verse

3.2.1 The Persian-Ottoman Turkish dictionary *Tuhfe-i Šāhidī* (The gift of Shahidī) (921/1514) contains the vocabulary of the work *Maḡnavī-yi Ma‘navī* by Jalāladdīn Muḡammad Balkhī (Mewlānā Rūmī, 13th century), one of the most studied and often copied works in Turkish studies (Paul de Lagarde 1884: 8–9, 29–30, Verburg 1997: 5–87, Kılıç 2007: 516–548).<sup>9</sup> It is included in two miscellanies in the Arabic script collection at the MM: Mss. N 741, N 937 and one volume under the N 325.<sup>10</sup> The ff.108v–127r of the Persian-Turkish-Arabic trilingual manuscript miscellany N 741 contain the dictionary *Tuhfe-i Šāhidī* consisting of 22 *qīṭa*’s. The miscellany was composed in 1196/1781 in Ankara and the scribe is Muḡammad al-Marašī.

The miscellany N 937, according to its colophon (f.47v), was written earlier, in 1033 /1623–1624, by the scribe ‘Amīd al-Ḥuseyn. In this manuscript miscellany, the scribe Farīd ad-dīn ‘Aṭṭār wrote the dictionary consisting of 21 *qīṭa*’s on the ff.27v–47r, after the work *Pandnāma*. The colophons say that the manuscript contains additional *beyts*: *almalḥaq*. The copy is unique for its working style: the missing *beyts* are added, explanations and commentaries are provided in the margins, and at the end there are probably the names of additional scribes.

The manuscript N 325 contains the entire dictionary *Tuhfe-i Šāhidī*, the scribe of which is Muṣṭafa Natīf. It consists of 24 pages written in the script *nasx*. The artistic design of the book gives evidence that it had aesthetic value and perhaps even a “commercial” purpose. The text of the manuscript is the most complete: preface, 27 *qīṭa*’s, a date calculated with *ebjed*.<sup>11</sup>

3.2.2 The other Persian-Turkish dictionary in verse kept in the Arabic Script collection at the Yerevan Matenadaran is the *Tuhfe -i Vehbī* (Gift of Vehbī), written in

9 Among many copies, see: Flügel, 1977: 135, n° 131, 136 n° 132; Cf. Rieu 1966: 513v–514r (Harl. 5494), etc.

10 About copies N 937, 741, 55, 983 kept in MM, see Avetisyan (2011: 148–167).

11 See descriptions in Kostikyan (2017, N 741, 237, N 937, 243, N 325, 134–135).

the 18th century (Kılıç 2007: 410–475; Öz 1997: 219–232).<sup>12</sup> It is included in the manuscript Miscellany N55, while the Mss N 983 is a separate copy.<sup>13</sup>

In the manuscript miscellany (ff.73v–96r), the dictionary is copied after the *Dīvān* of Šāʿib Tabrīzī, an Iranian poet of the 17th century. The dictionary has 57 *qīṭa*’s, one *maṣnavī* with 201 *beyts*, and a *tārīx* with 5 *beyts*, according to which the author wrote it in 1196/1781–1782. The name of the scribe of the manuscript N 55 is known, Šaʿīd Muḥammad Laṭīfī (096r), but the date of copying is unknown. The other copy is the separate volume N 983, which has 57 *qīṭa*’s. There is no information about either the date of writing or the scribe.

One may say that the history of medieval Persian-Turkish bilingual dictionaries ends with this dictionary written in the 18th century.

#### 4. Conclusion

The study of the medieval bilingual Persian-Ottoman Turkish dictionaries and the linguistic interrelations of the time attest to the great influence of the Persian language on Turkish, revealing numerous Persian, as well as Arabic words which were introduced into Turkish by their intermediary and became a part of the Turkish vocabulary. The manuscript dictionaries kept at the Matenadaran are considered to be *farhangs* and glossaries, and play a significant role in studying the medieval educational methodology. The marginal notes and the absence of scribes’ names in some of these dictionaries indicate that these copies were manuals used in *medreses*, which explains the existence of numerous copies. As manuscripts, they are considered to be copies of an early period and can be useful for the compilation of critical texts.

#### References

- Avetisyan A. 2011. Matenadarani parskeren-turk‘eren baġarannerə, [The Persian-Turkish dictionaries of the Matenadaran]. *Collection of articles of the Second International Armenological and Iranological Symposium*. Yerevan: Nairi. 148–167.
- Baevskii, S. (Revised and updated by Perry, J. R.) 2007. *Early Persian lexicography: Farhangs of the eleventh to the fifteenth centuries*. Translated by Killian, N. Folkestone: Global Oriental LTD.
- Brunschvig R. 1963. Kemāl Pāshāzāde et le persan [Kemāl Pāshāzāde and Persian]. *Mélanges d’orientalisme offerts a Henri Massé*, Téhéran: Imprimerie de l’Université.
- Dmitrieva, L. V. & Akimuškin, O. F. (eds.) 2002. *Katalog tjurkskix rukopisej Instituta vostokovedenija rossijskoj Akademii nauk* [Catalogue of Turkic manuscripts of the Institute of Oriental Studies of the Russian Academy of Sciences]. Moscow: Vostočnaja Literatura.
- Csató, É. Á. & Brendemoen, B. & Johanson, L. & Römer, Cl. & Stein, H. 2016. The linguistic landscape of Istanbul in the seventeenth century. In: Csató É. Á. & Menz, A. & Turan,

12 See copies in Flügel 1977: 143–44, n° 141; Rieu 1966: 515v (Add. 7687).

13 See descriptions in Kostikyan 2017, N 55, 30–31, N 983, 249.

- F. (eds.) *Spoken Ottoman in mediator texts*. (Turcologica 106.) Wiesbaden: Harrassowitz. 1–33.
- Flügel, G. 1977. *Die arabischen, persischen, türkischen Handschriften, der k.u.k. Hofbibliothek zu Wien*. Hildesheim: Olms Verlag. 1977.
- Fragner, B. 2006. Das Persische als Hegemonialsprache in der islamischen Geschichte: Überlegungen zur Definition eines innerislamischen Kulturraums. In: Johanson, L. & Bulut, Chr. (eds.) *Turkic-Iranian contact areas: Historical and linguistic aspects*. (Turcologica 62.) Wiesbaden: Harrassowitz. 39–49.
- Kılıç, A. 2007. Türkçe-Farsça manzum sözlüklerden *Tuhfe-i Şahidi*. Metin [One of the Turkish-Persian Verse Dictionaries: Tuhfe-i Şāhidī (Text)]. *Turkish Studies. International Periodical for the Languages, Literature and History of Turkish or Turkic* (Erzincan) 2: 4, 516–548.
- Kılıç, A. 2007. Türkçe-Farsça manzum sözlüklerden *Tuhfe-i Vehbî*. Metin [One of the Turkish-Persian Verse dictionaries: Tuhfe-i Vehbî (Text)]. *Turkish Studies. International Periodical for the Languages, Literature and History of Turkish or Turkic* (Erzincan) 2: 2, 410–475.
- Kostikyan, K. 2017. *Catalogue of Persian manuscripts in the Matenadaran*. Yerevan: Nairi.
- Lagarde, P. de 1884. *Persische Studien*. Göttingen: Dieteriche Verlags-Buchhandlung.
- Öz, Y. 1997. Tuhfe-i Vehbî Şerhleri [The commentaries of Tuhfe-i Vehbi]. *İlmî Araştırmalar* (İstanbul: İlim yayma cemiyeti) 5, 219–232.
- Öz Y. 2016. *Tarih boyunca Farsça-Türkçe sözlükler* [Persian-Turkish dictionaries during History]. (Türk Dil Kurumu Yayınları) Ankara: Türk Dil Kurumu.
- Richard, Fr. 1889. *Catalogue des manuscrits persans* (Ancien fonds persan de la Bibliothèque nationale). Paris: et VII Planches.
- Rieu, Ch. 1966. *Catalogue of Persian manuscripts in British Museum* 2. London: Gilbert & Rivington.
- Sāme'ī, H. 1995. Dictionaries. *Encyclopaedia Iranica* 7, fasc. 3, 395.
- Stein, H. 2006. Palatal-velar vocalism of Arabic-Persian loanwords in 16th century Ottoman Turkish. In: Johanson, L. & Bulut, Chr. (eds.) *Turkic-Iranian contact areas. Historical and linguistic aspects* (Turcologica 62.) Wiesbaden: Harrassowitz. 143–158.
- Tietze, A. & Lazard, G. 1967. Persian loanwords in Anatolian Turkish. *Oriens* 20, 125–168.
- Turan, F. 2014. Adventures of a mediaeval language book into modern times: Persistence of *Şihāhu'l-'Ajam* in Ottoman language learning and its textual problems. In: Demir, N. & Karakoç, B. & Menz, A. (eds.) *Turcology and linguistics*. Ankara: Hacettepe Üniversitesi. 441–441.
- Verburg, A. C. 1997. The Tuhfe-i Şāhidī: A sixteenth century Persian Ottoman dictionary in rhyme 1. *Archivum Ottomanicum* 15, 5–87.
- Yavuzarslan, P. 2009. *Osmanlı dönemi Türk sözlükçülüğü* [Turkish lexicography of the Ottoman period]. Ankara: Tiyem Yayıncılık.

# Krymchak language samples

Henryk Jankowski

Jankowski, Henryk 2018. Krymchak language samples. *Turkic Languages* 22, 175–193.

This article presents language samples recorded from a single informant, Ms Sultan Peisakh (1924, Simferopol; 2014, St. Petersburg). Ms Peisakh's language competence was limited, but there were no better native speakers available. She could only remember the language she spoke and heard in her childhood, and pronounce some expressions and simple sentences. Although Krymchak is a well-documented language, the available texts represent various written styles and folklore influenced by Crimean Tatar and Turkish. In contrast, documentation of spoken Krymchak is very poor. Therefore, even short recordings are worth publishing.

Keywords: Krymchak Turkic language, language documentation

Henryk Jankowski, Department of Asian Studies, Adam Mickiewicz University in Poznań, Niepodległości 24, PL-61-714 Poznań, Poland. E-mail: henko@amu.edu.pl

## 1. The Krymchaks and their ethnic language

The Krymchaks are a small Rabbinic Jewish community who once called themselves *beni Israel* 'Israel's sons' or *Yisraeler* 'the Israelites', but who since the beginning of the 19th century have been known by the Russian name 'Krymchak'. Their population was assessed to about 8,000 in 1939. During World War II, most of them were exterminated by the Nazis in the Crimea (Kupoveckij 1983: 86) and the 1959 Soviet census showed 1,500 Krymchaks (Kupoveckij 1989: 64). Their number decreased to 1,053 in 1970 (Kupoveckij 1983: 86). In Kizilov's (2008: 66) opinion, in 2002 there were only 204 Krymchaks in the Crimea and about 600–700 in Israel (2008: 68, 71), to which most of them migrated in the 1990s.

Krymchak is practically an extinct language. According to Polinsky (1991: 130), who carried out fieldwork among the Krymchaks in the 1980s, all native speakers were over 70 years old. There is no commonly accepted opinion about the Krymchak language. Some authors consider the Turkic language of the Krymchaks to be an ethnolect of Crimean Tatar (Ianbay, Erdal 1998: 1, Ianbay 2001: 502), or "the Jewish Krimchak ethnolect of the Crimean Tatar language" (Ianbay 2000: 4); for more details see Jankowski (2016: 455–456). The Krymchaks themselves had various names for their language, e.g. Crimean Tatar, Tatar, Krymchak, Turkish, Krymchak-Tatar, Tatar-Krymchak, Crimean Jewish and even Chaghatai (Jankowski 2016: 456). Therefore, they were not unanimous in their opinion about the Turkic language they spoke. This question will not be studied in this article, but it is clear that, from a purely structural point of view, the language spoken by Ms Peisakh may be considered to be a variety of Crimean Tatar strongly influenced by Turkish, though



with some distinctive features such as the -{GAy} optative and specific Krymchak words like *čeče* ‘dumpling’ or *arle* ‘drink of roasted flour’, not to speak of religious terms.

Krymchak literature is known from manuscripts with various contents, so-called *jönks*,<sup>1</sup> eight printed books with the translations of biblical books published at the beginning of the 20th century, a few primers and notebooks published in the Soviet Union in the 1920s–1930s (Kaja 1928, 1930), editions of fragments of old manuscripts by Rebi and V. M. Lombrozo,<sup>2</sup> re-editions of earlier printed texts by Rebi,<sup>3</sup> and the publication of proverbs and sayings by Ačkinazi.<sup>4</sup> Most of these publications appeared in the 1990s and 2000s, which was the period of an unsuccessful attempt to revive Krymchak culture and language in the Crimea.

As for studies, Filonenko (1972) was long the only available study with texts. In 1988 Polinskaja and Černin published an important study on Krymchak kinship terms, which was followed by two studies by Polinsky (1991, 1992), the former comprising eight texts with an introduction and discussion, and the latter a description of the language as recently documented and related questions. A very important event was the edition of three printed books of religious content (of the eight mentioned above), two of which were biblical translations and *targumim*, namely *Targum Shen* of the Book of Ruth of 1906 by Erdal & Ianbay (1998) and the Song of Songs printed in 1905 with a *targum*, edited by Ianbay (2017). The third was the *Book of Miracles and Wonders* printed in 1907 and edited by Erdal & Ianbay (2000). All three were originally published by Nissim N. Levi Chakhchir. In 2010 Rebi published the *Book of Daniel* from Gabay’s *jönk* (2010: 163–234).

In Ianbay (2000) there are fragments of some Krymchak manuscripts from St. Petersburg, in Ianbay (2001) some other short texts, and in Ianbay (2002) fifteen poems by Perich. Lastly, the Krymchak translations of Obadiah and some short texts from manuscripts preserved in Kiev were published by Shapira (2016, to appear). These recent studies by Shapira are very important, for he edits fragments of the 18th-century manuscripts,<sup>5</sup> some with the same texts as those printed in 1905–1907

1 *Jönk* ‘collection of folk and sometimes religious literature’ is referred to in the Russian form as *jonka* (*džonka*). At one time they were very popular. Filonenko (1972: 15) mentions *jönks* owned by Purim, Chapicho, Mizraxi and Zengin.

2 Fragments of Šolom Baxšij’s and Yosef Gabay’s *jönks*, such as the story of Ašik Ğariḅ with a Russian translation were published by Rebi in Rebi & Lombrozo (2000: 34–132), another variant being available in Rebi (2010: 9–134).

3 Rebi republished the texts of stories from Kaja (1930) in Cyrillic transcription and Russian translation in (1993: 27–35) and (2004: 31–51). Rebi (2004: 3) says that they were published at the end of the 1930s; he probably used a later reprinted version of this reader.

4 These proverbs and sayings were prepared for print by B. I. Ačkinazi in 1991 and published posthumously (Ačkinazi 2004).

5 Polinsky’s text 1 (1991: 135–136) is also old; it is dated to 1785. However, since it was republished from a non-critical edition by Kaya, its linguistic reliability is restricted. In

by Chakhchir. We can see that the language of these manuscripts is very archaic and similar to that of contemporary Karaim Bible translations.<sup>6</sup>

As far as the description of the language and the lexicon is concerned, in addition to Polinsky (1992) there is a sketch of grammar by Rebi & Ačkinazi & Ačkinazi (1997), and a more recent one by Rebi (2004: 4–26). We also have two dictionaries by Rebi (2004) and Ianbay (2016). These descriptions and dictionaries reflect Krymchak as known from the texts listed above. It is certain that after the publication of a more extensive corpus of earlier texts, as in Shapira (2016, to appear), there will be a need to write another grammar and another dictionary of pre-nineteenth-century Krymchak, not affected by Turkish.<sup>7</sup>

As for the texts recorded from the native speakers, it is important to distinguish between folk literature like proverbs, saying, riddles, songs and popular stories which were common to all Crimean Turkic peoples, i.e. also memorized, recited and copied by the Krymchaks,<sup>8</sup> and spoken Krymchak. The early researchers were mostly interested in folklore. For example, in 1928 Filonenko recorded nine songs sung by Z. Zengin in Simferopol. In 1947, Kaja recorded a poem spoken by M. Purim, all published by Filonenko (1972: 16–34).<sup>9</sup>

Samples of spoken Krymchak were first published by Polinsky (1991). However, there are in fact only five texts recorded from three native speakers, the other texts being published earlier and only checked by Polinsky's informants. This is because she admitted that all Krymchak native speakers were in fact semi-speakers, able to produce only a few sentences, relate short texts heard before, and carry on short practical dialogues. She added that their vocabulary was restricted (Polinsky 1991: 130–131). In 2006, Iryna Dryga & Yong-Song Li recorded Krymchak material from David (Davut) Rebi, the Krymchak scholar and author of many works on this lan-

fact, we see some SW Turkic features in this text that may or may not have been introduced by the editor.

- 6 There are also a few publications in Turkey, which unfortunately are unreliable, for example Altınkaynak (2006), which quotes various texts without reference to the source and shows "Arami yazı ile Kırımçak Türkçesi Metinleri", i.e. 'Krymchak Turkish texts written in Aramaic characters' which are reproduced from a Karaim(!) manuscript.
- 7 Especially manuscript Evr I 143, among the holdings of the Russian National Library in St. Petersburg, probably written on 15th-century paper, formerly described by Harkavy & Strack (1875: 168), and considered to be Karaim, should be examined for a possible Krymchak provenance which cannot be excluded at this stage of research.
- 8 Ananiasz Zajączkowski (1939) and Włodzimierz Zajączkowski (1961), two Turcologists of Karaim descent, were aware of this and referred to the songs and poetry once popular among the Karaims in the Crimea which they published as Tatar-Karaim.
- 9 Ianbay (2000) has distinguished the following genres of Krymchak literature: (1) Bible translations and other religious works; (2) poetry; (3) songs; and (4) prayers and dirges. In another article (Ianbay 2001), she presents a slightly different classification: (1) translations of the Bible and other religious works; (2) songs; (3) epic; and (4) poetry. For a recent outline of Krymchak literature, see Jankowski (2016: 461–462).

guage. Dryga (2016: 257) stressed that Rebi was the only native speaker able to speak his language fluently. She published fragments of their recordings, mostly words, and characterized them linguistically. This material will be very interesting when published in full, but one should not forget that Rebi is not an ordinary native speaker and his knowledge is a kind of extended, learned and self-studied competence.<sup>10</sup>

Therefore, because of a very limited corpus, it is worth extending our knowledge with new material. Since the competence of our informant was limited, I decided to add a few songs recited by Ms Peisakh as additional material for colloquial Krymchak.

## 2. Ms Sultan Peisakh and her life story

Ms Sultan Peisakh (Sultana Abramovna Pejsax), born in 1924 in Simferopol, Crimea, lived in Simferopol until the age of 16. When the Germans invaded the peninsula, she left with her mother and settled in Almaty (Alma-Ata), Kazakhstan, where they lived until 1960. Her father joined them later. In 1960 they all moved to St. Petersburg (then Leningrad) where she lived until her death in 2014.

I visited Ms Peisakh in her flat in St. Petersburg on 24 July 2010, having been introduced by her relative, Ms Irina Romodina,<sup>11</sup> to talk about the Krymchak past in the Crimea<sup>12</sup> and make recordings. Ms Peisakh said that she cannot speak Krymchak, but understands everything and if one tells her a Krymchak word, phrase or sentence she can translate it into Russian. This is because she was never interested in her ethnic language and spoke it rarely. She said that their parents, when living in the Crimea, talked to each other exclusively in Krymchak and also addressed her in this language quite often, although of course they did not use it outside the home. After her parents died, she lived in St. Petersburg alone, having no relatives or friends at all. She said that she only had an elder sister in Moscow, but could not speak to her even by telephone, for her sister was deaf.

Ms Peisakh normally referred to her native language as Tatar, i.e. Crimean Tatar. She maintained that the songs they once sang were Tatar, as was the national dance *çaytarma* they used to dance. However, when asked if they spoke Crimean Tatar or Krymchak, she answered that it was Krymchak. She stressed that these two languages were very similar, but Krymchak was distinct. During our conversation she also pointed to some differences between Krymchak and Crimean Tatar. It must also

10 I often visited Mr Rebi in the Crimea and found him an enthusiastic and very helpful elder colleague. He spoke Krymchak when he was a child and youngster in the 1920s and 1930s, but later he lived and worked in Russia and did not have any contact with his native language. After he retired, he returned to the Crimea and started studying it by himself, mastering the Hebrew writing and studying old handwritten documents.

11 I owe thanks to both of them to make me this work possible.

12 All Krymchaks find it deeply distressed to talk about their history, because of their traumatic memory of the genocide of their nation by Nazis during World War II.

be stressed that her opinion about the Crimean Tatars was very positive, and she always called them close friends and good neighbours.

Unfortunately, Ms Sultan Peisakh in fact could not speak Krymchak. She was only able to translate some words, phrases and sentences into Krymchak as well as to pronounce some idioms and expressions and give indirect information on grammar. As a physician, she did not have any contact with Krymchak culture, and used Krymchak only in conversation with her close relatives, in practice her parents while they were alive.

### 3. Language samples

The samples recorded are of various types. Firstly, there are a few phrases and sentences that Ms Peisakh uttered spontaneously during our conversation. Secondly, there are a few sayings and expressions pronounced by her at my request. Thirdly, there are some sentences translated by her from Russian. Finally, there are a few songs which Ms Peisakh had learned from her mother and the lyrics of which she recited to me without singing.

#### 3.1 Phrases and sentences

These phrases and sentences are provided in the order in which they were recorded. They refer to common everyday situations, which is why our informant remembered them.

- (1) *Ekı de dülber kizim var.*  
two too pretty daughter-1POSS existing  
'I have two pretty daughters.'
- (2) *Kapini yap!*  
door-ACC shut.IMP  
'Shut the door!'
- (3) *Men evdem.*  
I house-LOC-1COP  
'I am at home.'
- (4) *Men seni seviyim.*  
I you-ACC love-1PRES  
'I love you.'

- (5) *Ökmek*<sup>13</sup> *var.*  
bread existing  
'There is some bread.'
- (6) *Ökmek yoç.*  
bread non-existing  
'There is no bread.'
- (7) *\*Penjere aç!*  
window open.IMP  
'Open the window!'

Ms Peisakh uttered this sentence, but afterwards she doubted its correctness. Correctly, it should be *Penjereni aç*, for the object in a sentence so pronounced must be definite.

- (8) *Vay vay anam!*  
O o mother-1POSS  
'O my mother!'
- (9) *Korçma*<sup>14</sup> *balam.*  
fear-NEG.IMP child-1POSS  
'Don't be afraid, my child.'
- (10) *Yaymur yaya.*  
rain rain-3PRES  
'It is raining.'

### 3.2 Sayings and expressions

Ms Peisakh was able to recall the following sayings and expressions:

- (11) *Közleriñi öpeyim.*  
eye-PL-2POSS-ACC kiss-1PRES  
'I am kissing your eyes,' said when addressing a child with love.
- (12) *Say boluñız.*  
healthy be-IMP.POLITE  
'Thank you.'

13 Cf. Rebi (2004: 155) *ökmek*, Ianbay (2017: 20) *äkmek*, *ekimek*, *ekmek*, *okmek*, *okmäk*, *ökmäk* 'bread'; see also the Standard Crimean Tatar *ötmek* (Useinov 2005: 190).

14 Cf. Ianbay (2016: 166) *qorq-*, *qorx-* 'to fear, to be frightened', but also *korxli* 'terrible, dangerous' (2016: 166).

- (13) *Aš delige kaldi.*  
 food fool-DAT remain-3PAST  
 ‘The guests have eaten few.’

This saying is also known in Turkish, see *Aş deliye kaldı* (Aksoy 1965: 158).

- (14) *Koluğa sayliḡ.*  
 hand-2POSS-DAT health  
 ‘Thank you.’

Neither Rebi nor Ianbay give evidence for this expression; see Turkish *Elinize sağlık* ‘a form of thanksgiving said to one who has done something with his hand’ (Akalın ed. 2005: 619).

- (15) \**Xoš bolunuz.*  
 well be-IMP.2PL  
 ‘Welcome (POLITE).’

The correct form should be *Xoš keldiñiz* ‘Welcome’ (Ianbay 2016: 218), cf. Rebi (2004: 200) *Xoš keldiñ!* ‘Welcome.’

- (16) *Janı dovaçıḡ olay.*<sup>15</sup>  
 soul-3POSS prayer-DIMIN be-3OPT  
 ‘May he rest in peace.’

Ms Peisakh said *janim* ‘my soul’. Although this may also be considered correct as a form of address to a deceased man, the usual form is *janı*.

- (17) *Patlayay!*  
 burst-3OPT  
 ‘May (s)he break up!’

- (18) *Tuz közüñe.*  
 salt eye-2POSS-DAT  
 ‘Beware!’ said to protect someone against the evil eye.

- (19) *Başı var akili yoḡ.*  
 head-3POSS existing wisdom-3POSS non-existing  
*akili var kilifti yoḡ.*  
 wisdom-3POSS existing case-3POSS non-existing  
 ‘Stop playing the wise guy.’

15 For *dovaçıḡ*, see *duva ~ dua* ‘prayer’ (Rebi 2004: 93) and *dova* ‘prayer’ (Ianbay 2016: 65).

Ms Peisakh said *kifiri*, which resembles *küfürü* < *küfrü* ‘his cursing’. However, this makes no sense. The word *kifiri* should probably be amended to *kiliff* ‘his case; his cover’ by analogy to a similar Turkish saying *Başı var, aklı yok, dili var kabı yok* in which *kabı* means ‘his cover; his vessel’ (cf. Redhouse 1991: 571). This saying designates a foolish, talkative, indiscreet person.

- (20) *Kettî*                      *yürek kettî*                      *baş,*  
                                  go away-3PAST   heart   go away-3PAST   head  
                                  *kayttî*                      *yürek kettî*                      *baş.*  
                                  returned-3PAST   heart   go away-3PAST   head  
                                  ‘You should consider what you are saying.’

### 3.3 Sentences translated from Russian

While most of the phrases, sentences and expressions known by the informant from her natural linguistic experience have SOV word order, a number of the sentences which she translated from Russian are SVO. We can therefore conclude that Russian syntax affected the word order of these sentences. In other words, this kind of material does not represent Krymchak as spoken in its natural environment undisturbed by another language. A similar situation is known in translations from Hebrew, especially the canonical texts.

- (21) *Biz kirdik evge.*  
                                  we   enter-1PL PAST   house-DAT  
                                  ‘We have entered the house.’
- (22) *Bizge misafir keldi / Bizge keldi misafir.*  
                                  we-DAT   guest   come-3PAST / we-DAT   come-3PAST   guest  
                                  ‘The guests have come to us; a guest has come to us.’

Both variants were provided.

- (23) *Bu adam yaxşı.*  
                                  this   man   good  
                                  ‘This man is good.’
- (24) *Bu adam yaxşı dugul.*  
                                  this   man   good   not  
                                  ‘This man is not good.’
- (25) *Evge keldim.*  
                                  house-DAT   come-1PAST  
                                  ‘I have come home.’

- (26) *Ketti.*  
go away-3PAST  
'(S)he has gone.'
- (27) *Karasubazarya varyan.*  
Karasubazar-DAT go-3PERF  
'(S)he has gone to Karasubazar.'
- (28) *Kayda siz yaşaysız?*  
where you-POLITE live-2PRES.POLITE  
'Where are you living?'
- (29) *Men institutta işledim.*<sup>16</sup>  
I institute-LOC work-1PAST  
'I worked at the institute.'
- (30) *Men kirdim evge.*  
I enter-1PAST house-DAT  
'I have entered the house.'
- (31) *Men oña söledim<sup>17</sup> kırıñız mana şındı.*  
I (s)he-DAT say-1PAST enter-2IMP POLITE I-DAT now  
'I told him to come to me now.'
- (32) *Men sana čeče vereyim.*  
I you-DAT dumpling give-1OPT  
'I will give some dumplings to you.'
- (33) *Men xayttım evge.*  
I return-1PAST house-DAT  
'I have come home.'
- (34) *O kirdi evge.*  
(s)he enter-1PAST house-DAT  
'(S)he has entered the house.'

16 Ms Peisakh said *işıldım* 'I worked' which is incorrect; for *işle-* 'to work', see Rebi (2004: 101) and Ianbay (2016: 102).

17 Cf. *solemek* (Rebi 2004: 173) and *sole-*, *soyle-*, *söylä-*, *söyle-* (Ianbay 2016: 187) 'to speak, to tell'.



- (35) *Sen kırdın evge.*  
 you enter-1PAST house-DAT  
 ‘You have entered the house.’
- (36) *Siz kırdığınız evge.*  
 you enter-2PAST.POLITE house-DAT  
 ‘You have entered the house.’

### 3.4 Songs of Ms Sultan Peisakh’s mother

The first song is Crimean Tatar, for it is about Bakhchasaray where the Krymchaks did not live. It has the form of the Turkish *mani* and Crimean Tatar *mane*. The fourth (*Keten kôlmek kenarsız*) is of the same kind. The second piece is not a song, but a popular Turkish saying. The third is a Crimean Tatar *çîğ*, i.e. a couplet. Couplets of this type were recited by two or more people one after another, normally as a form of poetry competition. The last one is of the same genre.

<i>Bağçasaray bizimdir,</i>	Bakhchasaray is ours,
<i>İçî tok yüzümdür,</i>	It is abundant in grapes,
<i>Endî kartlar yol berse,</i>	If the elders agree,
<i>Endî leybet<sup>18</sup> bizimdir.</i>	We will be happy.
<i>Bu dünyanın üç [...],<sup>19</sup></i>	There are three [...] in this world,
<i>Biri asretlik,</i>	One is yearning,
<i>Biri xastalıx,<sup>20</sup></i>	Another is sickness,
<i>Biri ölüm.</i>	Still another is death.
<i>Ekî de elim pek semiz,</i>	Both my hands are very fat,
<i>Jeñime sıymay, alay,</i>	They do not fit in my pockets,
<i>Öleğim derdimizden,</i>	I will die from our sorrow,
<i>Kimse duymay, alay.<sup>21</sup></i>	Nobody will notice it.
<i>Keten kôlmek kenarsız,</i>	My linen shirt is very long,
<i>Ölsem menî anarsız,</i>	If I die you will remember me,
<i>Yüreğime ot düştü,</i>	Fire burnt my heart,

18 Cf. *rağbet* ‘esteem’ (Rebi 2004: 164), *rağbet* ‘sympathy, glory’ (İanbay 2016: 172).

19 Cf. Turkish *Bu dünyada üç şey vardır*.

20 Cf. *xastalıx* (Rebi 2004: 198), *xastalıq* (İanbay 2016: 216) ‘illness’.

21 There is a similar variant recorded in Dobruja: *Ekî kolıñ bembıyaz | Cebîñe sıymay | Olecekmen dertinden | Kimse tuymay* (Mahmut & Mahmut 1997: 84). Krymchak version with *semiz* ‘fat’ seems to be better than Dobruja Tatar *bımbıyaz* ‘all white’, since it is about the hands being too big to be contained in pockets.

<i>Kaçınız dostlar yanarsız.</i> <sup>22</sup>	Escape, my friends, lest you be burnt.
<i>Kınalı parmağ yez tırnağ,</i>	The fingers are hennaed, the nails are brass,
<i>Altında[n] oymağ, alay,</i>	A thimble of gold,
<i>Seni karşıdan kórdüm,</i>	I saw you just opposite,
<i>Pek janım sevdim seni.</i>	I loved you very much.

#### 4. Lexicon

Ms Peisakh's personal Krymchak vocabulary, as recorded during my work with her, is limited. She certainly knew more words, but collecting them would have required a longer study than was possible. My questions aimed to elicit basic vocabulary, but also specific Krymchak or Crimean elements of material and spiritual culture.

##### 4.1 Nouns

*Names of ethnic groups of the Crimea*

*Urus* 'Russian'; *Tatar* '(Crimean) Tatar'; *Aşkenaz* 'Jew';<sup>23</sup> *Čufut* '(derogatory) Jew';<sup>24</sup> *Karaim* 'Karaim'.

*Names of religious feasts*<sup>25</sup>

*Pesağ* 'Passover'; *Šabat* 'Sabbath'; *Purim* 'Purim'; *Yom Kipur* 'Yom Kippur, Day of Atonement'.

*Names of dishes*

*čeče* 'dumpling';<sup>26</sup> *čibörek* 'fried meat pie';<sup>27</sup> *lokma* 'small fried ball of dough with honey'; *mača* 'unleavened bread; matzah';<sup>28</sup> *követe* 'baked pie';<sup>29</sup> *kurabiye* 'bis-

22 Another variant of this *mani* ~ *mane* is known from Dobruja: *Keten kölmek kenarsız | Keniñ anañ bek arsız | Cüregimde ot cana | Kaşınız dostlar, canarsız* (Mahmut & Mahmut 1997: 45).

23 Cf. *aşkenazi* 'Ashkenazic' (Ianbay 2016: 13).

24 Identically Garkavec' (2000: 271) for Urum, though without signalling a derogatory style.

25 According to Ms Peisakh, people were not religious at the time of her childhood. In fact they were all atheists. Therefore, she has a very limited knowledge of religion and can say just a few quite general words about the feasts, e.g. that *Purim* was a very cheerful holiday.

26 Cf. *čoče* 'meat dumpling' Rebi (2004: 205) and *čoče* 'patty, pastry' Ianbay (2016: 51).

27 Absent from Rebi (2004) and Ianbay (2016). This is a Crimean Tatar pie; see (Useinov 2005: 304). It is well-known throughout the whole of Russia and in post-Soviet Central Asia as *čeburek*.

28 Identically Rebi (2004: 144); quite interestingly, Ianbay provides a "learned" form *masa* (2016: 124); cf. Crimean Karaim *maçça* (Aqtay & Jankowski 2015: 231), but in the old texts there is also *mača*.

29 This is also a dish known in the whole of Turkic Crimea, cf. *kubete* 'kubbeteh, dish of dough' (Ianbay 2016: 118), Crimean Karaim *köbeti* ~ *követi* ~ *küveti* 'round pie with raw

cuits', *pastel* 'small pie',<sup>30</sup> *sariy* ~ *sariĭx* 'rolls made from dough',<sup>31</sup> *süzme* 'pancake (served with spread butter and ground nuts)',<sup>32</sup> *şorpa* 'soup', *çalva* 'Turkish delight', \**χursöz*<sup>33</sup> 'nuts and dried fruits', *çavurma* 'fried mutton'.

#### *Names of drinks and beverages*

*arle* 'drink of roasted flour',<sup>34</sup> *boza* 'drink of fermented millet', *katik* 'yoghurt', *süt* 'milk', *sütlü tüblü* 'boiled milk with coffee'.<sup>35</sup>

#### *Names of fruits and vegetables*

*alma* 'apple', *armut* 'pear', *büber* 'pepper; paprika',<sup>36</sup> *erik* 'plum', *funduk* 'hazelnut', *inĭir* 'fig', *kavun* 'melon', *patliĭan* ~ *paltajan* 'eggplant', *sarimsaĭ* 'garlic', *soyan* 'onion', *çarbuz* 'watermelon', *yüzüm* 'grapes', *yüzüm-erik* 'fruits (collective word for small fruits like plums and grapes)'.

#### *Kinship terms*

*aya* 'elder brother', *ana* (*menim anam*) 'mother (my mother)', *apa* (*menim apam*) 'aunt (my aunt)', *ata* 'father', *bala* '1. child. 2. son; daughter', *kardaş* (\**menim kardaşlar*)<sup>37</sup> 'brother (my brothers)', *kartana* 'grandmother', *kartata* 'grandfather', *kelin* 'daughter-in-law; bride', *kizi* 'his/her daughter', *küyöv* 1. 'son-in-law; bridegroom'. 2. 'husband', *oylu* 'his/her son', *çatın* 1. 'woman'. 2. 'wife'. As forms of address: *apay* to a woman, e.g. Sultan *apay* 'Ms Sultan', *akay* ~ *açay* to a man, *çanum* to a lady.

meat' (Aqtay & Jankowski 2015: 226), Crimean Tatar *köbete* 'pie filled with meat' (Useinov 2005: 102) and Urum *köbete* ~ *kübete* (Garkavec' 2000: 279); one of Polinsky's texts (1991: 145) describes the preparation of this dish, which she calls 'large round meat-cake'.

30 Identically Rebi (2005: 161) and Ianbay (2016: 148).

31 Cf. *sariĭx* 'kind of sweets' (Rebi 2005: 161), absent from Ianbay.

32 Cf. *süzme* 'dumplings filled with meat, sprinkled with ground nuts' Rebi (2004: 175), *suzme* 'dumplings' Ianbay (2016: 190). Another of Polinsky's texts (1991: 146) is the recipe for preparing this dish; she calls it 'small mutton dumplings similar to tortellini'.

33 Cf. *qurı* '1. dried. 2. dried fruits' (Rebi 2004: 136); Ianbay (2016: 170) glosses the word *quru* only in the common meaning 'dry'; the second component *-söz* is probably identical with Turkish *çerez* also designating dried fruits (Akalın ed. 2005: 416), the initial part of which was deleted; the meaning 'dried fruits' is also evidenced in Urum *xuru* (Garkavec' 2000: 556).

34 Rebi (2004: 63) glosses this word as 'dessert beverage'; absent from Ianbay (2016).

35 Rebi (2004: 176) provides this word, spelled *sütlü-tüblü*, with a completely different meaning 'round dumplings'; absent from Ianbay (2016).

36 Cf. *buberĭx* 'pepper-DIM' (Polinsky 1991: 146); the round vowel which is affected by the preceding labial *b-* is also evidenced in Crimean Karaim *büber* 'pepper' (Aqtay & Jankowski 2015: 99); however, see Crimean Tatar and Urum *biber* (Useinov 2004: 44, Garkavec' 2000: 97).

37 Correctly *menim kardaşlarım*; Ms Peisakh was not sure if this form was correct.

*Parts of the human body*

*ayız* ‘mouth’, *ayaχ* ‘leg; foot’, *baš* ‘head’, *burun* ‘nose’, *kol* ‘hand’, *kóz* ‘eye (*kózüm* my eye(s))’, *parmaχ* ‘finger’, *kulaχ* ‘ear’, *sač* ‘hair’, *tıl* ‘tongue’, *tış* ‘tooth’,<sup>38</sup> *yürek* ‘heart’.

*Kitchen utensils*

*kaşıχ* ‘spoon’, *pıçaχ* ‘knife’.

*Other nouns*

*Alla verdı* ‘something lucky; good luck’, *azbar* ‘yard’, *eküm* ‘doctor, physician’,<sup>39</sup> *mişik* ‘cat’,<sup>40</sup> *penjere* ‘window’, *oylan* ‘boy’.

*Proper names**Krymchak surnames*

Ms Pesakh provided the following surnames: *Aškenazi*, *Baχči*, *Gabay*, *Gurji*, *Kaya*, *Kokuš* (← *Köküş*), *Levi*, *Lombrozo*, *Medini*, *Peisaχ*, *Piastro*, *Kuyumji*, *Šamaš*, *Šalom*, *Trevhoda*.

Of these, *Baχči*, *Gurji*, *Kaya*, *Kokuš* and *Kuyumji* are of Turkic origin, *Lombrozo* and *Piastro* of Italian origin, *Aškenazi*, *Gabay*, *Levi*, *Medini*, *Peisaχ*, *Šamaš* and *Šalom* of Hebrew origin. *Trevhoda* is of unknown origin, but this name is well known.

*Krymchak given names*

Female names: *Lunetta*, *Raxel*, *Rebeka*, *Sultan*, *Ventora*

Male names: *Avraham* ~ *Abraham*, *Davut*.

*Lunetta* and *Ventora* are of Italian origin. The former is a diminutive form of *luna* ‘moon’, the latter is an abbreviated form of *Bonaventura* ‘good fortune’. *Sultan* is Turkic, having been borrowed from Arabic, while the remaining names are Hebrew. However, *Rebeka* is provided in a non-Jewish form, since in Hebrew it is pronounced *Rivqa*; see Hebrew רִבְקָה. *Davut* was pronounced in a Turkic form which has its root in Arabic.

38 Note that Rebi (2004: 207) records the form *čış*, while Ianbay (2016: 50) gives the following variants: *čış*, *diš*, *tiš* and *tış*.

39 Cf. Crimean Karaim *heküm* ‘physician, doctor’ (Aqtay & Jankowski 2015: 182).

40 Cf. *mışıχ* (Rebi 2004: 150) and Ianbay *mişiq*, *mışiq*, *mışıχ* (2016: 129); note that this word is common to Crimean Tatar and Urum (Useinov 2005: 166, Garkavec’ 2000: 308), whereas in Crimean Karaim it sounds like *meçi* (Aqtay & Jankowski 2015: 236).

#### 4. 2 Verbs

The following verbs have been recorded:

*aç-* ‘to open’, *ayıra-* ‘to hurt’, *aşa-* ‘to eat’, *bol- ~ ol-* ‘to be; to become’, *işle-* ‘to work’ (corrected from *işil-*), *kal-* ‘to stay; to remain’, *kel-* ‘to come’, *ket-* ‘to go away; to leave’, *kır-* ‘to enter, to come in’, *korç-* ‘to fear’, *kór-* ‘to see’, *çayt-* ‘to return’, *öp-* ‘to kiss’, *sev-* ‘to love; to like’, *patla-* ‘to break out; to burst’, *söle-* ‘to tell; to say’, *ste-* ‘to want’, *var-* ‘to go’, *ver-* ‘to give’, *yay-* ‘to rain; to snow, etc.’, *yap-* ‘to close’, *yaşa-* ‘to live’, *yırla-* ‘to sing’, *yuxla-* ‘to sleep’.

The verbs from songs are not included. To express existence, phrases with *var* ‘there is’ and *yoç* ‘there is not’ are used.

#### 4. 3 Adjectives

*Qualifying adjectives*

*balaban* ‘big, great’, *biyük* ‘big, old (age)’, *delibaş* ‘fool, silly’, *kıčkene* ‘small, young’, *kulaçlar tolmaz* ‘noisy, inquisitive’, *çasta* ‘ill, sick’, *yağşı* ‘good’, *yaramay* ‘bad; wrong’.

Of these adjectives, *balaban* is typical of the whole of Crimea (Rebi 2004: 71, Ianbay 2016: 25, Useinov 2005: 37, Garkavec’ 2000: 78); *kıčkene* is attested in Ianbay (2016: 112). *Delibaş* is a compound of *deli*, known in the whole of Crimea, and *baş* ‘head’. In this composite form it is recorded in Urum by Garkavec’ (2000: 166). *Kulaçlar tolmaz* is known as an idiom, e.g. Turkish *kulakları dol-* ‘to get tired of hearing the same thing over and over again’ (Redhouse 1991: 683).

*Colour names*

*aç* ‘white’, *kırmızı* ‘red’, *çara* ‘black’.

#### 4. 4 Numerals

The following cardinal numerals have been recorded:

*bİR* ‘one’, *ekİ* ‘two’, *üç* ‘three’, *dört* ‘four’, *beş* ‘five’, *altı* ‘six’, *yedİ* ‘seven’, *sekiz* ‘eight’, *dokuz* ‘nine’, *on* ‘ten’, *yigİRİm* ‘twenty’, *otuz* ‘thirty’, *kırç* ‘forty’, *ellİ* ‘fifty’, *altmİş* ‘sixty’, *seksen* ‘eighty’, *doğsan* ‘ninety’, *yüz* ‘hundred’, *ekİ yüz* ‘two hundred’, *bİN* ‘thousand’.

The numerals from one to ten are identical to Turkish except for *ekİ* ‘two’, which in the Crimea is always used in this form, characteristic of NW Turkic. The tens are also usual, if we disregard the aspirated variants *kırç* ‘forty’ and *doğsan* ‘ninety’, and ‘twenty’ which has an unusual form *yigİRİm*. This numeral has a normal form

*yigirmi* in Ianbay (2016: 228); unfortunately Rebi does not record it. *Yigirim*, in addition to *yigirmi*, is also attested in Urum (Garkavec' 2000: 240).

For a half, *yarım* is used, e.g. *ekı yarım* 'two and a half'.

#### 4.5 Adverbs

Since only a few adverbs have been recorded, they are provided without further classification into semantic groups:

*endi* 'from now on; now', *şimdi* 'now'; *yavaş* 'slowly', often reduplicated, e.g. *men söledim yavaş-yavaş* 'I have spoken slowly'.

#### 4.6 Interjections

Two interjections were recorded, *Ay* 'O, Oh', e.g. *Ay Allam* 'O my God' and *Vay vay* 'O, Oh'.

### 5. Grammar

#### 5.1 Nouns

The plural, possessive<sup>41</sup> and case forms are similar to those in other Crimean Turkic languages. 1st singular possessive forms are often used as forms of address, e.g. *anaçıyım* 'my mummy', *balaçıyım* 'my little child', *jançıyım* 'my child', *sevgiliim* 'my darling'.

The following forms of predicative suffixes with nominals were attested:

Number/person	Singular	Plural
1	<i>men xəstam</i> 'I am sick'; <i>men yaxşıım</i> 'I am good'	<i>bız xəstamiz</i> 'we are sick'
2	—	<i>sız xəstasız</i> 'you are sick'
3	<i>o xəsta</i> '(s)he is sick'	—

#### 5.2 Verbs

*Moods and modality*

*Imperative*

Number/person	Singular	Plural
2	<i>aç</i> 'open'; <i>bol</i> 'be'	<i>açınız</i> 'open'; <i>bolunuz</i> 'be'

41 Ms Peisakh was not sure if the sentence she produced *menim yürek ayıray* 'I have a heart-ache; my heart is hurting' was correct.

*Optative*

The optative has the form {-GAY}, but only the form for 3rd person singular was recorded. It occurred in two expressions: *olay* ‘let it be’ and *patlayay* ‘may (s)he go to hell’.

*Volitional modality*

Number/person	Singular	Plural
1	<i>men steyim yuxlama</i> ‘I want to sleep’	–
2	<i>sen steysin yuxlama</i> ‘you want to sleep’	–
3	<i>o stey yuxlama</i> ‘(s)he wants to sleep’	–

*Necessitative*

The necessitative is expressed with *kerek* ‘needed; it is needed’.

*Tenses**Present tense*

Number/person	Singular	Plural
1	<i>men yirlayim</i> ‘I am singing’	<i>biz yirlaymiz</i> ‘we are singing’
2	<i>sen yirlaysin</i> ‘you are singing’	<i>siz yirlaysiz</i> ‘you are singing’
3	<i>ol yirlay</i> ‘(s)he is singing’	<i>olar yirlaylar</i> ‘they are singing’

*Past tense*

Number/person	Singular	Plural
1	<i>men kirdim</i> ‘I have entered’	<i>biz kirdik</i> ‘we have entered’
2	<i>sen kirdin</i> ‘you have entered’	<i>siz kirdiniz</i> ‘you have entered’
3	<i>o kirdi</i> ‘(s)he has entered’	–

*Future tense*

Number/person	Singular	Plural
1	<i>men ašayjam</i> ‘I will eat’	<i>biz ašayjamiz</i> ‘we will eat’
2	<i>sen ašayjasin</i> ‘you will eat’	<i>siz ašayjasiz</i> ‘you will eat’
3	<i>ol ašayjak</i> ‘(s)he will eat’	<i>olar ašayjaklar</i> ‘they will eat’

Note the unusual second person contracted forms, which are also shown as optional forms in Rebi (2004: 18). According to Polinsky’s samples, these forms should have been pronounced *ašajasin* and *ašajasz*; see *tašladžasın* and *jaghladžasın* (Polinsky 1991: 145).

*Negative forms*

The negative verb forms are formed with the negative suffix {-mA}, e.g. *Korxma* ‘don’t be afraid’, *Men šindı yuxlama stemeyim* ‘I do not want to sleep now’.

*Interrogative verb forms*

Quite interestingly, Ms Peisakh argued that the enclitic *mî* is not used and one should ask using intonation.

*Verbal noun*

Only one {-mA} verbal noun has been recorded; see above.

**5.3 Pronouns****3. 1. Personal pronouns**

The following pronouns occur in the samples recorded:

**Singular**

- (1) *men* 'I' (*mana* 'to me', *meni* 'me', *meniim* 'my');
- (2) *sen* 'you (SING)' (*sana* 'to you', *seni* 'you ACC');
- (3) *o ~ ol* '(s)he' (*oña* 'to him/her').

**Plural**

- (1) *biz* 'we' (*bizge* 'to us', *bizim* 'our');
- (2) *siz* 'you' (POLITE and PL);
- (3) *olar* 'they'.

**Conclusion**

Although there is satisfactory documentation of various written varieties of modern and pre-modern Krymchak, the spoken language is very poorly documented. In fact there are only Polinsky's (1991) texts collected in the 1980s. Polinsky was however only able to find three native speakers who could relate some stories. Therefore, even such a small corpus as in this article must be considered valuable. In the circumstances of ongoing language loss, it is naturally difficult to assess Ms Peisakh's competence in Krymchak, but we can characterize it as follows. Ms Peisakh's vocabulary was very restricted. She could not remember many Krymchak words and provided a few words in incorrect or doubtful forms, e.g. *işil-* 'to work' and *χursöz* 'nuts and dried fruits'. There were also some inadequacies in her grammar. She could not recall the full paradigms of some verb forms, e.g. the 3rd plural of *o stey yuxlama* 'he/she wants to sleep', though she knew how to say 'they are singing'. Nevertheless, she seldom made a mistake, rather admitting that she did not know a form than pronouncing it incorrectly. In the case of a few incorrect forms, she stressed that she was uncertain, e.g. *\*meniim kardaşlar* 'my brothers'. Morphologically, Krymchak as spoken by Ms Peisakh is a NW Turkic language. Phonologically, it is a mixed SW and NW Turkic language, e.g. *var-* 'to go', but *ket-* 'to go away'. There are some forms typical of common Crimean Turkic, less common in Crimean Turkish, some of which have not been included in grammars, e.g. *yigirim* 'twenty'. However, there are also some Crimean Turkish forms, e.g. the *-iy* present as in *seviyim* 'I love', but *öpeyim* 'I kiss'.



## References

- Ačkinazi, Boris Mixajlovič 2004. *Ƙrimčaxlarınñ atalar sözi. Poslovici i pogovorki krymčakov* [Krymchak proverbs and saying]. Simferopol: Tavrija.
- Akalın, Şükrü Halûk et al. (ed.) 2005<sup>10</sup>. *Türkçe sözlük* [Turkish dictionary]. Ankara: Türk Dil Kurumu.
- Aksoy, Ömer Asım 1965. *Atasözleri ve deyimler* [Proverbs and sayings]. Ankara: Türk Dil Kurumu.
- Altınkaynak, Erdoğan 2006. *Kırımçaklar. Kültür, tarih, folklor* [The Krymchaks. Culture, history and folklore] (Turquoise Series 15). Haarlem: Sota.
- Aqtay, Gulayhan & Jankowski, Henryk 2015. *A Crimean Karaim-English dictionary. 10 000 entries*. Poznań: Department of Asian Studies.
- Dryga, Iryna 2016. Lingvistyčni zapysy ostann'ogo nosija krymčackoji movy [Recordings of the language of the last Krymchak native speaker]. In: Bogomolov, O. V. et al. (ed.) *Zagroženi movy. Kryms'kotatars'ka ta inši tjurks'ki movy v Ukrajinu* [Endangered languages. Crimean Tatar and other Turkic languages in the Ukraine]. Kyiv: Instytut Sxodozavstva im. A. Ju. Kryms'kogo. 255–263.
- Erdal, Marcel & Ianbay, Iala 2000. The Krimchak Book of Miracles and Wonders. *Mediterranean Language Review* 12, 39–141.
- Filonenko, V. I. 1972. Krymčakskie etjudy [Krymchak studies]. *Rocznik Orientalistyczny* 35: 1, 5–35.
- Garkavec', Oleksandr 2000. *Urums'kyj slovnyk* [The Urum dictionary]. Alma-Ata: Baur.
- Harkavy, A. & Strack, H. L. 1875. *Catalog der hebräischen Bibelhandschriften der Kaiserlichen Öffentlichen Bibliothek in St. Petersburg* 1–2. St. Petersburg.
- Ianbay, Iala 2000. New data on the literature and culture of the Krymchaks. *Manuscripta Orientalia* 6: 4, 4–13.
- Ianbay, Jala S. 2001. Tjurkskaja literatura krymčakov [Turkic Krymchak literature]. *Materialy po arxeologii, istorii i etnografii Tavrii* 8, 502–508.
- Ianbay, Ialqyn 2002. Marcel Perich's poems. *Rocznik Orientalistyczny* 55: 2, 5–37 + 4 tables.
- Ianbay, Iala 2016. *Krimchak dictionary*. Wiesbaden: Harrassowitz, Ben Zvi Institute.
- Ianbay, Jala 2017. *Pesn' carja Solomona i Targum na jazyke krymčakov* [King Solomon's song and its Targum in Krymchak]. Jerusalem: Click.
- Ianbay, Iala & Erdal, Marcel 1998. The Krimchak translation of a targum *Šeni* of the Book of Ruth. *Mediterranean Language Review* 10, 1–53.
- Jankowski, Henryk 2016. Karaim and Krymchak. In: Kahn, Lily & Rubin, Aaron D. (eds) *Handbook of Jewish languages*. Brill: Leiden, Boston. 451–488.
- Kaja, I. S. 1928. *Qrımcaq balalar'ı cın ana tilinde alefbet ve oqu kitabı* [Primer and reader for the Krymchak children]. Simferopol: Qrım hukümet neşriyatı.
- Kaja, I. S. 1930. *Qrımcaq mekteblerinin ekinçi sınıfına maxsus oquv kitabı* [Reader for the 2<sup>nd</sup> class of Krymchak schools]. Simferopol: Qrım devlet neşriyatı.
- Kizilov, Mikhail 2008. The Krymchaks: Current state of the community. *Euro-Asian Jewish Yearbook* 5768 (2007/2008), 63–89.
- Kupoveckij, M. S. 1983. Dinamika čislenosti i rasselenie karaimov i krymčakov za poslednie dvesti let [Dynamics of population change and settlements of Karaites and Krymchaks in the last two hundred years]. In: Arutjunov, S. A. & Jerenkov, L. N. (eds). *Geografija i kul'tura etnografičeskix grupp tatar v SSSR* [Geography and culture of the

- Tatar ethnographic groups in the USSR]. Moscow: Moskovskij Filial Geografičeskogo Obščestva SSSR. 75–93.
- Kupoveckij, M. S. 1989. K ètničeskoj istorii krymčakov [On the ethnic history of the Krymchaks]. In: Krupnik, I. I. (ed.) *Ètnokontaktnye zony v evropejskoj časti SSSR (geografija, dinamika, metody izučenia)* [Areas of ethnic contact in the European part of the USSR]. Moscow: Akademija Nauk SSSR, Moskovskij Filial Geografičeskogo Obščestva SSSR. 53–69.
- Mahmut, Enver & Mahmut, Nedret 1997. *Dobruja Türk halk edebiyatı metinleri* [Turkic folklore texts in Dobruja]. Ankara: Kültür Bakanlığı.
- Polinsky, Maria S. 1991. The Krymchaks: History and texts. *Ural-Altaic Yearbook* 63, 123–154.
- Polinsky, Maria 1992. Crimean Tatar and Krymchak: Classification and description. In: Aronson, Howard (ed.) *The Non-Slavic languages of the USSR. Linguistic Studies. New Series*. Chicago: Chicago Linguistic Society. 157–188.
- Polinskaja, Maria S. & Černin, V. Ju. 1988. Terminologija rodstva u krymčakov [Krymchak kinship terms]. *Sovetskaja Tjurkologija* 3, 15–23.
- Rebi, D. I. 1993. *Krymčakskij jazyk. Kırımcağ tili* [The Krymchak language]. Simferopol.
- Rebi, David 2004. *Krymčakskij jazyk. Krymčaksko-russkij slovar'* [The Krymchak language. A Krymchak-Russian dictionary]. Simferopol': Dolja.
- Rebi, David 2010. *Pis'mennoe nasledie krymčakov* [The written heritage of the Krymchaks]. Simferopol': Dolja.
- Rebi, D. I. & Ačkinazi, B. M. & Ačkinazi, I. V. 1997. *Krymčakskij jazyk* [The Krymchak language]. In: Tenišev, È. R. (ed.) *Jazyki mira. Tjurkskie jazyki* [World's languages. The Turkic languages]. Bishkek: Kyrgyzstan. 309–319.
- Rebi, D. I. & Lombrozo, V. M. 2000. *Krymčaki* [The Krymchaks]. Simferopol: Krymskoe respublikanskoe kul'turno-prosvetitel'noe obščestvo krymčakov "Qırımcağlar".
- Redhouse 1991 = Alkim, U. Bahadır et al. (eds) 1991. *Redhouse Yeni Türkçe-İngilizce sözlük. New Redhouse Turkish-English dictionary*. İstanbul: Redhouse Yayınevi.
- Shapira, Dan 2016. A Krymchak Obadiah. *Karaite Archives* 4, 117–127.
- Shapira, Dan [to appear]. Preliminary notes on some Krymchak texts from the Ukrainian National Library, Kiev. *Karaite Archives*.
- Useinov, S. M. 2005. *Qırımtatarca-rusça luğat. 25 000-den ziyade söz ve söz birikmesi. Krymskotatarsko-russkij slovar'*. *Svyše 25 000 slov i slovosočetanij* [Crimean Tatar-Russian dictionary. Over 25,000 words and expressions]. Simferopol': Odžak.
- Zajączkowski, Ananiasz 1939. Tatarsko-karaimskie piosenki ludowe z Krymu (t. zw. *čij*) [Tatar-Karaim folk songs from the Crimea]. *Rocznik Orientalistyczny* 14, 38–65.
- Zajączkowski, Włodzimierz 1961. Z poezji ludowej tatarsko-karaimskiej na Krymie [Samples of Tatar-Karaim folk poetry from the Crimea]. *Rocznik Orientalistyczny* 24: 2, 119–147.

# 3D documentation of Old Turkic Altai runiform inscriptions and revised readings of the inscriptions Tuekta-V and Bichiktu-Boom-III

**Irina Nevskaya & Larisa Tybykova & Mikhail Vavulin & Olga Zaytseva & Evgeniy Vodyasov**

Nevskaya, Irina, Tybykova, Larisa Vavulin, Mikhail, Zaytseva, Olga & Vodyasov, Evgeniy 2018. 3D documentation of Old Turkic Altai runiform inscriptions and revised readings of the inscriptions Tuekta-V and Bichiktu-Boom-III. *Turkic Languages* 22, 194–216.

30 Old Turkic runiform inscriptions in the Altai Mountains were digitalized by means of 3D technologies during field research in 2017. How the inscriptions were carved as well as their locations in hardly accessible areas required a solution that would combine a detailed 3D documentation of rock surfaces with mobility and freedom from external power supply. Digital photogrammetry proved to be the optimal choice for achieving high-quality documentation using inexpensive hardware and free software for further data processing. The advantages of 3D documentation of Altai runiform inscriptions are illustrated by the examples of two Old Turkic runiform inscriptions: Tuekta-V and Bichiktu-Boom-III. The 3D results made it possible to propose new readings of these inscriptions based on the revised composition of their runiform signs. It turns out that the Bichiktu-Boom-III inscription might contain a very rare usage of a religious term that can shed light on the history of religious confessions in the area at the time of the creation of those inscriptions.

Keywords: Old Turkic, runiform inscriptions, Republic of Altai, 3D documentation, digital photogrammetry, Tuekta-V, Bichiktu-Boom-III

*Irina Nevskaya, Frankfurt University, Institute of Philology SD RAN, Tomsk State University, Senckenberganlage 3, 160325 Frankfurt, Germany, E-mail: nevskaya@em.uni-frankfurt.de; Larisa Tybykova, Gorno-Altai State University, Tomsk State University; Lenkina st., 1. 659700, Gorno-Altai, Russia, E-mail: fat@gasu.ru; Mikhail Vavulin, Olga Zaytseva, and Evgeniy Vodyasov, Tomsk State University, Lenin Ave, 36, Tomsk 634050, Russia, E-mail: lira-artefact@mail.tsu.ru*

## 1. Introductory remarks

The earliest sources of Turkic, Old Turkic runiform inscriptions, are found all over Eurasia: in Mongolia, the basin of the river Yenisei, Mountainous Altai, Kirgizstan, China and Kazakhstan, among other places. These inscriptions constitute a large corpus of several hundred inscriptions. They are an important part of the written heritage of the Turkic world, providing us with valuable linguistic and historic in-

formation. The so-called Orkhon inscriptions in Mongolia and the Yenisei inscriptions are the most famous and studied ones. As for the Old Turkic runiform inscriptions of the Altai Mountains, they had been considered to be small in number and fragmentary, to present no coherent texts, and to represent a dying runiform tradition (Ščerbak 2001: 45). Consequently, there have been practically no systematic attempts to document the Altai corpus as a whole until recently; see Tybykova & Tybykova (2015) on the history of their research. Such research began in the early twenty-first century with yearly expeditions by Larisa Tybykova and Irina Nevskaya, accompanied at various times by other Turcologists (e.g. Marcel Erdal in 2007 and 2008), archaeologists (e.g. Igor Kyzlasov in 2003), and art historians (Evgeniy Matochkin in 2011 and 2012), etc. Tybykova, Nevskaya and Erdal published the collected inscriptions as Tybykova et al. (2012), which was followed by Vasilev (2013), and Konkobaev et al. (2015), presenting collections of the various authors' data on Altai runiform inscriptions. All these editions partially overlap; however, none of them presents all the known Altai inscriptions in full. All in all, the number of inscriptions published in these collections exceeds 120, while the number of so far discovered inscriptions, both published and still unpublished, is even higher, and is approaching 130. (About sixty inscriptions in Mountainous Altai had been found and published by the beginning of the 21st century).

It is certainly true that the majority of Altai inscriptions are rather short, and many have not been preserved in full; but we should be careful when thinking in terms of a dying tradition. Numerous inscriptions have been found recently and the inventory of known inscriptions is expanding quickly, not to mention the many conceivable inscriptions that may not have survived.

Another widespread opinion is that the Altai runiform tradition differs from the classic Orkhon and Yenisei traditions in that the rules of the runiform script are often violated in the Altai tradition. Some authors explain this by the fact that the scribes were illiterate, while others claim on the contrary, that they were highly educated and knew different writing systems, leading to contamination of their rules. In the meantime, however, dozens of inscriptions violating rules of harmony and of the explicit appearance of vowels have also been found in Mongolia, and many of the Yenisei inscriptions are not at all "regular" either (Nevskaya & Erdal 2015). Thus, at present it is not possible to link the undated inscriptions to any historical process. A rather small number of official inscriptions of the Türk and Uyghur empires were produced by professionals; the authors of the funerary inscriptions of the basin of the Yenisei and its tributaries probably also had some training. However, the rest of the known inscriptions, and not only in the Altai Republic (many as yet unpublished), appear to be the occasional scribbles of the Turkic population at large. "Far from speaking of illiteracy, it seems that one should be impressed by the widespread literacy of a semi-nomadic medieval population engaged in animal husbandry and spread over such a huge area (beside the possibility that texts written down on perishable materials did not survive the climatic conditions)" (Nevskaya & Erdal 2015). All the extant Altai inscriptions are also acutely endangered because of in-

dustrial activities, natural phenomena (erosion caused by the wind and the water, earthquakes), agricultural activities and vandalism by tourists and local inhabitants making graffiti over the old inscriptions. Documenting them, researching them and taking measures for their preservation are all urgent tasks.

As has become clear, the majority of the readings of the Altai and other inscriptions that have been published to date should generally be considered preliminary. Further research on the Altai runiform tradition and on the modern documentation methods to which we will refer in later sections of this article will help us clear up some fundamental questions and, if necessary, propose new readings of the inscriptions.

This article will present our current research on Old Turkic runiform inscriptions in the Altai Republic. In Section 2, we will describe new documentation methods employed since 2017 that allow us to clarify the composition of the runiform signs of a particular inscription and propose its new reading. We then publish our new readings of the inscriptions Tuekta-V (Section 3) and Bichiktu-Boom-III (Section 4).

## **2. Methods for 3D documentation**

### **2.1 Altai runiform signs and difficulties with documenting and reading them using contact methods**

Deciphering the Altai inscriptions is significantly hampered by the way the inscriptions were made. Altai runiform inscriptions were mostly made on vertical rocks, on surfaces 0.02–1.0 m<sup>2</sup> in area. They differ in length, and can comprise anywhere from a few to a few dozen runes. Their lines are very thin (0.1–1.0 mm wide, less than 0.1 mm deep). Their signs are normally very small, about 1.0–3.0 cm high, which makes it very difficult to copy and read them. The freshly-cut lines used to contrast with the weathered rock surface, but have become less discernible, or almost invisible with time. Today, we can sometimes see them only from a specific angle with the sunlight coming from the side. In some cases, runes are also difficult to read because of more recent petroglyphs and graffiti. Also natural stone defects can easily be mistaken for man-made lines.

Available contact methods, traditionally used to document petroglyphs, are not suitable for providing a reliable documentation of runiform inscriptions. For instance, the method of copying petroglyphs on mica-coated paper, which has been widely used in Russia, has some serious limitations even for petroglyphs. Such copies are effectively useless in the case of Altai runiform inscriptions, as our attempts to copy these inscriptions by rubbing showed: not a single visible sign could be discerned on such rubbings. The method of copying inscriptions on transparent foliae proved to be more useful, and was used by Nevskaya and Tybykova in their expeditions of 2003–2016. However, it also has its limitations. Because drawings, strokes, and carvings were done at different times on the same rock, and practically

always are accompanied by natural rock splitting, the ways researchers see the signs of the same inscription (and consequently copy them) might differ so greatly that very different readings come into existence; see Sections 3 and 4. Sometimes, even one or two strokes can dramatically change our understanding of an inscription.

Altai runiform inscriptions therefore require not only high-resolution macro-photography but also rock relief capturing in order to be reliably documented. Non-contact methods using 3D scanning and digital photogrammetry are currently being developed.

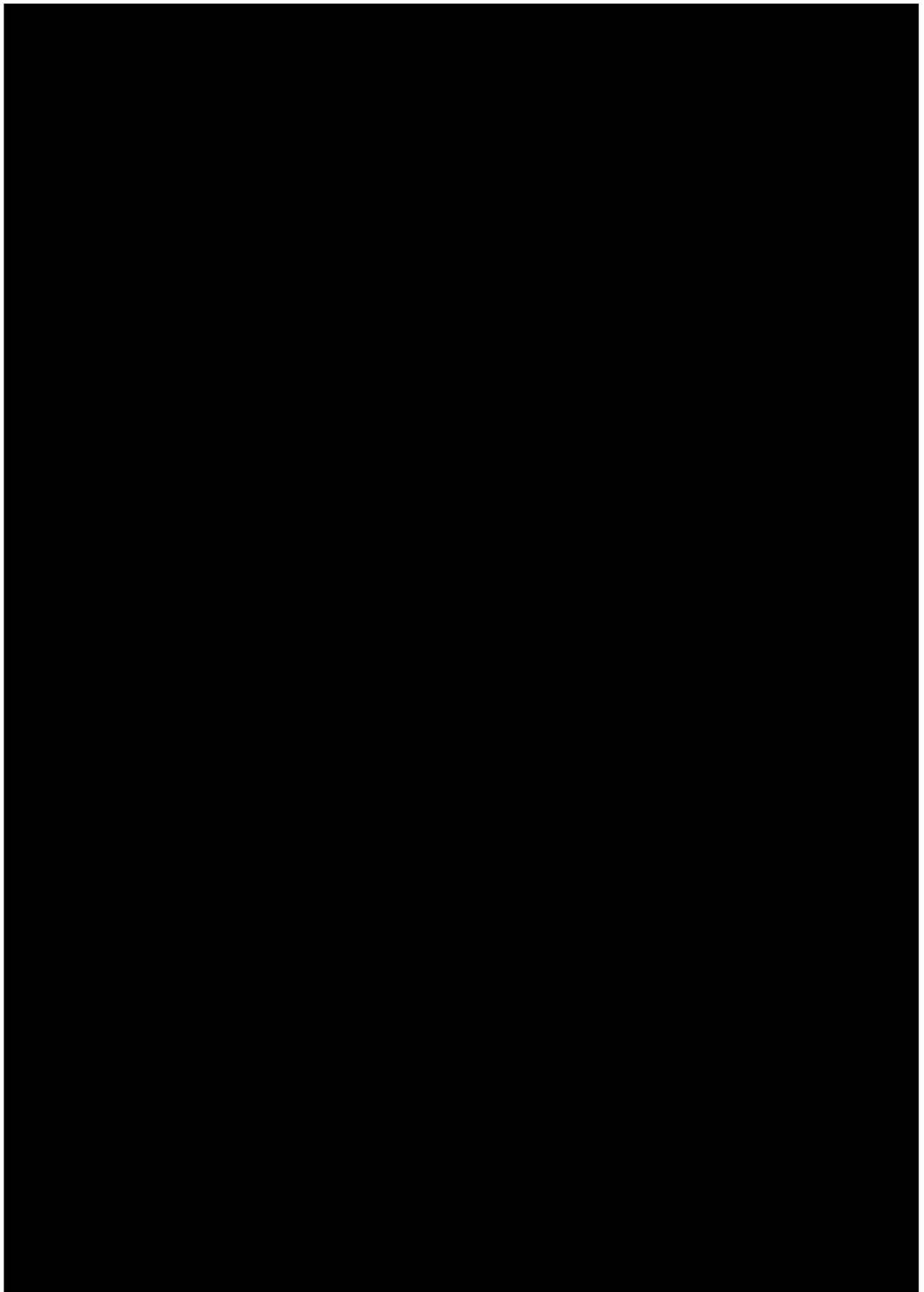
## 2.2 Choice of technology

Since 2017, the Laboratory of Digital Archaeology at Tomsk State University (Olga Zaytseva, head of the laboratory, Mikhail Vavulin and Evgeniy Vodyasov, staff members) is supporting the ongoing documentation of the Altai inscriptions by Nevskaya and Tybykova with up-to-date technologies. Alongside the already mentioned participants, students from Frankfurt University and the Gorno-Altai State University took part in the expedition in 2017.

Today there are numerous brands of 3D scanners, capable of scanning tiny details on the surface of objects. Various technologies are used: structured light, laser scanning, computer tomography, etc. Such scanners are very expensive, and either are stationary or depend on robot manipulators. Moreover, they require much energy and permanent external power supply for their operation. Altogether, this makes them unsuitable for use in inaccessible areas like the Altai Mountains.

Some portable 3D scanners have a resolution of 0.1 mm and higher, but are also very expensive. Among less expensive solutions (under EUR 20,000), only structured-light models can boast of specifications sufficient to document runiform inscriptions (Vavulin 2017). However, they depend on the ambient light and require an external power supply. Such equipment cannot be used to document Altai runiform inscriptions either. Therefore, for documenting Altai runiform inscriptions, it was decided to use the photogrammetric technology, which allows for both mobility and energy autonomy (Vavulin et al. 2018).

The highest-resolution photo camera and the highest-focal-length macro photo lens were selected from the available laboratory equipment to provide maximum detail-capturing and 3D models with resolutions down to 0.015 mm, which should be enough to document details as tiny as 0.05 mm on the surface, if required (Vavulin et al. 2014). Depending on the time of the day and weather conditions, shooting was performed either in natural light or in artificial shadow covering the whole surface. The camera was moving vertically, parallel to the digitalized plane, with a 70% image overlap; a side overlap of 60% was provided using a parallel shift between the images. Agisoft Photoscan Pro software was used to process the data and prepare 3D models of the inscriptions, QGIS Desktop to analyze the results and develop orthophotos and heightmaps, and Meshlab and CloudCompare for working with 3D models (see Vavulin 2017 for the technical details). 3D data analysis allows for analyz-



word dividers. The inscription is situated under the feet of the goats, but upside down. The authors supposed that the graffiti was made earlier. The scribe ought to have been lying on the stone while writing the inscription, trying to place it under the feet of the goats, which resulted in its positioning. It is read from right to left.

*Reading of I. A. Nevskaya and L. N. Tybykova (Tybykova et al. 2012)*

Runiform transliteration: : 1 𐰚 𐰢 𐰣 𐰤 𐰥 𐰦 𐰧 𐰨 𐰩 𐰪 𐰫 𐰬 𐰭 𐰮 𐰯 𐰰 𐰱 𐰲 𐰳 𐰴 𐰵 𐰶 𐰷 𐰸 𐰹 𐰺 𐰻 𐰼 𐰽 𐰾 𐰿

Transliteration:  $k^2 I y^1 d^1 g^1 : y^2 p / s^2 s^1 U \ddot{U} n^2 m s^2 :$

Transcription: (ä)ki y(a)d(a)g : y(e)p/y(i) š (a)šu ün(m)iš :

Translation: Two (persons/or other living beings, going) on foot traversed (a mountain) and descended (into the valley), (after) having eaten. Or: Two (persons going) on foot traversed a mountain and descended (into the valley).

Comments:

1. The eighth sign was not clear to the interpreters. It could be 1 or I. Consequently, they proposed two different interpretations. In the first case, they presupposed a sequence  $y^2 p s^1 U \ddot{U} n^2 m s^2$ , which gave the interpretation *y(e)p (a)šu ün(m)iš* ‘having eaten, traversed (a mountain) and descended’. In the second case, it was *y(i)š (a)šu ün(m)iš* ‘traversed a mountain and descended’.

2. The inscription was written in the classic runiform orthography.

*Reading by Konkobaev et al. (2015)*

In 2012, several Kirgiz Turcologists headed by Kadyraly Konkobaev studied this inscription on site. It was shown to them by a local teacher, Boris M. Kindikov. Their fixation was more precise; they discerned the number of the signs correctly—17. They agreed with some fragments of the first reading, but interpreted the verb *ün-/ön-* (DTS) as ‘grow’. Nevertheless, the first part of the inscriptions remained undeciphered, as traditional methods of copying did not give the necessary details for a correct reading of the first two signs.

Konkobayev et al. 2015: 290-291

Runes: 𐰚 𐰢 𐰣 𐰤 𐰥 𐰦 𐰧 𐰨 𐰩 𐰪 𐰫 𐰬 𐰭 𐰮 𐰯 𐰰 𐰱 𐰲 𐰳 𐰴 𐰵 𐰶 𐰷 𐰸 𐰹 𐰺 𐰻 𐰼 𐰽 𐰾 𐰿

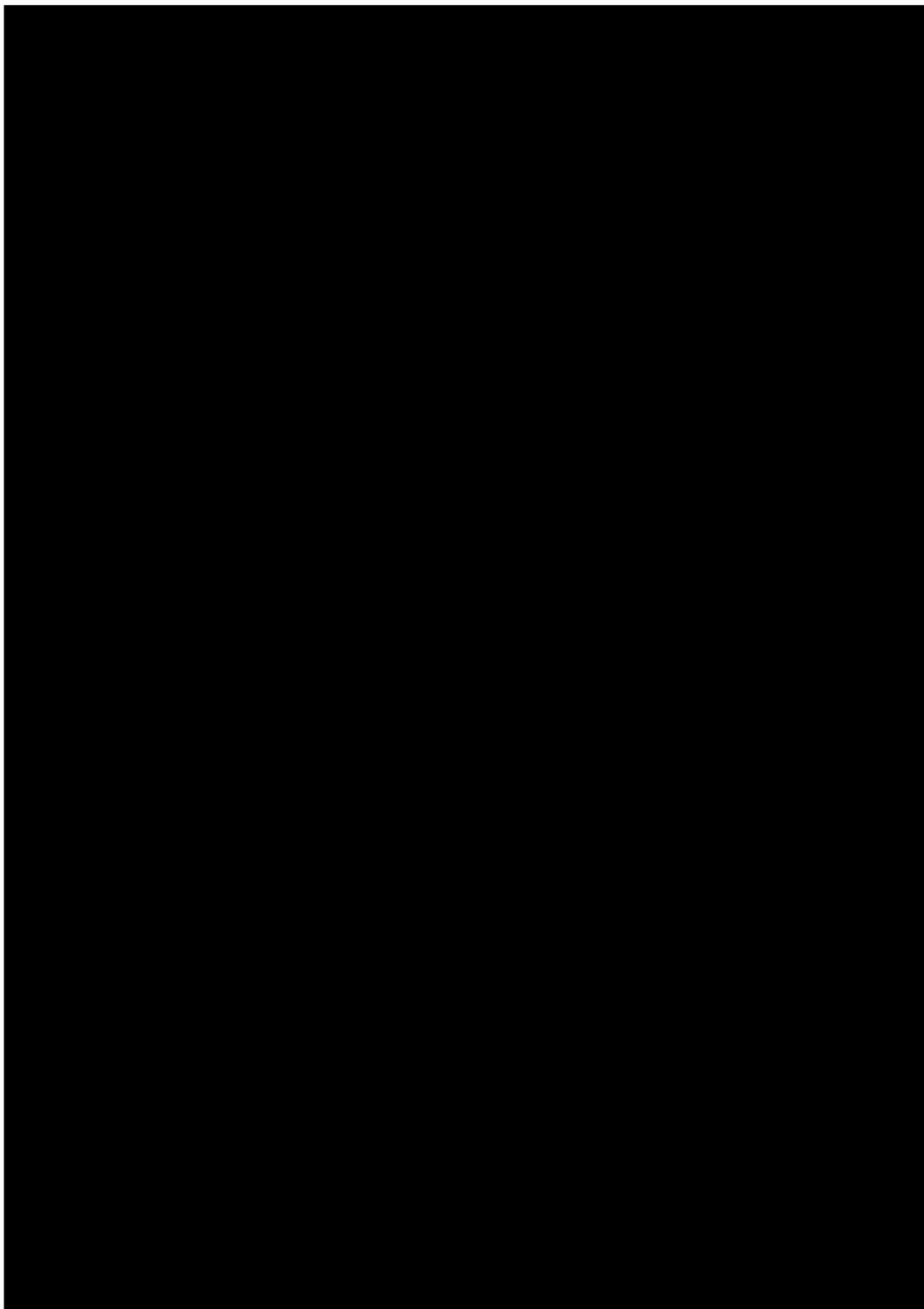
Transcription: y(e)m k(a)y(a)d(a) y(a)r(i)lu önmiš.

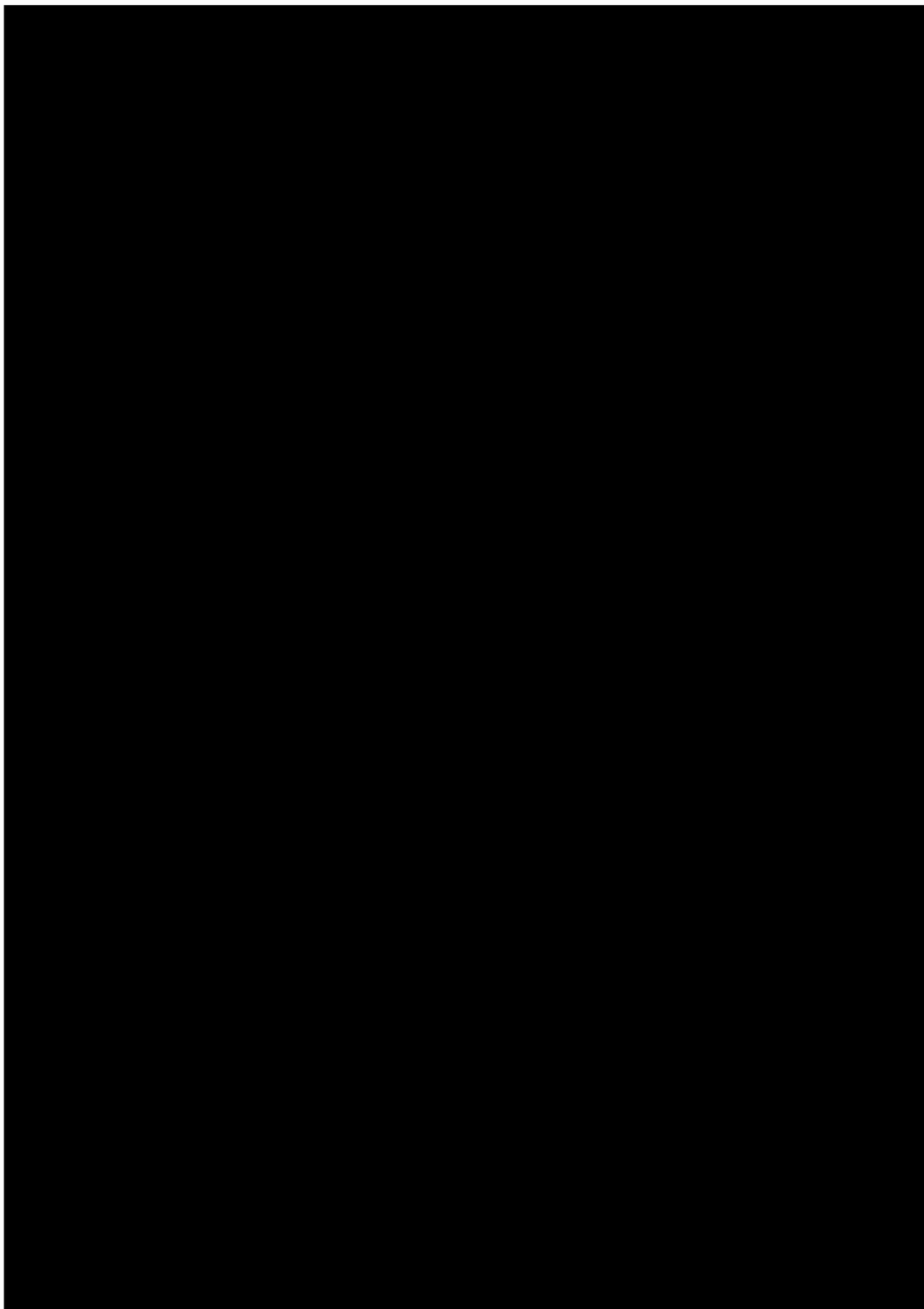
Translation: ‘The food (a plant) on the rock, has, apparently, split the rock and grown up.’

### 3.2 3D analysis and a new reading of the inscription

The area of the rock surface with the inscription is 0.03 m<sup>2</sup>. To cover this surface with the vertical and horizontal shifts of the camera (ensuring 70% and 60% covering of the previous photo respectively), 40 photos were made. The field fixation took 0.3 hours, and the subsequent computer processing in the program Agisoft Photoscan Pro took 8 hours. Vavulin et al. 2018 published the results of our 3D







*Reading by Tybykova and Nevskaya (Vavulin et al. 2018)*

Runes:  $\wedge \times \text{H} \text{H} \text{Y} \text{Y} \text{H} : \text{H} \times \text{D} \text{H} \times \text{J} \text{Q}$

Transliteration:  $y^2 l^1 m k^1 y^1 d^1 n^2 : y^2 r^1 l^2 U \ddot{U} n^2 m \text{š} :$

Transcription:  $y(a)l(\ddot{i})m k(a)y(a)d(a)n : y(a)r(\ddot{i})lu \ddot{u}n(m(i)\text{š}) :$

Translation: From the bare rock, they, having split (into smaller groups), descended.

Comments:

1. The 3D analysis has shown that there is no rune  $\text{J}$  A, as presupposed by Konkobaev et al. (2015), because the upper side line is a natural splitting of the rock which can easily be mistaken for a line without a 3D analysis. The three signs  $\times \text{J} \text{Q}$  ( $y^2 l^1 m$ ) can be read as  $y(a)l(\ddot{i})m$ , although the first rune should have been  $y^1$ , not  $y^2$ , and the harmonious type of this consonant is therefore wrong. This happens quite often in this region (Nevskaya 2011), especially with the two types of  $y$ .

2. The last sign before the first word divider ( $\text{H}$ ) is  $n^2$ . This is the most problematic sign in the inscription, even with the results of the photogrammetry, because the depth of one line was not typical for this inscription; it is marked with the broken line on the copy. However, it is neither a natural splitting of the rock, nor graffiti, so we have assumed that it is a part of the sign, but was made on a place with some hard inclusions in the rock, so it could not be incised more deeply.

3. The 4 signs  $k^1 y^1 d^1 n^2$  ( $\text{H} \times \text{D} \text{H}$ ) are read as  $k(a)y(a)d(a)n$ , which is logical here, as the verb  $\ddot{u}n$ - can govern the ablative case. The unwritten vowel of the ablative case could be either  $a$ , or  $\ddot{i}$ . Erdal (2004: 175) observes that a low vowel is more frequent in runiform inscription; therefore we prefer to transcribe it as  $a$ .

The second consonant of the ablative is  $n^2$ , not  $n^1$ , as one could expect; this could be an instance of the symbolic writing of affixes, quite typical for the runiform script where a row of affixes have the same graphical form regardless of the harmonic class of the word they are attached to.

4. The phrase *yalim kaya* is very frequent in runiform inscriptions; compare Tuekta-IV, originally situated on a separate small stone, just one hundred meters away from Tuekta-V (Tybykova et al. 2012). It is written with the correct  $y$ ; so we can suppose that two different scribes made these two inscriptions:

$b(\ddot{a})\eta(i)g\ddot{u} : y(a)l(\ddot{i})m k(a)y(a) : b(i)t(i)g ur(a)y(\ddot{i})n$

‘This is an eternal bare rock. Let me make an inscription (on it)!’

The inscription Tuekta-IV was brought by us to the Gorno-Altaysk National Museum, when a road construction project threatened its existence.

5.  $y^2 r^1 l^2 U$  ( $\text{Y} \text{Y} \text{H} \text{Q}$ ), read as  $y(a)r(\ddot{i})lu$ , is considered to be a vowel converb  $-U$  of the verb *yaril*- ‘split’ in Vavulin et al. (2018). The VATEC database confirms the labial variant of the vowel converb; the aorist of this verb is also  $-Ur$ , and the aorist and vowel converbs always have the same vowel in Old Turkic. Clauson (EDPT: 955) mentions this word with both back and front consonants, treating the two harmonic shapes as variants of one word. Thus, this sequence could also be read as

*y(e)r(i)lū*. Erdal considers *yaril-* and *yeril-* not to be variants, but different verbs; one is dealt with in OTWF: 684, the other in OTWF: 686, and they have different meanings. *yeril-* ‘be separated from’ seems to be more suitable here. Thus, we should correct the reading of this word.

6. The last word, *ünm(i)š*, is *ün-* in the indirective form *-mlš*. Nadeljaev et al. (1969) and Clauson (1972) give different transcriptions of this word—with *ö* and *ü* respectively; this verb governs the ablative case and means ‘to get out, to rise, to grow’: *ol baliqtin öntüm* ‘I went out of that city’ (DTS 1969: 385). OTWF: 624 quotes a Brâhmî source, which documents the vowel of this verb undoubtedly as /ü/. Thus, the quoted sentence should be corrected as *ol baliqtin üntüm* ‘I went out of that city’.

We suppose that in this inscription, the verb should be understood as a verb of motion, as it has a spatial argument in the ablative case form. Then, the inscription means that a group of people got separated and went out of (a location), or ascended (a mountain), in contrast to the proposed reading in Vavulin et al. 2018.

*New reading proposed here:*

Transcription: *y(a)l(i)m k(a)y(a)d(a)n : y(e)r(i)lū ün(m)iš* :

Translation: From the bare rock, they, having got separated (from the rest), ascended (a mountain) / got out (of a location).

The interpretation proposed by Konkobaev et al. (2015) is also possible: something (a tree?) grew up out of the bare rock, having split it.

Regardless of these and further possible interpretations of this inscription, the correct composition of the signs is the most important prerequisite for its reading. This was achieved only with the help of the 3D technologies.

An example of the significant data that Altai runiform inscriptions can contribute to our understanding of the history of religious traditions in this area is provided by the inscription Bichiktu-Boom-III (presupposing that our interpretation is correct).

#### 4. Bichiktu-Boom III

##### 4.1 Inscriptions with religious content in the Altai Mountains

The religion of the ancient Turks who created the Old Turkic runiform monuments in South Siberia is a disputed question. Nevskaya (2015) has summarized the viewpoint of Igor Kyzlasov, who considers practically all Siberian runiform inscriptions to be Manichaean; she analyzed some of his questionable readings and presented linguistic facts that mirror religious beliefs of the creators of Old Turkic inscriptions in Siberia. Here, we briefly present Igor Kyzlasov’s view.

In his works, Igor Kyzlasov states that Old Turkic runiform writing as a whole “has religious roots, as actually any writing does” (Kyzlasov, I. L. 2015).

According to him (Kyzlasov, I. L. 2001), archaeological monuments of the Sayan-Altai Mountains bear witness to two spheres of ritual usage of the runiform

writing in the ancient Kyrgyz state—epitaphic inscriptions on steles and prayers written on rocks. However, the tradition of erecting steles facing the east or south-east was much older; it was connected with the Tashtyk culture of the 1st century BC (see Kyzlasov, L. R. 1960, 1965; Kyzlasov, I. L. 1975; Grjaznov 1979: 128-142) where stone steles symbolized the deceased. “Later such steles were put in front of stone mausoleums called kurgans. An epitaph was the last component of such memorials; it was carved on the already erected stele, as the forms of the lines and the location of the signs show” (Kyzlasov, I. L. 2015).

Following the views of his father, L. R. Kyzlasov, Igor Kyzlasov supposes that the spreading and use of the runiform script in the Altai Mountains was connected with spreading of Manichaeism in medieval Turkic states (Kyzlasov, I. L. 2001, 2004, 2015; compare to Kyzlasov, L. R. 1998, 1999, 2001). According to him, Manichaeism became the successor of an unknown ancient religion of the Turks that may already have used the runiform writing for sacral purposes. The authority of this ancient religion and its writing was overtaken by Manichaeism. He states that numerous Old Turkic runiform inscriptions on steles appeared in South Siberia in the late eighth century. This is also the dating of brick temples found there. He further writes, “A comparative analysis of runiform texts written on steles found in this area and identified as Manichean proves that these temples should also be considered Manichean ones.” (Kyzlasov, I. L. 2015). According to him, the content of rock inscriptions is almost exclusively Manichaean (Kyzlasov, I. L. 1994. Chapter 5, Section 4; 2001).

In fact, this assumption is based on a very problematic reading of a religious term found in the basin of the river Yenissei, namely the word *äzrua* ‘Zerwan’ in E-29,3 (Kormušin 1997: 67–75, 2008: 120–121), which the interpreter himself considered questionable<sup>1</sup> as the inscription is very damaged.<sup>2</sup> It is obvious that 3D documentation methods are necessary for a more reliable reading of this inscription. In this connection, Nevskaya (2015) has noted that in Old Turkic texts in the database VATEC, the word *äzrua* ‘Zerwan’, also meaning ‘Brahma’, is encountered 36 times, most of which are in Buddhist texts including Altun Yarok and Maitrisimit (<http://>

1 In our personal communication, Marcel Erdal commented on this the following: “What he interprets as *äzrua* is *zr<sup>l</sup>w*. The inscription was edited in the 1970s, 1980s and 1990s by Kljashtornyj, Mori and Talat Tekin, and none have such a reading but they are, of course, all based on Radloff/Malov. Note that Kormushin (1997) has conditional *ärsäm* before that, which is quite impossible for Old Turkic. Instead of this word, Tekin writes *atīm* ‘my name’, which is compatible with what Thomsen and Wulff saw. The W. Thomsen / K. Wulff’s text, based on original material, has *nr<sup>l</sup>w* instead of these 3 characters, which could be *aṇaru* (attested also in Tonyukuk) / *īṇaru* (attested also in Šine Usu) ‘to that’. But of course *zr<sup>l</sup>w* is not yet *äzrua*, as final vowels are supposed to be written, and as the context is impossible.”

2 The authors could check this place using contemporary photos of this inscription which were kindly sent to them by Igor Kormushin.

vatec2.fkidg1.uni-frankfurt.de/). See also a definition of this term and examples of its use in Old Uyghur in UW (2000). Even supposing that this term had been used in one of Yenissei inscriptions, just one single usage of a religious term that could be Manichaean or Buddhist does not seem to be a solid foundation for such far-reaching conclusions.

Nevskaya (2015) referred to a more reliable fact of the usage of a religious term in runiform inscriptions found in the Altai Republic. In one Old Turkic runiform inscription, on the rock Kalbak-Tash, which contains thousands of petroglyphs, graffiti of the Old Turkic period, Old Turkic *tamga* signs, and a remarkable collection of more than thirty Old Turkic runiform inscriptions, we find an instance of the word *ṛṣi* (the inscription Kalbak-Tash VIII (A-28)), see Tybykova et al. 2012. The use of this word is extremely interesting for Old Turkic studies. We have not encountered it in the runiform writing system before; see the word *arži* in Röhrborn's dictionary (UW, VI: 221). In the VATEC database, we encounter two different writings of this word in two different writing systems: the Sogdian and the Uyghur ones. The Sogdian script has *ṛṣ'yl'r* and the Uyghur script *'rṣyl'r*. However, is it proof that the ancient Turkic population at the time of creation of those inscriptions had Manichaean religion? It is hard to say so. Of 13 instances of *arži* in the VATEC database, all have appeared in Buddhist texts, including *Maitrisimit* and *Huanzang's biography* (which might be rooted in a much higher percentage of preserved Old Turkic Buddhist texts, both overall and in the database VATEC). The word is, of course, connected with religious contents; however, it cannot be associated only with Manichaeism (Nevskaya 2015).

Our recent 3D analysis of a well-known inscription Bichiktu-Boom-III has yielded a new understanding of the composition of signs in this inscription, which has resulted in a tentative new reading providing astonishing evidence of a further religious term found in Old Turkic runiform inscriptions in this area.

#### 4.2 Previous readings of Bichiktu-Boom III (A-16)

This inscription is situated to the south of the village. It is made on a small rock plane 65.0x52.0 cm facing south, almost at the foot of the rock. It was discovered by T. I. Sabaeva in the 1950ies, and studied by K. Sejdakmatov in 1964, by E. I. Ubrjatova in 1974, by D. D. Vasil'ev in 1980, by I. L. Kyzlasov in 1994, by I. A. Nevskaya and L. N. Tybykova in 2003–2005, and also by M. Erdal in 2006.

The inscription has two vertical lines. Its number in the DTS is A8b. The inscription got the name “inscription to the right” when it was first published by K. Sejdakmatov. E. I. Ubrjatova saw 43 signs, 40 in the first line and 3 in the second. According to the reading by Tybykova, Nevskaya and Erdal (Tybykova et al. 2012), the inscription has 38 characters: 3 in the first line and 35 in the second line. The length of the first line is 4.5 cm; the length of the second line is 45.0 cm; the height of the characters is 1.0–3.2 cm. The beginning of the inscription is only 6.0 cm from the earth. It is possible that the scribe made while lying on his side. The

inscription consists of two vertical lines that are read from bottom to top, (respectively, right to left).

*Reading of K. Sejdakmatov*<sup>3</sup>

Transliteration:

1) Ёб Гб Ы

2) Ёб Гб Тб М Ш А Кб Ш Лб Гб А Кб Гб Лб Ы А Тб Гб А Сб Нб Ы А Н ч М Дб Ё  
Нб А

Кб Ы Кб З Дб

Transcription:

1) *йегу*

2) *йег этмиш а киши элиг а кэ эли а тег а сени анчам эд й не кик эд*

Translation:

1) His goodness

2) ... did good (things). Man, people ... you so much....

(Sejdakmatov 1964: 101)

*Reading of E.I. Ubrjatova* (from top to bottom)

Transliteration:

$d^2 z t l t q'' : n^2 j \gamma m n\check{c} q l \check{s} y q n q b r q l l^2 k^{\ddot{u}} \ddot{u} s^2 \ddot{a} g^{\ddot{a}} l^2 \check{s} t q l b i l p k^{\ddot{u}} j^2 l j^2$

Transcription:

$\check{a}d^{\check{z}} z tol t^{\check{u}} q, \check{a}n j^{\check{a}} \gamma^{\check{y}} m^{\check{i}} n\check{c} q^{\check{y}} l, \check{s}yq, {}^y n^{\check{a}} q b^{\check{a}} r q^{\check{y}} l, {}^{\check{o}} l k\ddot{u}s\check{a}g^{\check{i}} \check{c}, \check{s}at q^{\check{y}} l, b^{\check{a}} j^{\check{a}} p k^{\check{o}} j^{\check{a}} l k^{\check{o}} j$

Translation:

Get full to the top, the abundance,

My main enemy, make peace;

Fortune—friendship, come about (lit.: make it present);

Wish (at least) the first degree of saturation.

Bring about satisfaction;

Rich, strong Köy (abundant?);

Holy Köy (Abundant?)

(Ubrjatova 1966: 161)

*Reading of Ė. R. Tenišev*

Transcription:

1) *jäg aŋ*

2) *jäg ätmiš ä käš älig käg äli ä ätig ä säni ä ança ma (?) jan<sup>2</sup>a kik azad<sup>2</sup>*

1) Good hunting,

2) Organized well, a quiver, 50 (?), nobility, detachment, you, so (?), hostility decreased.

(Tenišev 2006: 140–143)

<sup>3</sup> We use the transliteration and transcription systems of each author of the presented readings; however, we translate their Russian versions into English.

Transliteration:

2)  $j^2 g t^2 m \ddot{s} \ddot{a}(a) k \ddot{s}^2 l^2 g \ddot{a}(a) k^2 g l^2 i(\ddot{i}) \ddot{a}(a) t^2 g \ddot{a}(a) s^2 n^2$

i(ï) ä (a)nčmu(o)d<sup>2</sup>j1n<sup>2</sup> ä(a)ki(ï)kzš<sup>2</sup>

1)  $j(e)g(i)$

2) j(e)g(i) (e)tm(i)š k(i)š(i)l(i)g ekg(ä)li teg (i)s(i)ni inč(?) (a)mul (?) j(i)n (i)ki  
k(e)z(i)š

1) Follow the goodness.

2) Having done goodness, in order to spread humanity, having got encouraged, look for humility (= quietness) and together (with me, with God in one's thoughts), wander!

Reading by L. N. Tybykova, I. A. Nevskaya and M. Erdal (2012)

[illegible]
$$1) \gamma^2 g^2 I$$

2)  $y^2 g^2 t^2 m \dot{s}^1/t^2 A k^2 \dot{s}^1 l^2 g^2 A s^2 \eta^2/k^2 g^2 l^2 I A? t^2 g^2 A s^2 n^2 I A n\check{c} m U d^2 y^1 n^2 A \eta^2/n^1/k^2 I \eta^2/k^2 z$

1) y(e)gi

2) y(e)g (e)tm(i)š k(e)šl(i)g s(ö)kg(ä)li täg! (ä)s(ä)ni (e)nč mu? ....

1) his/their wellness

2) (You) who have beautiful (beautifully made) quivers, attack (the lines of the enemies), in order to break them apart! Is their mind quiet, at rest?

Transcription:

1) j(e)g (i)g(ä)

2) j(i)g(i)rm(i) (o)q / (e)ŋ (a)š (e)l(i)g / s(a)ŋ(i)g (e)l(i) (e)t/d (e)t(i)g s(e)ni (e)r  
(a)nčm(a) (e)d(e) (a)ja (a)nšin (i)d(i)m

1) the good ruler/owner

2) Twenty arrows. First of all—(there is) a land of fifty monasteries, prosperity, well-being. Oh, how much I have paid tribute to your memories, adorning you, hero, by goodness!



*Reading by Konkobaev et al. (2015: 54)*

Transliteration:

1) ı/i g<sup>2</sup> y<sup>2</sup>

2) z<sup>2</sup> k<sup>2</sup> ı/i k<sup>2</sup> e/a n<sup>2</sup> y<sup>1</sup> d<sup>2</sup> m nç e/a ı/i n<sup>2</sup> s<sup>2</sup> e/a ı/i l<sup>2</sup> g<sup>2</sup> k<sup>2</sup> s<sup>2</sup> e/a g<sup>2</sup> l<sup>2</sup> ş<sup>2</sup> k<sup>2</sup> e/a ş<sup>2</sup> m t<sup>2</sup> g<sup>2</sup> y<sup>2</sup>

Transcription:

1) y(e)gi

2) y(e)g(e)tm(i)ş k(e)şl(i)g s(ö)kg(e)li t(e)g (e)sni (ı)nç(ı)m (e)d (a)y(ı)n (e)ki k(e)z

Translation:

1) well-being

2) (Warriors) with glorious (bringing victory) quivers, in order to defeat (the enemies), attack (them). Remember (take into consideration). (My) worried ones, be frightened two times.

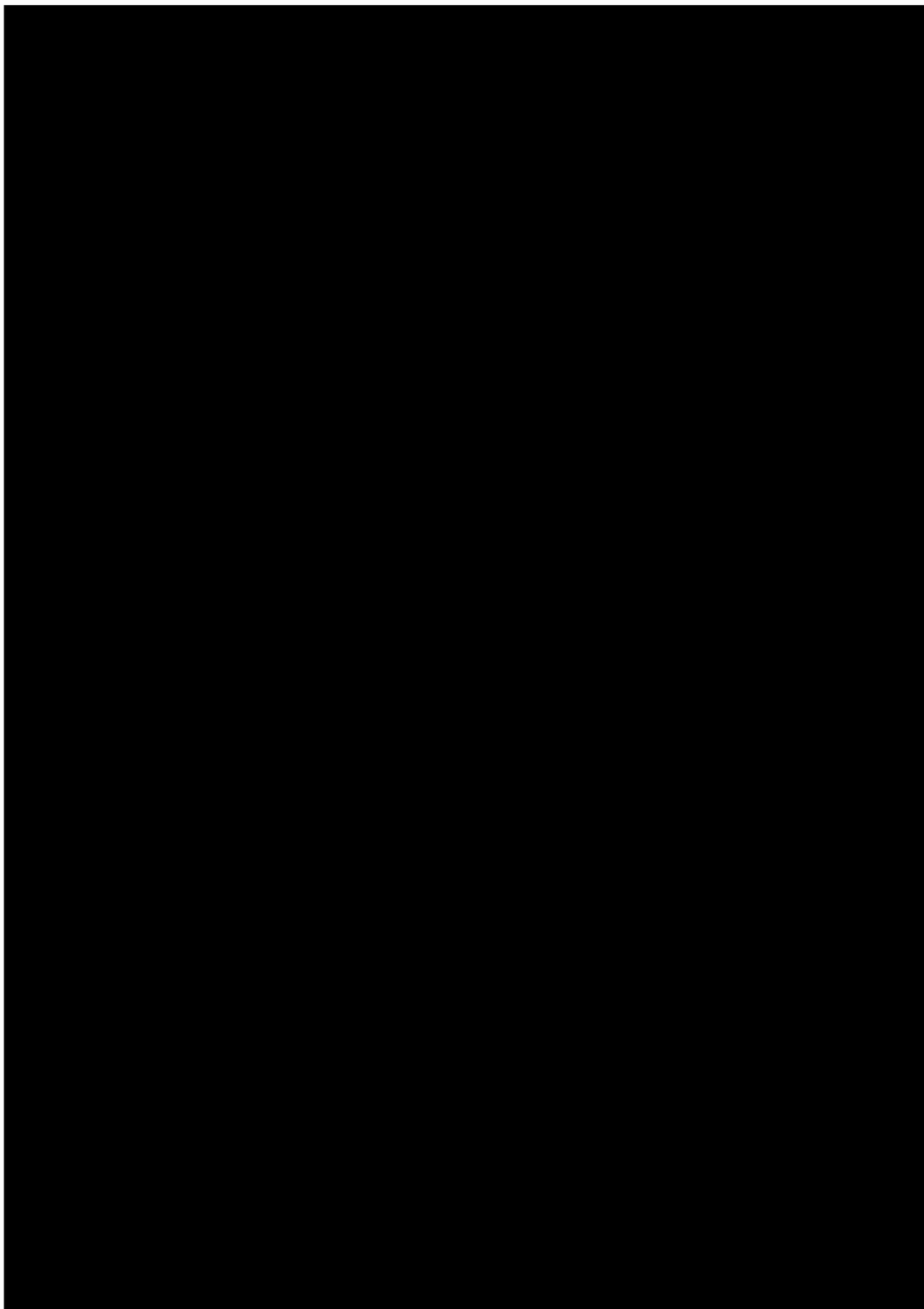
We do not analyze each of these readings because, quite obviously, they are based on how each author interprets and identifies the signs of the inscription. It is enough to say that none of the readings, including the one proposed by the authors of this article, gives a coherent and comprehensive understanding of this inscription.

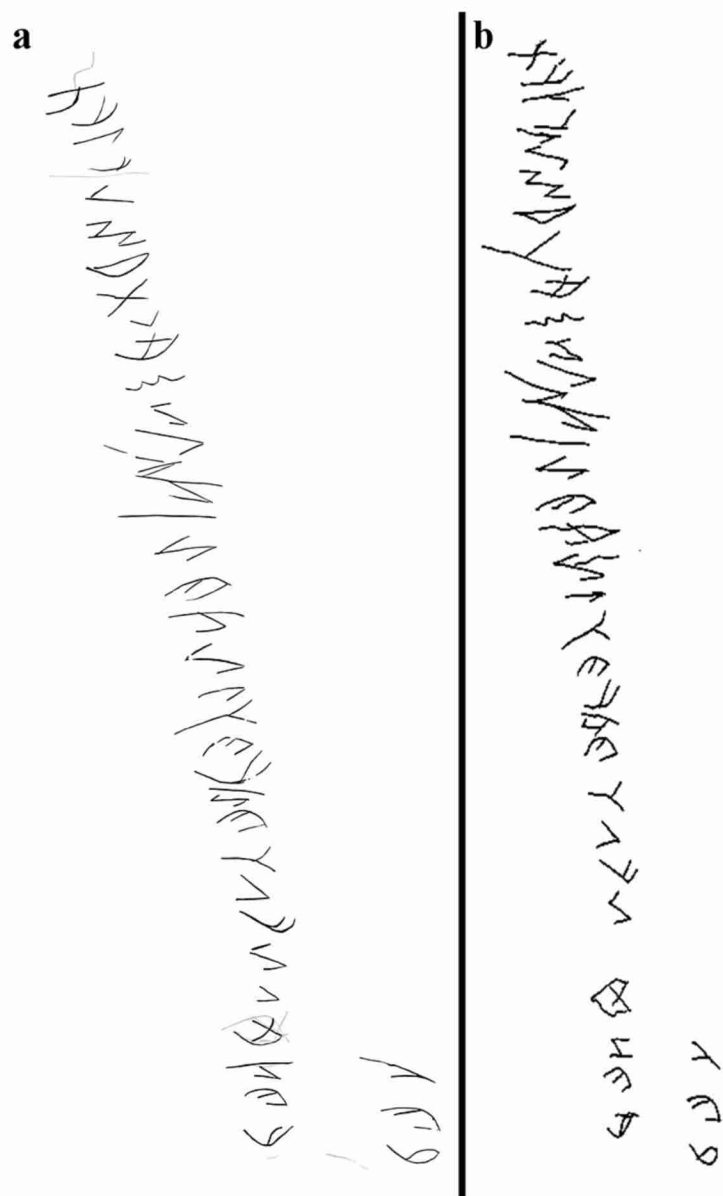
#### 4.3 3D analysis and a tentative new reading of the inscription

Several types of lines on the rock surface with this inscription were discovered using 3D technology (Vavulin 2017).<sup>4</sup> In most cases, the runiform lines can easily be identified; they are not more than 5.0 cm long, 0.6 mm wide and 0.06–0.12 mm deep. They are considerably lighter than the background rock surface (Picture 5, a). The second type of lines shares most characteristics with the first one. These lines have the same length, width and color. However, they are much shallower, about 0.03–0.06 mm deep. It is possible that all or some of them also make up part of the runiform inscription (Picture 5, a). Lines of the third type have the same width, depth and color as the first one, yet they are considerably longer, sometimes reaching 14 cm (Picture 5, b). The fourth type differs sharply from the previous three. The arch-shaped lines are 3.0–12.0 cm long, 0.5–1.0 mm wide, and 0.01–0.05 mm deep. Their color is practically the same as that of the background rock surface (Picture 5, c). Finally, lines of the fifth type represent scratches rather than carvings. They are easily distinguishable due to their lighter color; they have almost no depth at all. Their sizes and forms vary (Picture 5, d).

Lines of each type were rasterized individually in relevant vector layers. The resulting image contributes a great deal to the rasterized data from previous surveys (Picture 6).

<sup>4</sup> Vavulin (2017) presents a 3D data analysis of this inscription done by him. Pictures 5 and 6 were also first published in that article.





Picture 6. Copies of the inscription: 6a (Vavulin 2017), 6b (Tybykova et al. 2012)

The new copy of this inscription has become the foundation for our revised reading of the inscription.

## Transliteration

1)  $y^2 g^2 I$ 2)  $s^1/y^2 g^2 t^2 m \text{ š } A k^2 \text{ š } l^2 g^2 A s^2 \eta g^2 l^2 I A t^2 g^2 A s^2 n^2 I A n\check{c} m U d^2 y^1 n^2 l^1 k^2 I$   
 $s^1/\eta z/t^2$ 

## Transcription

1)  $y(e)gi$ 2)  $(a)s(i)g / y(e)g (e)tm(i)\text{š } A k(i)\text{š}(i)l(i)g A s(a)\eta(i)g (e)li A t(\ddot{a})g A (\ddot{a})s(\ddot{a})ni A$   
 $(a)n\check{c}(a)m(a) ud(u)y(u)n (a)lk\ddot{i}\eta(i)z / (a)lk\ddot{i}\text{š } (e)t!$ 

## Translation

1) Its prosperity (lit.: its goodness, i.e. what is good for it).

2) Similar to those people belonging to the monastery community consisting of (lit.: with) individuals making contributions (or doing good deeds), following (them) bless/praise its well-being beyond limits (lit.: to such an extent)!

## Comments:

1. In this new reading, the first line appears to be the title of the rest of the inscription. Its content is explained in the second line: what is good for it (the monastery community).

2. Almost all Turcologists who worked on this inscription saw the first character of the second line as  $y^2$ . However, it is very clearly visible on the orthophoto that the loop in the upper part of the sign has not been closed. There is another vertical line that could be taken for the missing part of the loop. The author of the inscription probably also saw the character as completed. If that is the case, we can read the first 5 characters of the second line as  $y(e)g (e)tm(i)\text{š}$  ‘doing good (deeds)’. If we follow the signs that are there, without trying to reconstruct the intentions or ideas of the scribe, we have to read this part as  $(a)s(i)g (e)tm(i)\text{š}$  ‘making use/profit/contributions, i.e. being useful’ because the open loop of  $y^2$  gives us the character  $s^1$ . The word *asġg* ~ *asag* is encountered in VATEC 72 times, in both Buddhist and Manichaean texts.3. We read characters 12–14 of the second line  $s^2 \eta g^2$  as  $s(a)\eta(i)g$  ‘belonging to monastery community’, following in this Vasilev’s reading of this inscription (2013: 87). Together with the previous and subsequent characters, we get ‘similar to the people belonging to the monastery community consisting of individuals making contributions (or doing good deeds)’.4. We understand the word *el-i* standing after the word ‘belonging to monastery community’ as referring to the people of the community, not land or country, as Vasilev (2013: 87) does. The meaning of *el* ‘people, nation’ is quite frequent in runiform inscriptions.5. The next word *täg* is a simulative postposition ‘like, similar to’.6.  $(\ddot{a})s(\ddot{a})n-i$  could be the direct object of the verb *alka-/alki-* ‘praise, bless’ (or *alkiš et-*, see comment 8), while  $(a)n\check{c}(a)m(a)$  is its degree adverb: to praise its (com-

munity's) well-being (prosperity) to such a (high) degree. In reading  $(a)nč(a)m(a)$ , we again follow Vasilev's proposal (2013: 87). Because *äsän* has a possessive suffix of the third person, the accusative of it should have been *äsänin*. In case *äsäni* is not the direct object of the verb 'to praise', it could be an interjection 'Its well-being (should be blessed)! ~ Let it be safe and sound!'; compare Erdal (2002, footnote 27) on the use of this word in Yenisei inscriptions. This assumption is supported by the fact that *äsäni* is bracketed by two runes A serving as punctuation in Altai runiform inscriptions (Nevskaya 2011). In this case, the interpretation of the inscription would be: Similar to those people belonging to the monastery community consisting of (lit.: with) individuals making contributions (or doing good deeds),—let it (the community) be safe and sound!—following (them) bless/praise (it) beyond limits (lit.: to such an extent)!'

7. The converb *ud(u)y(u)n* has a rather rare converb affix  $-(X)yXn$ , found in *iyin* 'following' (contracted through haplology from *iy-iyin*) in Maitrisimit, *tilä-yin* 'wishing' in Xuanzang-Biography (VATEC data).

8. The last word,  $(a)lk̄ñ(ī)z$ , could be an irregular imperative form  $-(X)ñ(X)z$  of the 2nd person plural of the verb *alka-/alki* 'praise, bless'. According to Erdal (2004: 237), in some Old Turkic texts,  $-(X)ñ$  is exclusively used for polite address in the singular, and  $-(X)ñlAr$  for plural addressees. However, DLT fol. 289 quotes a verse with the second person plural imperative forms *koymañiz* and *k̄ymañiz* corresponding to what in his language would be *kodmañ* and *k̄idmañ*, saying that this is how the Oghuz and Kipchak tribes form this imperative. The forms of 2nd person plural imperative, in use in Oghuz languages until now, are constructed by analogy with personal pronouns and possessive suffixes, whereas the  $+lAr$  of other Turkic languages comes from nominal inflexion (Erdal 2004: 237). If this reading is correct, can we probably speak of a dialectal feature in this inscription? This assumption seems to be very problematic, since these forms are practically restricted to Oghuz, and the assumption of an Oghuz feature in an Old Turkic inscription in Siberia could be too far-fetched.

The signs  $ñ$  and  $s^l$  are written in this inscription in a very similar way. Moreover, the last character could be read either as  $z$  or  $r^2$  because the upper section of the vertical line going from left to right in  $z$  is problematic. Thus, we could suppose that the sequence  $l^1 k^2 I s^l/ñ z/r^2$  can be read either as  $(a)lk̄ñ(ī)z$  or  $(a)lk̄š(e)t$ . This second version would refer to one addressee: 'Similar to those people belonging to the monastery community consisting of individuals making contributions (or doing good deeds), following (them) bless/praise its well-being beyond limits (lit.: till such extent)!'

Irina Nevskaya has recently had an opportunity to discuss this inscription with Dieter Maue, who considered whether this inscription could refer to an addressee who should obtain well-being for himself in a way the people belonging to the monastery community did when they had joined it. To acquire new members for the community would also be a great merit for the scribe of this inscription himself.

Following this idea, we could think of the following interpretation of this inscription. This reading would presuppose a usage of the similative postposition *täg* as the first part and *ančama* (or *anča* plus the particle *mu*) as the second part of a conjunction ‘similarly to ... likewise ...’. The word *äsän-i* could be a direct object of the verb *ud-*: ‘follow the well-being (of the community members) and do favor/acquire blessing (to yourself)’. The reading of the inscription would be: ‘Similar to the people belonging to the monastery community consisting of individuals making contributions (or doing good deeds), you likewise, through obtaining (lit.: following) their well-being, obtain blessing (to yourself)!’ In case *äsäni* is an exclamation and not a direct object, the interpretation would be: ‘Similar to the people belonging to the monastery community consisting of individuals making contributions (or doing good deeds), let it (the community) be safe and sound, you likewise, following (them), obtain blessing (to yourself).’

Thus, another religious term seems to be found in this inscription, supporting the idea of Kyzlasov that there was a monastery in the vicinity of the village called Bichiktu-Boom. We encounter numerous inscriptions here; this is the reason why the Altai people gave this name to the village and the location around it. Bichiktu-Boom in Altai means ‘a mountain with inscriptions’. We can only guess how many inscriptions were on the rock originally, at the time when the monastery existed in this area. But what was the religion of the supposed monastery community? The word *saŋig/saŋik* is found in the VATEC database two times in the Buddhist context. Thus, the question of the religious beliefs of the ancient Turkic population of this area remains unanswered. However, it is quite clear now that this area used to be affected by the spreading of world religions already in the Old Turkic period.

## 5. Conclusion

Thirty Altai runiform inscriptions of the Old Turkic era were documented in 2017 by 3D technology. Photogrammetry proved, in this case, to be not only the cheapest solution but also the optimal choice. The resulting data represents a first attempt at runiform monument digital documentation, which will help solve long-lasting disputes about specific runes in rock inscriptions. It will also make it possible to preserve ancient written monuments in their current state, at least in digital form. As soon as they are posted online, they will be accessible to all Turcologists around the world, thereby giving a new impetus to their deciphering. In addition, these efforts will bring the long-term benefits of using 3D recording for the preservation and popularization of archaeological and epigraphic heritage in Russia, where such projects are extremely rare today (Plets et al. 2012; Zaytseva 2014; Plisson & Zotkina, 2015; Pushkarev et al. 2016; Zaytseva 2016 ).

Our further plans include the following:

To document with this technology most Altai runiform inscriptions, ideally, all of them.

To publish the results on the Internet in open access format.

To also publish 3D models, orthophotos and heightmaps.

To prepare a new edition of the *Catalogue of Altai runiform inscriptions*.

To foster further research on Altai runiform inscriptions and non-canonical inscriptions elsewhere.

### Acknowledgements

The authors greatly appreciate the support and insightful and valuable remarks of Marcel Erdal, with whom we have discussed this article at different stages of our work on it.

We are also very grateful to the financial support of our work through Ministry of Education and Science of the Russian Federation, Project Number 33.8576. 2017/9.10. The project is headed by Anna Dybo.

### References

- DLT = Dankoff, R. & Kelly, J. = Maḥmūd al-Kāšġarī, 1982–1985. *Compendium of the Turkic dialects (Dīwān Luyāt at-Turk)*. Turkish Sources 7: 1–3. Cambridge, Mass: Harvard University.
- DTS = Nadeljaev, V. M. & Nasilov, D. M. & Tenišev, Ė. R. & Ščerbak, A. M. (eds.) 1969. *Drevnetjurkskij slovar'* [An Old Turkic dictionary]. Moscow: Nauka.
- EDPT = Clauson G. Sir 1972. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon.
- Erdal, M. 2002. Anmerkungen zu den Jenissei-Inschriften [Comments to the Yenisei inscriptions]. In: Ölmez, M. & Raschmann, S.-Chr. (eds.) *Splitter aus der Gegend von Turfan. Festschrift für Peter Zieme*. Istanbul & Berlin: M. Ölmez. 51–73.
- Erdal M. 2004. *A grammar of Old Turkic*. Leiden & Boston: Brill.
- Grjaznov, M. P. 1979. Taštykskaja kul'tura [Tashtyk culture]. In: *Kompleks arxeologičeskix pamjatnikov u gory Tepsej na Enisee*. Novosibirsk. 128–142.
- Konkobaev, K. & Useev, N. & Šabdanaliev, N. 2015. *Atlas drevnetjurkskix pis'mennyx pamjatnikov Respubliki Altaj* [An atlas of Old Turkic written monuments in the Altai Republic]. Astana: International Turkic Academy.
- Kormušin, I. V. 1997. *Tjurkskie enisejskie ėpitafigi. Teksty i issledovanija* [Turkic Yenisei epitaphs. Texts and studies]. Moscow: Nauka.
- Kormušin, I. V. 2008. *Tjurkskie enisejskie ėpitafigi. Grammatika i tekstologija* [Turkic Yenisei epitaphs. Grammar and textology]. Moscow: Nauka.
- Kyzlasov, I. L. 1975. Pominal'nye pamjatniki taštykskoj ėpoxi [Epitaphs of the Tashtyk time]. *Sovetskaja Arxeologija* 2, 200–208.
- Kyzlasov, I. L. 1994. *Runičeskaja pis'mennost' evrazijskix stepej* [Runic writing of the Eurasian steppe]. Moscow: Vostočnaja literatura.
- Kyzlasov, I. L. 2001. Smena mirovozzrenija v Južnoj Sibiri v rannem srednevekov'e (idei edinobožija v enisejskix nadpisjax) [A change in the worldview in South Siberia in early medieval times (the idea of one god in Yenisei inscriptions)]. In: *Drevnie civilizacii Evrazii. Istorija i kul'tura*. Moskva. 243–270.
- Kyzlasov, I. L. 2002. *Pamjatniki runičeskoj pis'mennosti Gornogo Altaja 1*. [Written runic monuments of Mountainous Altai 1.]. Gorno-Altajsk: Gorno-Altajsk State University.

- Kyzlasov, I. L. 2004. Manixejskie monastyri na Gornom Altae [Manichaean monasteries in the Altai Mountains]. In: *Drevnosti Vostoka. Sbornik k 80-letiju professora L. R. Kyzlasova*. Moscow. 111–129.
- Kyzlasov, I. L. 2015. Religioznaja priroda enisejskoj pis'mennosti (pervoe približenie) [Religious origin of the Yenisei writing (the first approach)]. In: Nevskaya, I. & Erdal, M. (eds.) *Interpreting the Turkic runiform sources and the position of the Altai corpus*. Berlin: Klaus Schwarz.
- Kyzlasov, L. R. 1960. Novaja datirovka pamjatnikov enisejskoj pis'mennosti [New dating of Yenisei written monuments]. *Sovetskaja Arxeologija* 3, 152–161.
- Kyzlasov, L. R. 1965. O datirovke pamjatnikov enisejskoj pis'mennosti [Dating Yenisei written monuments]. *Sovetskaja Arxeologija* 3, 38–49.
- Kyzlasov, L. R. 1998. Severnoe maniyejstvo i ego rol' v kul'turnom razvitii narodov Sibiri i Central'noj Azii [The Northern Manichaeism and its role in the cultural development of the peoples of Siberia and Central Asia]. *Vestnik MGU*, 8. Istorija, 3, 8–35.
- Kyzlasov, L. R. 1999. Manixejskij xram v kotlovine Sorga (respublika Xakassija) [A Manichaean temple in the Sorga depression (Republic of Khakassia)]. *IRossijskaja arxeologija* 2, 181–206.
- Kyzlasov, L. R. 2001. Sibirskoe manixejstvo [Siberian Manichaeism]. *Ėtnografičeskoe obozrenie* 5, 83–90.
- Nadeljaev, V. M. 1984. Drevnetjurkskie nadpisi Gornogo Altaja [Old Turkic inscriptions of the Mountainous Altai]. In: *Altajskij jazyk na sovremennom étape ego razvitija*. Gorno-Altaysk. 81–103.
- Nevskaya, I. 2011. Some paleographic and orthographic features of Altay Runic inscriptions. In: Çelik Şavk, Ü. (ed.) *Orkhon yazıtların bulunuşundan 120 yıl sonra. Proceedings of the Third Runic Symposium 2*. Ankara: Hacettepe Üniversitesi. 589–599.
- Nevskaya, I. A. 2015. Old Turkic runiform inscriptions of religious content in the Republic Altai: Myths and linguistic evidence. In: Ragagnin, E. & Wilkens, J. (eds.) *Kutadgu nom bitig. Festschrift für Peter Laut zum 60. Geburtstag*. Wiesbaden: Harrassowitz. 313–326.
- Nevskaya, I. & Erdal, M. 2015. Introduction In: Nevskaya, I. & Erdal, M. (eds) *Interpreting runiform sources and the position of the Altai corpus*. (Studien zur Sprache, Geschichte und Kultur der Türkvolker 21.). Berlin: Klaus Schwarz. 8–17.
- Nevskaya, I. & Tybkova, L. 2011. Old Turkic runiform inscriptions in the vicinity of the village Tuekta, Altai Republic, Russian Federation. In: *Ötüken'den İstanbul'a Türkçenin 1290 yılı (720–2010) / From Ötüken to Istanbul, 1290 years of Turkish (720–2010)*. İstanbul. 309–324.
- OTWF = Erdal, M. 1991. *Old Turkic word formation: A functional approach to the lexicon*. Wiesbaden: Harrassowitz.
- Plets G. & Gheyle W. & Verhoeven G. & De Reu J. & Bourgeois J. & Verhegge J. & Stichelbaut B. 2012. Three-dimensional recording of archaeological remains in the Altai Mountains. *Antiquity* 86, 884–897.
- Plisson, H. & Zotkina, L. V. 2015. From 2D to 3D at macro and microscopic scale in rock art studies. *Digital Applications in Archaeology and Cultural Heritage* 2–3, 102–119.
- Pushkarev, A. A. & Zaytceva, O. V. & Vavulin, M. V. & Skorobogatova, A. Y. 2016. 3D recording of a 19th-century Ob River Ship. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Prague. 41-B5, 377–381.
- Ščerbak, A. M. 2001. *Tjurkskaja runika, Proisxoždenie drevnejšej pis'mennosti tjurok, granicy ego rasprostraneniya i osobennosti ispol'zovaniya*. [Turkic runic writing. The origin



of the most ancient Turkic writing, the borders of its spreading and peculiarities of its usage]. Saint Petersburg: Nauka.

Sejdamatov, K. 1964. Drevnetjurkskie nadpisi v Gornom Altae [Old Turkic inscriptions in the Mountainous Altai]. *Materialy po obščej tjurkologii i dunganovedeniju*. Frunze. 95–101.

Tenišev, È. R. 2006. Drevnetjurkskaja èpigrafiika Altaja [Old Turkic epigraphy of Altai]. In: *Izbrannye trudy. Kniga pervaja*. Ufa: Gilem. 140–143.

Tybykova, L. N. & Nevskaya, I. A. 2011. Novaja drevnetjurkskaja nadpis' okolo sela Tuekta [A new Old Turkic inscription near the Tuekta village]. *Rossijskaja Tjurkologija* 1: 4, 3–6.

Tybykova, L. N. & Nevskaya, I. A., & Erdal, M. 2012. *Katalog drevnetjurkskix runičeskix pjamjatnikov Respubliki Gornyj Altaj*. [A catalogue of Old Turkic runic written monuments of the Altai Republic]. Gorno-Altaysk: Gorno-Altayskoe knižnoe izdatel'stvo.

Tybykova, A. T. & Tybykova, L. N. 2015. A history of documentation and research on Old Turkic runiform inscriptions of the Altai Mountains. In: Nevskaya, I. & Erdal, M. (eds.) *Interpreting runiform sources and the Altai corpus*. Berlin: Klaus Schwarz Verlag.

UW = Röhrborn, K. 1977–1998. *Uigurisches Wörterbuch. Sprachmaterial der vorislamischen türkischen Texte aus Zentralasien*. Lieferungen 1–6. Wiesbaden: Steiner.

Vasil'ev, D. M. 2013. *Korpus tjurkskix runičeskix nadpisej Južnoj Sibiri* 1. *Drevnetjurkskaja èpigrafiika Altaja* [A corpus of Turkic runic inscriptions of South Siberia 1. Old Turkic epigraphy of Altai]. Astana: International Turkic Academy.

Vavulin, M. V. 2017. Documentation of Old Turkic runiform inscriptions of the Altai Mountains using photogrammetric technology. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 62-2-W8, 257–261.

Vavulin, M. V. & Zaytseva, O.V. & Pushkarev, A. A. 2014. 3D scanning techniques and practices used for different types of archaeological artifacts. *Siberian Historical Research* 4, 21–37.

Vavulin, M. V. & Zaytseva, O. V. & Vodyasov, E. V. & Tybykova, L. N. & Nevskaya, I. A. 2018. Dokumentirovanie drevnetjurkskix runičeskix naskal'nyx nadpisej Gornogo Altaja na osnove tehnologii fotogrammetrii [Documentation of Old Turkic runic inscriptions on rocks in the Mountainous Altai basing on the photogrammetry technology]. In: *Virtual'naja arxeologija (s vozduxa, na zemle, pod vodoj i v muzeje)*. *Materialy Meždunarodnogo foruma, sostojavšegosja v Gosudarstvennom Ėrmitaže 28–30 maja 2018 goda*. Saint Petersburg: Gosudarstvennyj Ėrmitaž. 29–37.

Zaytseva = Zaitceva, O. V. & Vavulin, M. V. & Pushkarev, A. A. & Vodyasov, E. V. 2016. Photogrammetry: From field recording to museum presentation (Timiryazev burial site, Western Siberia). *Mediterranean Archaeology and Archaeometry* 16: 5, 97–103.

Electronic resources

VATEC: [vatec2.fkidg1.uni-frankfurt.de](http://vatec2.fkidg1.uni-frankfurt.de)

# A Turcological gem: The Tuhan language of Northern Mongolia

**Elisabetta Ragagnin**

Ragagnin, Elisabetta 2018. A Turcological gem: The Tuhan language of Northern Mongolia. *Turkic Languages* 22, 217–229.

This paper presents some new linguistic data on Tuhan, a moribund Sayan Turkic language presently spoken by a handful of old speakers and semi-speakers in the county of Tsagaan Üür in northern Mongolia's East Khövsgöl region. Linguistically, besides displaying typical Sayan Turkic characteristics, Tuhan preserves archaic features otherwise not documented in the rest of Sayan Turkic. At the same time, it shares some isoglosses with Tofan and others with Altay and Yenisey Turkic. Despite the very small number of its speakers, Tuhan appears to be a highly important language for comparative Turcology: a precious linguistic gem among South Siberian Turkic languages.

Keywords: Siberian Turkic languages, Sayan Turkic languages, copula particles, demonstratives, language contact, Turco-Mongolic relations

*Elisabetta Ragagnin, Institute for Turcology, Free University Berlin & Department of Asian and North African Studies, Ca' Foscari University of Venice.*

*Email: Elisabetta.Ragagnin@fu-berlin.de, ragagnin@unive.it*

## Introduction

This paper<sup>1</sup> presents new data from my field expeditions to document speakers of Tuhan, a moribund Sayan Turkic language of northern Mongolia.<sup>2</sup>

The county of Tsagaan Üür, located east of Khövsgöl Lake at a distance of 185 km from the *aimag* city Mörön, is home to a small Turkic-speaking group, known in Mongolian sources as *Uygar-Uriankhay*. Their endonym, however, is *Tuha* [tu<sup>h</sup>ha], a variant of *tuba/tubo*, an ancient ethnic name that is common across Sayan Turkic speakers and various groups in the neighboring areas. The specific form *tuha* is shared by the reindeer herding Dukhan of West Khövsgöl and by Tuvans of the

1 I wish to thank the anonymous reviewers for valuable comments and suggestions.

2 Some linguistic aspects discussed in this article are mentioned in Ragagnin (in press) and were presented by the author on March 3, 2018, at the International Symposium “Current Topics in Turkic Linguistics/Aspects of Agglutination in Turkic Languages: An Integrated Approach of Phonology, Morphosyntax and Semantics: The 3rd meeting” held at the Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies, Japan.

Tere-Khöl and Toju areas in Tuva.<sup>3</sup> In Mongolia, Tuhans are also named *Üüriyn Uriankhay* ‘Uriankhay of the Üür River’, *Ĵinkhene Uriankhay* ‘real/true Uriankhay’ and *Tuha Uriankhay*. Further names found in the literature are *Soyot* or simply *Uriankhay*. The name Uriankhay—an ancient ethnic denomination used to refer to various ethnic groups and also documented in the Secret History of the Mongols (*uriangqai*) and in other sources (Rybatzki 2006: 153–154, Janhunen 2014: 76–77)—is furthermore in use to denote three Mongol-speaking groups of this area, namely *Övör Širkheten Uriankhay*, lit. ‘the Uriankhay of the Southern Flea’, *Arig Uriankhay* (from the name of the adjacent Arig River), and *Görööčün Uriankhay* literally ‘Hunting Uriankhay’; see Ragagnin (2009: 226–227). Officially, the total number of Uriankhay in Tsagaan Üür County amounts to 600. Out of the 2400 inhabitants, 650 are registered as Buryat, whereas the rest are Khalkha. Publications on aspects of the Tuhan language include, alphabetically, Bold (i.a. 2013, 1987, 1982, 1978, 1975), Eriksonas (2013) and Ragagnin, (e.g. 2009, 2013 and in press).

Presently, Tuhan is spoken by a handful of fluent older speakers and some semi-speakers, living in the village of Tsagaan Üür and adjacent areas. The youngest speaker is 67 years old. In my fieldwork sojourn in September 2017, I looked for additional Tuhan speakers in Čandaman Öndör, Bulnay and Khankh, where other Tuhan-speaking groups used to live. However, I could not find any still alive. As for clan names, the Tuhans with whom I have been interacting during my fieldwork belong to the following clans: *Danĵila*, *Derdelei*, *Yamaaday*, *Solgoy* and *Tsagaan Khuular*. The last clan name, Tsagaan Khuular, is also documented for Darkhat Mongols (Badamxatan 1987: 57). Eriksonas (2013: 225), on the basis of personal information provided by Professor V. I. Rassadin, who conducted fieldwork in these areas in the 1980s, mentions five additional Tuhan clan names, namely (in his transcription) *Aryamut*, *Dül’heĵik*, *Ĵohomdoi*, *Qu’rtqaahš* and *Sarihaahš*. Mongush (2003: 167–168) also mentions the subclan Čüs-Danĵila. On the other hand, Dukhan clan names include *Soyan*, *Balikšĵi*, *Urat*, *Čoodu*, *Salčak*, *Dodot*, *Höyüg*, *Demči* and *Darga*; see Ragagnin (2011: 19) for details.

As for lifestyle, Tuhans herd Mongolian-style cattle, especially cows, since their pastures are very lush. My Tuhan informants often proudly refer to their homeland as *Burhan ĵu’rtĵi* ‘God’s homeland’. With regard to the spiritual world, Tuhan beliefs are a combination of Buddhist and animist beliefs. Worth noting in this respect is the special meaning that the county of Tsagaan Üür holds for both Buddhism and animism. The Buddhist monastery dedicated to the well-known shaman deity Dayan

3 Though phonetically the ethnonym is the same, three different terms are in use in scholarly publications to refer to these three Sayan Turkic varieties: “Tuhan” with reference to the East Khövsgöl Steppe Sayan Turkic variety spoken in the county of Tsagaan Üür, “Dukhan” for the West Khövsgöl Taiga Sayan Turkic variety spoken by the reindeer-herding Dukhan community, and “Tere-Khöl Tuvan” for the variety spoken in the Tere-Khöl region in Tuva. For a description of the Dukhan language, see Ragagnin (2011). For the Steppe vs. Taiga subdivision of Sayan Turkic, see Ragagnin (2009: 241).

Deerh lies less than 40 km from the Tsagaan Üür Village. Also connected to the cult of Dayan Deerh is a fertility rite associated with a holy cave located approximately 15 km east of the monastery; for details, see Birtalan (2005 and 2011).

Historically, Tuhāns are descendants of those Sayan Turkic peoples who formed the Lake Khövsgöl *Uriankhay Khošun* in Manchu Qing time; further see Ewing (1981: 186–187), Højer (2003: 171–174), Eriksonas (2013: 228–229), Mongush (2014: 200–201) and Žamcarano (1991). An early 20th-century source providing eye-witness information on Tuhāns, referred to as “Soyotes”, is Haslund-Christensen’s travelogue (1934). According to a widespread story circulating among the Tuhān community, the Tuhān people are supposed to have come from Khan Khökhii in Uvs Aimag in Manchu times; in this respect, also see Njamaa (2011: 17–18). One of my best informants, Gombosüren, who passed away three years ago, added in this regard that Tuhān people came from the Baruun Turuun area in Uvs region. Worth noting in this respect is that in Tuvan folklore the region of Khan Khökhii (Tuvan *xaan kögei*) has specific holy connotations (Žanna Yuša, personal communication). Finally, Natsagdorj (2016) argued that Tuhān speakers historically represent former reindeer herders who changed their lifestyle from reindeer nomadic pastoralism to (Mongolian-style) pastoralism. Regarding this, also compare Ragagnin (in press).

### Distinctive Tuhān linguistic features

As shown in Ragagnin (2009), Tuhān holds a very special position within Sayan Turkic. Besides displaying those typological features that characterize Sayan Turkic as such, and sharing some specific isoglosses with Tofan, Tuhān displays unique lexical and grammatical features that distinguish it from the rest of Sayan Turkic. Such features include (a) absence of word-initial spirantization of initial velar stops in Turkic words, (b) occurrence of the converb {-GAs} instead of common Sayan Turkic {-GAš}, and (c) occurrence of the low-focal intraterminal verbal suffix *-iir* as a variant after polysyllabic vowel-final verbal stems, which is neutral with respect to synharmonism, and shows similarities with Khakas and Yakut. Compare, for instance, Khakas *uzir* ‘(s)he will sleep’, from *uzu-* ‘to sleep’ (Baskakov 1975: 228), and Tuhān *udi:r* ‘id.’ In the present contribution, I would like to add that Tuhān shares the occurrence of the 2PL possessive suffix {- (X)gAr} with Altay Tuvan varieties spoken in Mongolia (Monguš 2009 [1983]: 143) and China (Yuša 2017: 68), and with Altay Turkic (Baskakov 1997: 184). On the other hand, the rest of Sayan Turkic displays {- (X)ŋAr}. Moreover, {- (X)GAr}, in accordance with Sayan Turkic grammatical rules, also marks 2PL of the past {-D-} as well as 2PL imperatives, e.g. Tuhān *šay i<sup>h</sup>šager* [tea drink-IMP.PL/courtesy form] ‘please drink tea’, *geldäger* [come-PAST2.PL] ‘you (pl) came’, Chinese Altay Tuvan (Jungar Tuvan) *siler kördüger* [you see-PAST2.PL] ‘you (pl) saw’ vs. standard Tuvan *siler kördünher* (Yuša 2017: 68), Mongolian Altay Tuvan *malıgar* ‘your (pl) cattle’ vs. Standard Tuvan *malınar* (Monguš 2009 [1983]: 143).

A further noteworthy feature of Tuhan is the occurrence of a four-way distinction in demonstratives: *bo* ‘this’ (proximal), *ol* ‘that’ (distal), *no:* [that (the other one) (distal)] ‘yonder’, and *de:gi* (more distal and (possibly) not visible) ‘that’.

The Tuhan form *de:gi* goes back to the pronoun *dee* augmented with the relational suffix *-ki*. The rest of Sayan Turkic displays cognates void of *-ki*, e.g. Tuvan *döö* ~ *doo* ~ *düü* and *duu* (Isxakov & Pal’mbox 1961: 231–232), Tofan, Soyot and Dukhan *dee* (Ragagnin 2011: 99). Menges (1959/1960: 108) hypothesized that this distal pronoun may originate from the Mongolic distal demonstrative pronoun *\*te-*. Cognates of the distal demonstrative *dee* also occur beyond Sayan Turkic, e.g. Altay *tigi* (Verbitskij 2005[1884]: 354), Khakas *tigĩ* (Baskakov 1975: 148–149), Tatar *tege* (Dewletschin et al. 1989: 235), Chulym Turkic *teg* (Birjukovič 1981: 19). Turkish *öte-ki* ‘the other, the other one’ may also belong here (Marcel Erdal, personal communication).

As far as I know, the distal pronoun *noo*—displaying word-initial *n-*, which is untypical of Turkic—is unique to Tuhan and not hitherto documented in any other South Siberian Turkic language. It is quite likely related to Western Yugur/Yellow Uyghur *naqĩ* ‘yonder’, referring to people and objects at a distance from the speaker (Roos 2000: 83–84), and possibly related to Common Turkic *narukĩ* (< *añarukĩ*) ‘situated beyond, on the other side’ (Roos, personal communication). A relation either with the distal stem *in+* documented in Old Turkic (Erdal 2004: 205–207) or Mongolic *\*nögee* ‘other’ (Nugteren 2011: 463) might also be considered.

Finally, Tuhan holds a special position within Sayan Turkic regarding Mongolic lexical copies. For Mongolic copies that clearly display more archaic features than the rest of Sayan Turkic, pointing to earlier contacts with Mongolic varieties (possibly of Khamnigan type), see Ragagnin (2009), (in press) and (forthcoming).

### Copula particles

Unique to Tuhan is the occurrence of the assertive copula particle *erĩ* continuing East Old Turkic *erür*, which is otherwise not documented in the rest of Sayan Turkic. Two examples are quoted below:

- (1) *Men bi<sup>h</sup>šek ñon-mas erĩ men.*  
 I knife whet- NEG.INTRA COP I  
 ‘I don’t whet knives (indeed).’ (Field notes)

- (2) *Bağši gı<sup>h</sup>şĩ erĩ*  
 teacher person COP  
 ‘He is (indeed) a teacher.’ (Field notes)

On the other hand, the copula particle *duru* ~ *DXrX* is void of assertive contents. One example is quoted below:

- (3) *Hös mal semis duru.*  
 cow cattle fat COP  
 ‘Cows (lit. cow-cattle) are fat.’ (Field Notes)

The copula particle *duru* ~ *DXrX*, as is common in Turkic languages, can also be omitted:

- (4) *ǰuga:la-j-ǰir dil Tuha dil.*  
 speak-COOP-INTRA language Tuhan language  
 ‘The spoken language is the Tuhan language’ (Field notes)

Yet, in the Tuhan material I have analyzed so far, *erī* also occurs in affirmative answers bearing the meaning ‘yes, true, correct’. Close formal and functional correspondences of Tuhan *erī* are found exclusively in Western Yugur/Yellow Uyghur. One example is quoted below:

- (5) *Senəng dioma jer er.*  
 You-GEN opinion right COP  
 ‘Your opinion is (indeed) right’ (Cheng 2009: 251)

For Yellow Uyghur copulative particles, also see Roos (2000: 148–149).

Another formal correspondence of Tuhan *erī* is the Khalaj copula particle *+Ari* /*+r* / *+rī*; see Doerfer (1988: 200–201) and (1989) and Doerfer & Tezcan (1980: 111a). Two Khalaj examples are quoted below:

- (6) *Biz-üm baluq ēydi-si ho’ul-ari.*  
 we-GEN village master-POSS3 good-COP  
 ‘Unser Dorfherr ist gut.’ (Doerfer 1988: 200–201)
- (7) *Bo pādīšā-r.*  
 this padishah-COP  
 ‘Der ist ein Herrscher.’ (Doerfer 1988: 201)

For traces of *er-* in modern Turkic languages, see Doerfer (1989), Johanson (2000) and Brendemoen (2010).

The set of Tuhan particles originating from *\*er-* also includes the conditional and concessive copula particles *erse* and *erip*, shared with Tofan and Soyot, but otherwise not found in the rest of Sayan Turkic. Some examples are quoted below:

- (8) Tuhan  
*Ol bis-təŋ ǰooyaš bol-gan erse men göskīs-er men.*  
 that we-GEN near become-POST COP I show-INTRA I  
 ‘If we were near, I would show it (to you).’ (Ragagnin 2009: 239)

- (9) Tuhan  
*Tuha g<sup>h</sup>šī erip ĵuga:la-vas.*  
 Tuhan person COP speak-NEG.INTRA  
 ‘Although she is Tuhan, she does not speak.’ (Field notes)
- (10) Tofan  
*Īndīy čaraš erip, šin-nīy bol-dīy.*  
 such small COP truth-ADJ.DER become-PAST2SG  
 ‘Although you are so small, you were right.’ (Rassadin 2016: 585)
- (11) Tofan  
*Sen men-den burun gel-dīy erse,*  
 you I-ABL before come-PAST.2SG COP  
*bod-uŋ bile sooda-ar sen.*  
 REFL-POSS3SG with speak-INTRA you  
 ‘If you come before me, then you will tell (it) yourself.’ (Rassadin 2016: 585)
- (12) Soyot  
*Ol kiši kel-gen erse.*  
 that person come-POST COP  
 ‘If that person came.’ (Rassadin 2010: 225a)
- (13) Soyot  
*Aĵiri sooq bol-yan erip, bud-īm doŋ-maan iik.*  
 extremely cold become-POST COP self-POSS1 freeze-NEG.POST PTC  
 ‘Although it was very cold, my legs did not freeze.’ (Rassadin 2010: 43)

Yet, another copula-like particle that occurs in Tuhan and displays a high functional load is *GXy*. It occurs after nominals (nouns, adjectives and verbal nouns), and functionally it appears to focus the psychological interest. On this important function of copula particles, cf. Karakoç (2009). See the Tuhan examples below:<sup>4</sup>

- (14) *Men de:gi g<sup>h</sup>šī Tsevelma: guai bile ĵuga:li:r giy men.*  
 I that person Tsevelmaa title with speak:INTRA COP I  
 ‘I speak with that person, with Mrs Tsevelmaa.’ (Field notes)
- (15) *Monu:n aru:l al-gas hota-ġīdī id-īp sa<sup>h</sup>t-ar ġīy.*  
 this:ABL aruul take-CONV city-DIR send-CONV sell-INTRA COP  
 ‘After the *aruul* (type of cheese) is got from here, it is sent to the city (i.e. Ulaanbaatar) and sold’ (Field notes)

4 A contrastive analysis of Tuhan copula particles will be the subject of a forthcoming publication.

- (16) *Tuha uruγ-lar surgu:l-ga gir-ges aday dīl bil-bes*  
 Tuha child-PL school-DAT enter-CONV Mongolian language know-INTRA  
*Tuha dīl-līy gir-īp tur-gan ġīy.*  
 Tuha language-DER enter-CONV stand-POST COP  
 ‘When Tuhan children started (lit. entered) school, they did not know Mongolian. They started school (only) knowing Tuhan (and not Mongolian).’ (Field notes)
- (17) *Öske bo habī-de ĵuga:li:r g<sup>h</sup>šī ĵok ġīy.*  
 other this surrounding-LOC speak:INTRA person inexistent COP  
 ‘There is nobody else speaking (Tuhan) around here.’ (Field notes)
- (18) *Bo me:s-te ay ün-er.*  
 this forest-LOC lily-bulb exit-INTRA  
*Ay ġas-īp ĵi:r*  
 lily-bulb dig-CONV eat:INTRA  
*gašga ay de-es bir ĵime dur-ar ġīy.*  
 white lily-bulb say-CONV one thing stand-INTRA COP  
*Am bo me:s-te bar.*  
 now this forest-LOC existing  
 ‘In this forest there are lily-bulbs (Lat. *lilium pumilum*). We dig and eat lily-bulbs. There is a thing called white lily-bulb.<sup>5</sup> Now you can find it in this forest.’ (Field notes)
- (19) *Qor<sup>h</sup>-ar ĵime ĵok ġīy.*  
 fear-INTRA thing inexistent COP  
 ‘There is nothing to be afraid of.’ (Field notes)
- (20) *Ĵarim ĵime-sin men bil-bes ġīy men ĵoŋ!*  
 some thing-POSS3.ACC I know-INTRA COP I EMPH.PTC  
 ‘I certainly don’t know some things (mentioned before)!’ (Field notes)
- (21) *Eremsi:r<sup>6</sup> ġīy men.*  
 be thankful:INTRA COP I  
 ‘Thank you.’ (Field notes)

In Tuhan, the participle *-GAn* and *gXy* often, especially in rapid speech tempo, merge into *-GAy*, as seen in the following example:

5 Cf. Mongolian *tsagaan töms* ‘white potatoes’.

6 Tuhan *eremse-* ‘to be thankful’ most likely is diachronically related to East Old Turkic *ögir-* ‘to rejoice’ (Clauson 1972: 113–114) augmented by the simulative suffix *{-msA}*. Tofan, Dukhan and Soyot display the corresponding item *öör-* ‘to rejoice’, whereas Tuvan has *öörü-*. The absence in Tuhan of an initial long vowel, however, is difficult to explain.



- (22) *Halha gi<sup>h</sup>ši bile dur-gas tuha-sın jugaala-vas*  
 Khalkha person with stand-CONV Tuha-POSS3.ACC speak-NEG.INTRA  
*bol-ı ber-gey men [...]*  
 become-CONV give-POST.COP I  
 ‘I married a Khalkha person and I stopped speaking in Tuhan (lit. ‘I became a non-speaker’).’ (Field notes)

On the other hand, in such copular syntactic function the rest of Sayan Turkic displays the auxiliary noun ‘(some)thing’: Tuvan *čüve*, Tofan and Dukhan *jüme* ~ *jime* ‘thing’, inspired by Mongolic *\*yauma*<sup>7</sup> (cf. Written (Literary) Mongol *yaɣuma*, Khalkha *yüüm* and Buryat *yüüme(n)*) both formally and functionally). In Tuhan, *jüme* only occurs as a noun meaning ‘(some)thing’ and never in such copular position (cf. examples 20–22). Tuvan, Tofan and Dukhan examples displaying the auxiliary noun ‘(some)thing’ in copular function are quoted below:

- (23) Tuvan  
*Süür-ool baza seen dugay-ın čugaala-vaan čüve*  
 Süür-ool also you:GEN concern-POSS2 speak-NEG:POST thing  
 ‘Also Süür-ool did not speak about you.’ (Monguš 2009 [1983]: 110)
- (24) Tofan  
*Unuun bääri kiši-ler ol ool-nı ay kiši-si*  
 that:ABL since person-PL that boy-ACC moon person-POSS3  
*de-p ülegäär-lä-är jüme.*  
 say-CONV story-V.DER-INTRA.VBN COP  
 ‘The thing is that, thereafter, people named him moon boy.’ (Rassadin 1996: 10)
- (25) Dukhan  
*A<sup>h</sup>šta-p suksa-p jora:š ol i<sup>h</sup>t-ı börü*  
 be hungry-CONV be thirsty-CONV move:CONV that dog-POSS3 wolf  
*bol-a ber-ıen i<sup>h</sup>öögü-lüj jime.*  
 become-CONV give-POST.VBN history-ADJ.DER COP  
 ‘It was constantly hungry and thirsty, and that’s the story of how the dog began to turn into a wolf.’ (Ragagnin 2011: 229)

The Mongolian parallel syntactic function of *yum* ‘(some)thing’ can be seen in the two examples below:

- (26) Mongolian  
*Tzengxer ool-s-ii.g dab-aad yab-sen youm.*  
 azure mountain-PL-ACC cross-CONV move-POST COP  
 ‘We traveled crossing azure mountains.’ (Janhunen 2012: 229)

7 Reconstructed Mongolic forms are quoted according to Nugteren (2011).

## (27) Mongolian

*En' unen ug youm.*  
 this true word COP

'This is the truth.' (Janhunen 2012: 230)

Finally, following a participle suffix, *yum* is generally cliticized in the spoken language. For instance, the past participle {-*SON*} and *yum* merge into *-siim* (Janhunen 2012: 229). On Mongolian *yum*, also see Mönkh-Amgalan & Kan Šin (2014: 324).

**Further functions of *GXY***

The item *GXY* can also occur as a general noun referring both to animates and to inanimates, as seen in the examples below:

(28) *Üš giy hota-da bar, hi giy mında bar.*

three GN city-LOC existent two GN here existent

'Three "ones" are in the city (i.e. Ulaanbaatar), two "ones" are here.' (Field notes)<sup>8</sup>

(29) *Men hi giy-ni al-gan men.*

I two GN-ACC take-POST I

'I took two "items".' (Field notes)

Furthermore, *GXY* combined with demonstratives (*bo*, *ol*, *de*: and *no*:) forms a special class of extended demonstrative pronouns widely used in Tuhan and not found in any other Sayan Turkic variety, e.g. *bo gıy* ~ *bo gıy* ~ *boğıy* 'this one', *ol gıy* ~ *ol gıy* ~ *olğıy* 'that one', *no: gıy* ~ *no: gıy* ~ *no:gıy* 'yonder one' and *de: giy* ~ *de:giy*, 'that one (more far away)', with reference to both animates and inanimates. Such extended demonstratives are then inflected regularly, e.g. *no: gıylar* ~ *no:gıylar* 'yonder ones', *bo gıyğa* ~ *boğıyğa* 'to this one', *olğıynı* 'that one (ACC)', *boğıy bile* 'with this one' etc.

**Origin of *GXY***

As for the origin of Tuhan *GXY*, in Ragagnin (2009: 236) I proposed to trace it back to Turkic *\*qañu* 'which'. I was obviously wrong. Now it seems to me more plausible to relate Tuhan *GXY* with Western Yugur *kô* / *kə* 'this, that', which is mostly used as a personal pronoun (Roos 2000: 84), and to the Chuvash proximal demonstrative *ku* <ky>. Historically, Tuhan *GXY* could then be related to Bolgar Turkic *ꞑ*, underlying *kö* / *kü* / *gö* / *gü* and occurring as a comment of a nominal sentence, exactly like in the Tuhan sentences presented above. See the discussion on an Altaic

<sup>8</sup> The informant is speaking about her children.

particle *gU* in Erdal (1991). Whether or not the Kazakh particle *goy* <ғой>, occurring in sentence-final position after nominals, including participles, but also after finite verbal forms, belongs here is an open question that needs a detailed analysis.

### Acknowledgements

The data discussed in this article belong to the cultural heritage of the Tuhan people and were collected with the permission of the Tuhan community. I would like to thank the Tuhan community for sharing their knowledge with me and for welcoming me very warmly every time I visit them. I also owe my deepest gratitude to Marti Roos for fruitful feedback on Western Yugur copula particles, Žanna Yuša for inspiring discussions on Sayan Turkic folklore, Benjamin Brosig for feedback on Mongolic “things”, and Altangül Bolatqızı for brainstorming on Kazakh particles.

### Transcription and abbreviations

The transcription used in this contribution follows Lars Johanson’s principles outlined in (Johanson forthcoming). Abbreviations occurring in the grammatical glosses are: ABL: ablative, ADJ: adjectival, ACC: accusative, CONV: converb, COP: copula, COOP: cooperative, DAT: dative, DER: derivation, DIR: directive, EMPH: emphatic, GEN: genitive, GN: general noun, IMP: imperative, IND: indirective, INTRA: intraterminal viewpoint operator, LOC: locative, NEG: negative, PAST: past, PL: plural, POSS: possessive, POST: postterminal viewpoint operator, PTC: particle, REFL: reflexive, SG: singular, V: verbal, VBN: verbal nominal.

### References

- Badamxatan, S. 1987. *Les chamanistes du Bouddha vivant* (Études Mongoles et Sibériennes 17[1986]). [Traduit et adapté par Marie-Dominique Even]. Nanterre: Centre d’Études Mongoles et Sibériennes.
- Baskakov, N. A. (ed.) 1975. *Grammatika xakasskogo jazyka* [Grammar of the Khakas language]. Moscow: Nauka.
- Baskakov, N. A. 1997. Altajskij jazyk [Altay language]. In: Tenišev, È. R. et al. (eds.) *Jazyki mira: Tjurkskie jazyki* [Languages of the world: Turkic languages]. Biškek: Izdatel’skij Dom “Kyrgyzstan”. 179–187.
- Birjukovič, R. M. 1981. *Morfologija čulymsko-tjurskogo jazyka 2*. [Morphology of the Chulym-Turkic language 2.] Saratov: Saratov State University.
- Birtalan, Á. 2005. An invocation to Dayan Deerx collected from a Darkhad Shaman’s descendant. In: Grivelet, S. & Meserve, R. & Birtalan, Á. & Sary, G. (eds.) *The Black Master. Essays on Central Eurasia in Honor of György Kara on his 70th birthday*. Wiesbaden: Harrassowitz. 21–32.
- Birtalan, Á. 2011. The representation of the Mongolian shaman deity Dayan Deerh in invocations and in a Buddhist scroll painting / La représentation de la divinité chamannique Dayan Deerh dans les invocations et dans un rouleau peint bouddhique. In: *Variations tibétaines, Et autres...* (Études mongoles et sibériennes, centrasiatiques et tibétaines 42.) 1–39. (<http://emscat.revues.org/index1800.html>)

- Bold, L. 1975. Uygur-uriankhay khelniy egšig avia [The vowels of the Uigar-Uriankhay language]. *Khel Zokhiol Sudlal* [Language and Literature Studies] 11, 133–145.
- Bold, L. 1978. *Osobennosti ujuguro-urjanxajskogo jazyka* [Distinctive features of the Uygur-Uriankhay language]. [Ph.D. dissertation, Mongolian Academy of Sciences, Ulaanbaatar].
- Bold, L. 1982. Uygur-uriankhay khelniy khöömiişsen egšgiyn asuudal [The problem of pharyngealized vowels in the Uygur-Uriankhay language]. *Khel Zokhiol Sudlal* [Language and Literature Studies] 15, 54–61.
- Bold, L. 1987. Uygur-uriankhay khelniy tiin yalgal [The case system of the Uigar-Uriankhay language]. *Khel Zokhiol Sudlal* [Language and Literature Studies] 20, 58–67.
- Bold, L. 2013. Uygur Uriankhay khelend orson mongol üg [Mongolic loanwords in the Uighur-Uriankhay language]. In: Sükhbaatar, Na. (ed.) *Tuva: Tüükh, khel, soyol* [Tuva: History, language, culture] (Bibliotheca Oiratca 29.) Ulaanbaatar: Soyombo. 157–164.
- Brendemoen, B. 2010. Some remarks on the copula in the Eastern Black Sea dialects. In: Boeschoten, H. & Rentzsch, J. (eds.) *Turcology in Mainz: Turkologie in Mainz* (Turcologica 82.). Wiesbaden: Harrassowitz. 47–54.
- Cheng, Z. 2009. Modal meanings of two copulas in Western Yugur. *Turkic Languages* 13, 246–255.
- Clauson, Sir G. 1972. *An etymological dictionary of pre-thirteenth-century Turkish*. Oxford: Clarendon.
- Dewletschin, T. & Dewletschin, I. & Tezcan, S. (eds.) 1989. *Tatarisch-Deutsches Wörterbuch*. Wiesbaden: Harrassowitz.
- Doerfer, G. 1988. *Grammatik des Chaladsch* (Turcologica 4.) Wiesbaden: Harrassowitz.
- Doerfer, G. 1989. Die Kopula im Chaladsch. In: Heissig, W. & Sagaster, K. (eds.) *Gedanke und Wirkung. Festschrift zum 90. Geburtstag von Nikolaus Poppe*. Wiesbaden. 106–112.
- Doerfer, G. & Tezcan, S. 1980. *Wörterbuch des Chaladsch. Dialekt von Charrab*. (Bibliotheca Orientalis Hungarica 26.). Budapest: Akadémiai Kiadó.
- Erdal, M. 1991. An Altaic particle gU? ~ ! In: Brendemoen, B. (ed.) *Altaica osloensia. Proceedings from the 32nd Meeting of the Permanent International Altaistic Conference, Oslo 1989*. Oslo: Universitetsforlaget. 125–139.
- Erdal, M. 2004. *Old Turkic grammar*. (Handbuch der Orientalistik 8,3.) Leiden & Boston: Brill.
- Eriksonas, T. 2013. On the phonology and morphology of the Taiga-Sayan Tuha language. In: Erdal, M. & Koç, Y. & Cengiz, M. (eds.) *Dilleri ve kültürleri yok olma tehlikesine maruz Türk toplulukları konulu 4. [On endangered Turkic languages and cultures]*. (Uluslararası Türkiyat Araştırmaları Bildirileri). Ankara: Hacettepe Üniversitesi Türkiyat Araştırmaları Enstitüsü Yayınları. 225–246.
- Ewing, T. E. 1981. The forgotten frontier: South Siberia (Tuva) in Chinese and Russian history, 1600–1920. *Central Asiatic Journal* 25/3–4, 174–212.
- Haslund-Christensen, H. 1934. *Tents in Mongolia (Yabonah). Adventures and experiences among the nomads of Central Asia*. New York: Dutton & Co.
- Højer, L. 2003. *Dangerous communications: enmity, suspense and integration in postsocialist northern Mongolia* [Ph.D. dissertation, Department of Social Anthropology, University of Cambridge].
- Isxakov, F. G. & Pal'mbax, A. A. 1961. *Grammatika tuvinskogo jazyka. Fonetika i morfologija* [Grammar of the Tuvan language. Phonetics and morphology]. Moscow: Izdatel'stvo vostočnoj literatury.

- Janhunen, J. 2012. *Mongolian* (London Oriental and African Language Library 19.). Amsterdam & Philadelphia: Benjamins.
- Janhunen, J. 2014. On the ethnonyms Orok and Uryangkhai. *Studia Etymologica Cracoviensia* 19, 71–81.
- Johanson, L. 2000. Traces of a Turkic copula verb. *Turkic Languages* 4, 235–238.
- Johanson, L. (forthcoming) Turkic. Cambridge: Cambridge University Press.
- Karakoç, B. 2009. Notes on subject markers and copular forms in Turkish and in some Turkic varieties of Iran: A comparative study. *Turkic Languages* 13, 208–224.
- Menges, K. H. 1959–1960. Die türkischen Sprachen Süd-Sibiriens 3: Tuba (Sojon und Karayās), 2: Zur Charakteristik einer einzelnen sibirisch-türkischen Gruppe. *Central Asiatic Journal* 5, 97–150.
- Monguš, D. 1983. O jazyke Tuvincev severo-zapadnoj Mongolii [Concerning the language of the Tuvans of north-western Mongolia]. In: *Voprosy Tuvinskoj filologii* [Issues on Tuvan philology]. Kyzyl. 127–145.
- Monguš, D. 2009 [1983]. O služebnyx funkcijax slov *kiži*, *ulus* i *čüve* v tuvinskom jazyke [On the auxiliary functions of the words *kiži*, *ulus* and *čüve* in the Tuvan language]. In: Bičeldej, K. A. (ed.) *Tuvinskij jazyk i pis'mennost'* [Tuvan language and writing]. Kyzyl: Tyvapoligraf. 96–117. [Reprint of Monguš, D. 1983. O služebnyx funkcijax slov *kiži*, *ulus* i *čüve* v tuvinskom jazyke. In: Čeremisina, M. I. (ed.) *Tjurkskie jazyki Sibiri* [Turkic languages of Siberia]. Novosibirsk. 12–35]
- Mongush, M. V. 2003. The Tuvans of Mongolia: Peculiarities of contemporary ethnic development. *Inner Asia* 5, 163–176.
- Mongush, M. V. 2014. Tuvans outside of Tuva: The problem of ethnic self-conservation. In: Lkhagvasuren, I. & Konagaya, Y. (eds.) *Oirat People: Cultural uniformity and diversification*. (Senri Ethnological Studies 86.) 197–213.
- Mönkh-Amgalan, Y. & Kan Šin 2014. *Orčin tsagijn mongol xelnij büteever sudlal* [Morphology of modern Mongolian]. Ulaanbaatar: Soyombo.
- Natsagdorj, B. 2016. Khövsgöl nuurīn uriankhay khošuunī üüsel: tsaatnaas malčīn bolson üyl yavts [The Uriyangqai of Khövsgöl: from reindeer nomadic pastoralist to pastoralist]. In: Sampilondov, C. & Hurcha, A. B. & Hiroki, O. (eds.) *Eurasian nomadic pastoralism: History, culture, environment*. Sendai: Center for Northeast Asian Studies, Tohoku University. 135–168.
- Njamaa, M. 2011. *Züünij magnaj: Xojmor' nutag. Xövsgöl ajmgijn Tsagaan-Üür sum*. [The Eastern vanguard: Honorable homeland. The Tsagaan-Üür county of Khövsgöl region] Ulaanbaatar.
- Nugteren, H. 2011. *Mongolic phonology and the Qinghai-Gansu languages*. Utrecht: LOT, Netherlands Graduate School of Linguistics.
- Ragagnin, E. 2009. A rediscovered lowland Tofan variety in northern Mongolia. *Turkic Languages* 13, 225–245.
- Ragagnin, E. 2011. *Dukhan, a Turkic variety of northern Mongolia: Description and analysis* (Turcologica 76.) Wiesbaden: Harrassowitz.
- Ragagnin, E. 2013. Secluded in the valleys of Kubsugul and forgotten in Mongolian libraries: The Uyghur-Uriankhay language and its speakers. In: Pang, T. & Raschmann, S.-C. & Winkelhane, G. (eds.) *Unknown treasures of the Altaic World in libraries, archives and museums: 53rd annual meeting of the Permanent International Altaistic Conference, Institute of Oriental Manuscripts RAS St. Petersburg, July 25–30, 2010*. Berlin: Klaus Schwarz. 358–363.

- Ragagnin, E. (in print). On Old Uyghur traces. In: Gulácsi, Zs. (ed.) *Language, government, and religion in the world of the Turks: Festschrift for Larry Clark at seventy-five*. (Silk Road Studies 19.) Turnhout: Brepols.
- Ragagnin, E. (forthcoming). The puzzling Tuhān language. To appear in: *Proceedings of the International Symposium "Current Topics in Turkic Linguistics/Aspects of Agglutination in Turkic Languages: An Integrated Approach of Phonology, Morphosyntax and Semantics: The 3rd meeting"* held at the Research Institute for Languages and Cultures of Asia and Africa (ILCAA). Tokyo: Tokyo University of Foreign Studies.
- Rassadin V. I. 1996. *Legendy skazki i pesni sedogo Sayana: tofalarskij fol'klor* [Legends, tales and songs of white Sayan: Tofalarian folklore]. Irkutsk: Kom-t po kulture Irkutskoj obl. administracii.
- Rassadin, V. I. 2010. *Soyotica*. Edited by Béla Kempf. (Studia Uralo-Altaica 48.) Szeged: University of Szeged.
- Rassadin, V. I. 2016. *Tofalarsko-russkij slovar' / To'fa-orus soottari* [Tofalarian-Russian dictionary] Moskow: Izdatel'skij Dom JaSK.
- Roos, M. 2000. The Western Yugur (Yellow Uyghur) language. Grammar, texts, vocabulary. Unpublished Ph.D. dissertation. Leiden University.
- Rybatzki, V. 2006. *Die Personennamen und Titel der mittelmongolischen Dokumente: Eine lexikalische Untersuchung*. (Publications of the Institute for Asian and African Studies 8.) Helsinki: University of Helsinki.
- Yuša, Ž. M. 2017. Fol'klor i obrjad tvincev Kitaja v načale XXI veka: struktura, semantika, pragmatika [Folklore and rite of Tuvans of China at the beginning of the XXI century: structure, semantics, pragmatics] (unpublished habilitation thesis). Novosibirsk: Institut filologii Sibirskogo otdelenija Rossijskoj akademii nauk.
- Žamcarano, C. Ž. 1991. The Darqad and the Uriyangqai of lake Köbsögöl. *East Asian History* 1, 55–80.

# The multi-functional converb *-GAš* and related forms in Sayan Turkic

Gregory D. S. Anderson

Anderson, Gregory D. S. 2018. The multi-functional converb *-GAš* and related forms in Sayan Turkic. *Turkic Languages* 22, 230–285.

The multi-functional converb in *-GAš* is a salient and defining feature of complex sentence and narrative structure of the Sayan Turkic. The range of functions associated with this converb element includes temporal interpretations of complex sentences, including anteriority (and simultaneity), and some situations in individual Sayan Turkic languages that appear to be pushing the *-GAš* form into the finite verbal system as an anterior TAM marker as well. Perhaps the most salient and common feature across the Sayan Turkic dialects is its propulsive function in advancing narrative discourse. In some of the taiga varieties of Sayan Turkic, further specialization is found, namely the grammaticalization of the element as a same-subject marker within a system of switch-reference, and another path in yet other Sayan Turkic varieties is found in a system of tail-to-head linkage. In some of the endangered Sayan Turkic lects, *-GAš* is replacing the *-p* converb in certain auxiliary verb constructions. Other functions include both causal and purposive formations expressed by forms using the *-GAš* converb. The paper concludes with case-marked forms of the *-GAš* converb and a comparison of the Sayan Turkic uses of *-GAš* with cognate forms in other Turkic languages.

Keywords: Sayan Turkic, converbs, same-subject, different-subject, tail-to-head linkage, narrative propulsion, temporal formations, case-marking

Gregory D. S. Anderson, *Living Tongues Institute for Endangered Languages & University of South Africa (UNISA), Department of Linguistics and Modern Languages, Theovan Wijk 9–88, Muckleneuk campus, Pretorius Street, Pretoria, South Africa. E-mail: livingtongues@gmail.com*

## 1. Introduction

The Sayan Turkic languages (Schönig 1997, 1999) are a cluster of closely related Turkic lects found in a region roughly co-terminous with the Altai-Sayan mountain complex in a wide area in the Russian Federation ranging from Nizhneudinskij rajon of Irkutsk oblast' (Tofa) and nearby parts of Buryatia (Soyot) to a dense block covering the entirety of the Republic of Tuva (Central Tuvan, other peripheral Tuvan dialects of Tuva and Todzhu as well as Standard Tuvan), in nearby parts of northern Mongolia (Altai-Tsengel (Monchak) Tuvan, Dukhan and Tuhun) and in the extreme northern part of China (Jungar Tuvan). While sociolinguistically distinct, and exhibiting some considerable internal variation with respect to specific features and their quite diverse and different contact *milieux*, some of which are quite in-

tense—Tofa with Russian, Soyot with Buryat and Russian, Todzhu with Tuvan, Tsengel-Monchak, Dukhan and Tuhan with Mongolian, and Jungar Tuvan with Kazakh, Chinese and Mongolian—these various Sayan Turkic lects nevertheless share a number of features that distinguish them from other subgroups of Turkic varieties. One such feature is the pervasive use of the multi-functional converb in *-GAš* in its various allomorphic realizations, and derivatives thereof. To be sure, a cognate element is found in a number of other Turkic languages and groups, but these show either only a subset of the functions found in the Sayan Turkic varieties or even single functions associated with it, and the multifunctional use of *-GAš* can be considered one of the defining characteristics of the Sayan Turkic group.

In this study I discuss some of the most typical functions of the *-GAš* converb and various morphologically complex converb forms that are derived from this basic element across the different Sayan Turkic varieties. We first start with a basic overview of the functional domains in which this converb element is found in Sayan Turkic (1.1) and offer some background information on it and on converb formations as a whole in Turkic, and then turn to a brief presentation of the morphophonology of the converb in the Sayan Turkic varieties in 1.2. Section 2 examines temporal interpretations of complex sentences with this element, including anteriority and quasi-simultaneity, and section 3 discusses formations that appear to be pushing the *-GAš* form into the finite verbal system as an anterior TAM marker. Section 4 presents perhaps the most salient and common feature across the Sayan Turkic lects: its propulsive function in advancing narrative discourse. Section 5 discusses a further specialization: a functional opposition that has developed in various Sayan Turkic varieties and is suggestive of a grammaticalization of the element as a same-subject marker within a system of switch-reference. Section 6 discusses a related discourse function also found in a subset of Sayan Turkic lects that can be described as a system of tail-to-head linkage. Section 7 introduces examples of the replacement by *-GAš* of the *-p* converb in auxiliary verb constructions in certain endangered Sayan Turkic varieties. Section 8 examines causal functions and section 9 purposive functions of *-GAš*-marked forms. Section 10 examines case-marked forms of *-GAš*. Section 11 details how the functional characteristics of the *-GAš* converb in Sayan Turkic varieties differ from, or have analogs in, other Turkic languages where the element is found. Section 12 summarizes these findings.

### 1.1 On the functions of the converb *-GAš* in Sayan vs. other Turkic languages

Before launching into a discussion of what the functions of the multi-functional converb *-GAš* in Sayan Turkic languages are, some comments on the nature of converbs in Turkic languages as a whole need to be made to situate this element within the typology of Turkic converbs and converb formations. According to Johanson (1995: 313), converb segments (usually = clauses) are formally non-finite units constructionally subordinated syntactically to a base segment (a finite clause) provided with suffixed subjunctors (the converb). In terms of the referential domains of the



converb clause, Johanson distinguishes several levels of their distribution and use, relating to the relationship between the actants in the converb clause and the base clause, and the degree of integration of the two clauses. The converb clause can have a separate and individually expressed subject at “Level 1” of this formal-functional cline (Johanson 1995: 313). Level 2 (Johanson 1995: 314) requires both the converb clause and the base clause to have the same “first actant” (= subject). This in turn feeds fusion and monoclausality, either in lexicalized or conventionalized formations like Turkish *al-<sup>o</sup>p gel-* [take-IP.CV come] ‘bring’<sup>1</sup> where little to no material is permitted between the two elements at Level 3 (Johanson 1995: 315), but where both parts clearly contribute lexical meaning to the resulting form, i.e., within a serialization-like structure. There are also ambiguous forms allowing for lexical and grammatical interpretations at Level 4, as in Kyrgyz (Johanson 1995: 315) *oqu-p tur-du* which can mean either [read-IP.CV stand-DI.PST] ‘he read and then stood up’ or mean [read-IP.CV STAND.AUX-DI.PST] ‘he kept reading’ (although these can most likely be distinguished intonationally). Such converb+base clause forms thus ultimately can become fully conventionalized with the base segment undergoing semantic bleaching and yielding postterminal and intraterminal TAM forms (Johanson 1995: 315) as grammaticalized auxiliary verb formations (Anderson 2004), and also encoding a range of voice/version (Anderson 2001) or orientation meanings.<sup>2</sup> Further, specific converb forms—especially ones based on *de-* ‘say’ but also on *bol-* ‘be’—can be grammaticalized as markers of cause or purpose as well (Johanson 1995).

According to Johanson (1995: 316), converbs bear no mood markers, no ordinary tense markers, and only a restricted set of aspect markers and seldom person/number markers for subjects.<sup>3</sup> All of these are strictly speaking true of the *-GAŠ* converb in Sayan Turkic, which appears to largely be a canonical Level-2 converb. Functionally speaking Johanson (1995: 317) states that “converb markers...are, as a rule, strictly monofunctional”.<sup>4</sup> Functional expansions or paths of development how-

1 Fully lexicalized and unverbated forms are also found, such as Xakas *ayıl-* < \**a(lip) kil-* where the back vocalism of the first root wins out but Tuvan *ekel-* < \**a(lip) kel-*, where it is rather the vocalism of the second root that determines the vocalism of the resulting word, while in both instances, the stem-final *-l* and the converb of the original first word in these lexicalized forms are elided.

2 Some of these grammaticalized constructions (i.e., typically from Level 4) indeed reflect serialized deictic motion formations in their origins (i.e., rather coming from Level 3).

3 But note that subject marking is possible with this in the cognate element in the Kyzyl sub-dialect of the Eastern dialect of Bashkir:

(i) Bashkir, Eastern dialect, Kyzyl sub-dialect (Maksjutova 1976: 279)

<i>al-gas-ım</i>	<i>eště-mě-gěs-ěm</i>
take-GAŠ.CV-1	work-NEG-GAŠ.CV-1
‘I having taken, taking’	‘I having not worked, not working’

4 This is however definitely not true of the *-GAŠ* converb in Sayan Turkic, even with the normal paths of development that Johanson mentions (1971, 1975, 1988, 1990, 1991,

ever are found (Johanson 1995: 321) such as relative anteriority, adverbial “after” clauses coming to encode postterminality or terminality, and causal or purposive semantics. All such developments characterize the *-GAš* converb across the Sayan Turkic varieties. When non-modifying, converb clauses can encode events of equal narrative rank (Johanson 1995: 322),<sup>5</sup> and thus one finds a mismatch between the structural relations, i.e., one of subordination, and the semantic relations, i.e., one of coordination of the predicated events. It is with this latter domain that the present study is largely concerned, as this is the most salient function of the *-GAš* converb in Sayan Turkic varieties.

Some converbs are unanalyzable, historically speaking, in terms of known etymological origins. While lacking in the runic Old Turkic sources, *-yač* (etc.) constitutes one of these so-called “group 1” (historically unanalyzable) type of converbs in Turkic like *-A/I/-y* and *-p* (Johanson 1995: 317).<sup>6</sup> These converbs typically belong to the Level 2 functional distribution mentioned above, i.e., with first actant/subject coreference between the converb clause and the base clause. Johanson’s so-called “group 3” converbs consist of what otherwise appear to be prototypical participles that are marked by case and may permit marking of the subject of the converb clause; unsurprisingly these typically operate at functional Level 1, where first actant co-reference is not required (Johanson 1995: 318). Characteristically a given converb shows strong statistical tendencies towards one or the other preferred contexts, either with or without co-reference of actants. How this type of coreference of actants between converb clauses and base clauses is defined can vary in specific languages, as shown below, even within the cluster under investigation here, Sayan Turkic.

1993, 1995, 1996, 2005), as this element covers many informational domains including ones not enumerated in those studies.

- 5 While Johanson mentions that non-modifying converb clauses (Johanson 1995: 322) can be interpreted as foregrounded or new information, and *-GAš* in Tuvan and related varieties is described as a non-modifying terminal converb (Johanson 1995: 324), albeit one that vacillates between postterminal and terminal semantics, and between circumstantial and narratively equal events (Johanson 1995: 326), the use of *-GAš* in some Dukhan (Ragagnin 2011) texts to encode tail-to-head linkage shows that, at least in certain Sayan Turkic varieties, it does precisely the opposite, i.e., encodes backgrounded or old information, although such forms are indeed non-modifying in Johanson’s sense by being under the temporal, modal and illocutionary force of the base clause (Johanson 1995: 323). Both these and the terminal satellites conveyed for example by the *-ARGA.CV* in Todzhu mentioned below serve this same function, see below.
- 6 A similar converb form in *-yat* is found in Mongolic, which could suggest a copy phenomenon or borrowing scenario, but the directionality of this process is debatable, i.e., from Mongolic into Turkic or from Turkic into Mongolic. Of course, within the Altaic or Transeurasian theory, these could simply be considered to instantiate cognate reflexes of a Proto-Transeurasian (Proto-Altaic) converb form as well.

The functions of the *-GAš* converb in Sayan Turkic languages cover a somewhat wider range of functional domains than discussed by Johanson (1995) but they all form a logical network of extensions of the meanings discussed there. Narrative propulsion and anteriority are logically connected: X happens/happened and then Y. A further related nuance in this domain is the causal construction: because of X, then Y can/could happen. In some sense related to this, but with the inverse relation in the predicated event semantics between the two clauses—that is, one in which the relation of the event predicated by X, or the converb clause, is dictated by that of Y, the base clause—is what underlies the purposive marked clause; i.e., Y occurs, in order for X to occur. Such interpretations were sufficiently common in Sayan Turkic varieties to have lexicalized or grammaticalized certain constructional uses of the verb *de-* ‘say’ in this converb form, viz. *deeš/dääs/tääs/teeš*, etc., particularly with a verb in the aorist participle/intraterminal verbal noun form in *-Vr*, to become specialized as a purposive subordinator across the Sayan Turkic varieties (sometimes as a causal subordinator as well). In Central and Standard Tuvan *deeš* now permits nominal complements as well, and functions as a postposition. In its narrative propulsive function, the *-GAš* converb is typically a Level 2 element in Johanson’s (1995) typology, with a strong preference for first actant/subject continuity co-reference between the converb clause and the basic clause. As such it is not surprising that it has been reinterpreted as a same-subject marker in such Sayan Turkic lects as Todzhu and Tofan. A different but related functional specialization seems to have occurred in Dukhan where strings of such *-GAš*-marked forms chain into sets of sentences in chunks of discourse where each subsequent sentence chains a *-GAš* form of the verb and some attendant arguments—i.e., a non-finite converb clause that recapitulates the preceding sentence’s finite clause—in a pattern known in the typological literature as tail-to-head linkage. A specific further realization, instantiation or extension of this across many of the Sayan Turkic lects is the use of the *-GAš* form of the pronominal verb stem *inĵa-* or various derived forms thereof, e.g., *inĵa-n* or *inĵa-l*, and indeed case-marked forms of the *-GAš* forms of these too, to function as a kind of default narrative event chaining device, and often rendered in translations as ‘(doing) thus, (doing) so’, ‘it being so’, ‘therefore’, ‘thus’, etc. A further but different extension of the propulsive semantics of the *-GAš* converb is temporal anteriority of course, and in turn, this has enabled certain Sayan Turkic lects to permit sentences that have a predicate marked only by the *-GAš* form in an emergent, quasi-finite function, as a kind of anterior TAM marker. In addition to temporal anteriority, some varieties permit temporal simultaneity to be expressed by the *-GAš* form as well, more like *to Y (basic clause) while X-ing (converb clause)*.

Terminologically, the *-GAš* element is known in Tuvan linguistics as the “past converb form” (Isxakov & Pal’mbox 1961: 330). This is also what Rassadin (1978: 178) calls the same form in Tofa(lar), while Rassadin (2010: 25) calls this element the *converbum perfecti* in his analysis of Soyot. Mawkanuli (1999: 194) just calls it a converb and gives no other details about it in Jungar Tuvan. Ragagnin (2009: 234)

also refers to the element *-GAs* in Tuhan simply as a converb,<sup>7</sup> while Ragagnin (2011: 141) defines the *-GAš* element in Dukhan as “a syndetic and non-modifying converb that refers to events of equal narrative value with the event of the head clause”, i.e., something approximating a semantically coordinative but syntactically subordinating element.

## 1.2 On the morphophonology of *-GAš*

As a velar initial element, in Sayan Turkic the converb *-GAš* undergoes some significant morphophonological alternations. First consider the initial consonant: After stems ending in voiceless consonants, archiphonemic *-G* is realized as a voiceless front or back velar/uvular stop in most of the Sayan Turkic varieties, e.g., as *q-* or *k-* *at-qaš* [throw-*GAš*.CV] or *čet-keš* [reach-*GAš*.CV]. After stems ending in nasals or liquids/rhotics, there is a voiced realization of the initial consonant, varying between front and back articulations based on the vocalism of the stem, and alternatively appearing as a voiced fricative (*ɣ*, sometimes phonetically *ʁ*) or a voiced stop (*g*, also sometimes phonetically a voiced uvular stop *ɣ*) *kel-geš* [come-*GAš*.CV], *al-yaš* [take-*GAš*.CV], with the default realizations across the various Sayan Turkic lects being *ɣ/ʁ* with back vocalism stems and *g* with front vocalism in such contexts; i.e., the fricative realizations are more common overall as a whole with back vowel stems, with individual variations attested in specific varieties.

After stems ending in short vowels, the velar is elided and a secondary long vowel produced, usually with unrounded low vowels, thus replacing high vowels in the stem, e.g., *čoru* > *čora-aš* < \**čoru-yaš*, with \*/*uɣa*/ > [*aa*]. Sometimes, as in Tofan, a stem-final voiced velar will be elided but with no coalescence of the vowels, e.g. *čug-* ‘wash’ > *čuaš*. Other times there is degemination, e.g., *deg* > *degeš* with reduction of geminate /*gg*/ > [*g*] (Rassadin 1978: 179). Stems that end in long vowels take the voiced consonant-initial variants, as triple long vowels are not permitted in Sayan Turkic languages, as in Tuvan *xuvaa-yaš*, cf. *xuvaa-yan* ‘divided’ < *xuvaa+GAN* [divide-*GAN*.PST] (Anderson & Harrison 1999: 40). With a small number of short front vowel stems of the shape CV-, the resulting long vowel is not *-ee-* but rather *-ää-* in Tofan (Rassadin 1978: 179), e.g., *či-* ‘eat’ > *čääš* and *dä-* ‘say’ > *dääš*.

## 2. Temporal interpretations of *-GAš* clauses: Anteriority, simultaneity

A range of different temporal interpretations can be found with respect to the relationship of the event encoded in the *-GAš* marked converb clause to that of the base clause. Most typically this is a relationship of anteriority, where the converb clause event precedes the event of the base clause. Such interpretations are found

7 But she notes the unexpected change of final *-š* > *-s* in this form not shared by other Sayan Turkic varieties, nor indeed other words in Tuhan, where one finds the expected final *-š*.

across the Sayan Turkic lects, for example in Central Tuvan (1), Dukhan (2), Monchak or Altai-Tsengel Tuvan (3), Soyot (4), and Tofan (5)–(6).

- (1) Central Tuvan of Tuva (Voinov 2014: 153)  
*Bagay kadarčī meni öřše-ep kör-üñer... at+siv-ijar-ni*  
 lowly shepherd I:ACC show.mercy-IP.CV AUX.ATT/POL-2PL name-POSS2PL-ACC  
*dijna-aš kayga-ar-īm at-tīg bol-du ijar.*  
 hear-GAŠ.CV amazed-INTRATERM.VN-POSS1SG name-ADJ be-DI.PST POL  
 ‘Please (try to) forgive me, a lowly shepherd...(when) I heard your name, it was a name that amazed me.’
- (2) Dukhan (Ragagnin 2011: 140)  
*Po uləs-tar jora-aš gotə pa<sup>h</sup>t-a pa-ar pis.*  
 this people-PL move-GAŠ.CV down sink-A.CV go-AOR 1PL  
 ‘After these people have gone, we move downwards.’
- (3) Altai-Tsengel Tuvan (Aydemir 2009: 57)  
*Jariš do:z-ul-yaš jariš-ya gir-gen bar amitan-nar*  
 bet end-PSV-GAŠ.CV bet-DAT enter-POSTTERM.VN all living-PL  
*jan-ip olur-arda ...*  
 return-IP.CV SIT.AUX-ARDA.CV  
 ‘After the bet came to an end all the living beings who participated in the bet were returning home...’
- (4) Soyot (Rassadin 2010: 54)  
*Jala-aš kel-gän.*  
 invite-GAŠ.CV come-GAN.PST  
 ‘They invited him, and he came.’
- (5) Tofan (Field Notes)  
*...korña-a-n tut-kaš am ĩnda čerle-j ber-gen ihi-jään.*  
 wife-POSS3-ACC take-GAŠ.CV now there live-A.CV AUX.INCH-GAN.PST two-COLL  
 ‘...he took [her] as his wife and started to live there together with her.’
- (6) Tofan (Field Notes)  
*Čil bol-yaš čaa bol-u ver-gen.*  
 year be-GAŠ.CV war be-A.CV AUX.INCH-GAN.PST  
 ‘A year passed and the war started.’

In some instances the relationship between the converb clause and the base clause is one of immediate anteriority, such that there is very little elapsed time between the event of the converb clause and that of the base clause. This interpretation can also be found in a range of Sayan Turkic lects, such as Monchak/Altai-Tsengel Tuvan (7) and Soyot (8).

- (7) Altai-Tsengel Tuvan (Taube 2008: 119)  
*Taraq jaž-ï möndüre-eš xar-ya jüge düž-er-il?*  
 eye tear-POSS3 pour-GAŠ.CV snow-DAT why fall-AOR-3PRON.COP  
 ‘Why do tears pour out and fall onto the snow?’
- (8) Soyot (Rassadin 2010: 54)  
*Čihälä-p<sup>8</sup> kel-dır-gän bod-ın-dan aray burın*  
 go.directly.before-IP.CV arrive-CAUS-GAN.PST self-POSS3-ABL just front  
*eerän šäs-ı-ŋa eşir kel-ıp qon-yaš*  
 shaman’s.amulet directly.in.the.middle-POSS3-DAT eagle come-IP.CV land-GAŠ.CV  
*bod-ın-gıdı gö-ör bol-yan.*  
 self-POSS3-ALL see-AOR AUX.INCH-GAN.PST  
 ‘Directly before her arrival, an eagle flew in and landed on the middle of the amulet and began looking at itself.’
- In yet other instances, the two actions are in a relationship of simultaneity. Such forms are found in such Sayan Turkic lects as Soyot (9)–(10) or Standard Tuvan (11)–(12).
- (9) Soyot (Rassadin 2010: 53)  
*Ašta-aš jımä-sın-den ji-p tur-ar bol-yan.*  
 hunger-GAŠ.CV thing-POSS3-ABL eat-IP.CV STAND.AUX-AOR BE.AUX-GAN.PST  
 ‘Being hungry, he ate from that thing.’
- (10) Soyot (Rassadin 2010: 54)  
*Ät čet-käš čora-an.*  
 horse lead-GAŠ.CV move-GAN.PST  
 ‘They set out, leading the horse on.’
- (11) Standard Tuvan of Tuva (Aydemir 2009: 55; Isxakov & Pal’mbax 1961: 332)  
*Čarāz-ı-n čiyla-aš it tot-pas.*  
 saliva-POSS3-ACC lick-GAŠ.CV dog satiated-NEG.AOR  
 ‘A dog will not get full licking its spit.’
- (12) Standard Tuvan of Tuva (Aydemir 2009: 55; Isxakov & Pal’mbax 1961: 332)  
*Damıraq-tar čiıl-yaš xem bol-ur.*  
 source.stream-PL gather-GAŠ.CV river become-AOR  
 ‘When streams gather together, they become a river.’

8 The so-called pharyngealized (phonetically low pitch) vowels in Sayan Turkic are represented here as in Anderson & Harrison (1999) by *i*, *ù*, *à*, etc.

This simultaneity is sometimes best rendered in translation by an adverbially subordinate temporal clause headed by “while” in English translations. This function of *-GAš* can be found in Central Tuvan (13).

- (13) Central Tuvan of Tuva (Anderson & Harrison 1999: 58)  
*Kizil čor-up or-gaš oruk-ka xoy-nu kör-dü-m.*  
 Kyzyl go-IP.CV AUX-GAŠ.CV road-DAT sheep-ACC see-DI.PST-1  
 ‘While going to Kyzyl, I saw a sheep in the road.’

### 3. Finite functions of *-GAš*?

The notion of anteriority may be at the heart of what appears to be finite uses of *-GAš* converb forms in various Sayan Turkic lects, where no other verb appears. Thus it seems that in a small number of instances, there appears to be an incipient grammaticalization of this element, in such forms, into an anterior TAM marker. Such quasi-finite uses of *-GAš* forms can be found in Monchak or Altai-Tsengel Tuvan (14), or in Standard Tuvan (15).

- (14) Altai-Tsengel Tuvan (Taube 2008: 198)  
*Seziktig serviŋne-eš ĵärlïy sïrvaŋna-aš.*  
 suspicious look.here.and.there-GAŠ.CV mangy wag.tail-GAŠ.CV  
 ‘The suspicious one looks here and there, the mangy one wags its tail.’
- (15) Standard Tuvan (Delger-Ool 1960: 122)  
*Biče-m-de sayazanakta-p čora-an*  
 childhood-POSS1SG-LOC play-IP.CV AUX.DUR-POSTTERM.VN  
*čer-ler-im-ni kör-geš.*  
 place-PL-POSS1SG-ACC see-<GAŠ.CV=ANT>  
 ‘I saw places where I used to play in my childhood.’

It is possible that there is at least one instance of this in Dukhan, based on the following example (16):

- (16) Dukhan (Ragagnin 2011: 195–196)  
*Höngen gïl-ärda e<sup>h</sup>xe-le-eš-təŋ suy i<sup>h</sup>sə-t-keš*  
 höngen.bread make-ARDA.CV begin-V.DER-GAŠ.CV-GEN water hot-CAUS-GAŠ.CV  
*tus gut-kaš.*  
 salt put(.in)-GAŠ.CV  
 ‘To make *höngen* bread, first of all one heats up water and puts salt in.’

### 4. Narrative propulsion

Far and away the most common function of the *-GAš* converb across the various Sayan Turkic lects is as a propulsive narrative device used to advance the events of a plot line or a conversation. In such a function, it is found statistically overwhelm-

ingly in same-subject contexts, thus it canonically instantiates a Level-2 converb in Johanson's (1995) typology. These forms therefore show a mismatch in the morpho-syntax and semantic relations between the converb clause and the base clause: semantically they are coordinate, but syntactically subordinate.

This narrative propulsion use of the converb *-GAš* occurs with a wide range of inflectional forms of verbs in the base clauses, including imperatives (17), for example, in Central Tuvan.

- (17) Central Tuvan of Tuva (Voinov 2014: 138)

*Ol nom-nu a-p al-gaš olur=am.*  
that book-ACC take-IP.CV SBEN-GAŠ.CV sit=POL  
'Please take that book and sit down!'

It very frequently occurs in past forms, as in Altai-Tsengel Tuvan (18), Soyot (19), Dukhan (20) or Jungar Tuvan (21):

- (18) Altai-Tsengel Tuvan (Taube 2008: 150)

*Bay-niŋ baylı-ı-n gör-geš Balyınaq ıziq-ba-an.*  
rich-GEN wealth-POSS3-ACC see-GAŠ.CV Balginaq fall-NEG-GAN.PST  
'Balginaq saw the wealth of the rich and did not fall down.'

- (19) Soyot (Rassadin 2010: 54)

*Šej dül-gäš jime tavaqta-aš sal-ıp*  
tea put.to.boil-GAŠ.CV thing pour.into.bowl.as.meal-GAŠ.CV put-IP.CV  
*ber-gän.*  
BEN-GAN.PST  
'They set tea to boil, laid out something to eat.'

- (20) Dukhan (Ragagnin 2011: 107)

*On jetə gar je<sup>h</sup>t-keš gesə-l-ə ün-gen.*  
17 snow reach-GAŠ.CV wander-MED-A.CV exit-GAN.PST  
'I reached 17 and started wandering around.'

- (21) Jungar Tuvan (Mawkanuli 2005: 115: 1)

*Onu al-gaš žor-up olur-up xanas-ka bar-gaš bis-tiŋ*  
3.ACC take-GAŠ.CV move-IP.CV SIT.AUX-IP.CV Kanas-DAT go-GAŠ.CV we-GEN  
*xuda-bis bol-ur giži-niŋ urug-lar-ı-niŋ žada-a-ŋga*  
in-law-POSS1PL be-INTRATERM.VN person-GEN girl-PL-POSS3-GEN dorm-POSS3-DAT  
*bar-dı-m.*  
go-DI.PST-1SG  
'After I took it (the letter) I set off and went towards Kanas and I went to the dormitory of the girls whose father is one of our in-laws.'



So too does it occur with non-past marked forms with present tense interpretations, e.g., in Altai-Tsengel Tuvan (22), Dukhan (23) or Soyot (24).

- (22) Altai-Tsengel Tuvan (Aydemir 2009: 57)

*Γagy-īm gel-geš šīda-vayn duru men.*  
 laughter-POSS1SG come-GAŠ.CV CAP-NEG.CV STAND.AUX 1SG  
 ‘When I start laughing I can’t stop.’

- (23) Dukhan (Ragagnin 2011: 118, 74)

*Am po tayga-nəŋ urəŋ-lar-ə pol-sa on üš nasən-dan*  
 now this taiga-GEN child-PL-POSS3 become-COND 13 age-ABL  
*gar-lī-īn-dan e<sup>h</sup>xe-le-eš aŋ-nə öören-əp*  
 snow-ADJ.DER-POSS3-ABL begin-V.DER-GAŠ.CV hunt-ACC learn-IP.CV  
*e<sup>h</sup>xe-le-er.*  
 begin-V.DER-AOR  
 ‘Well, as for the children of this taiga, they start to learn hunting from the age, from the age of thirteen.’

- (24) Soyot (Rassadin 2010: 55, 117)

*Ol-güdi gīr-gäš olīr-īp turī*  
 that-ALL enter-GAŠ.CV sit-IP.CV STAND.AUX[:AOR]  
 ‘He goes into it and sits down.’

This also includes so-called narrative present forms, which have non-past marking but logically speaking must be interpreted as past events, as in the following examples from a Dukhan text (25)–(26).

- (25) Dukhan (Ragagnin 2011: 113, 6)

*Pil-əp ga-aβət-kaš iŋgay jō-y pa-ar.*  
 know-IP.CV throw-PRF-GAŠ.CV further move-A.CV go-AOR  
 ‘He notices it and moves away.’

- (26) Dukhan (Ragagnin 2011: 113, 50)

*Hāāg-ə-n sal-aβət-kaš jīnger-te-n-əp pa<sup>h</sup>t-a pa-ar.*  
 willow-POSS3-ACC leave-PRF-GAŠ.CV tumble-V.DER-MED-GAŠ.CV sink-A.CV go-AOR  
 ‘He lets go of his willow-twigs and goes tumbling down.’

The *-GAš* form of the converb also occurs in formations of this type, quite typically with non-past forms with future interpretations in Altai-Tsengel Tuvan (27), Todzhu (28), Dukhan (29) or Tuvan (30).

- (27) Altai-Tsengel Tuvan (Taube 2008: 126)  
*Ĵávaazaq-ti azıra-aš ĵal-ĩ-n ĵüge gez-er=il?*  
 two.year.old-ACC strip-GAš.CV mane-POSS3-ACC why cut-AOR=3PRON.COP  
 ‘After one has shorn the two-year-old, why will one cut the mane?’
- (28) Todzhu (Bergelson & Kibrik 1995: 376)  
*Ava-m inek-ti saap-kaš čan-ĩp kel-ir.*  
 mother-POSS1SG cow-ACC milk-GAš.CV return-IP.CV CLOC-AOR  
 ‘My mother will milk the cow and return home.’
- (29) Dukhan (Ragagnin 2011: 113, 46)  
*Irey-əβas üš hon-gaš ñan-əp ge-er.*  
 old.man-POSS1PL three spend.night-GAš.CV return-IP.CV come-AOR  
 ‘Our old man will be back in three days.’
- (30) Tuvan (Ragagnin 2009: 234)  
*Hi hon-gas gel-ir.*  
 two spend.night-GAš.CV come-AOR  
 ‘(S)he will come in two days.’

Long strings of semantically coordinated events can be marked by this conjunctive or narratively propulsive converb in *-GAš* in the Sayan Turkic varieties. One such example can be seen in the following sentence from Jungar Tuvan (31).<sup>9</sup>

- (31) Jungar Tuvan (Geng Shimin 2000: 52, 43)  
*Qayay ö-ön-gä ĵügür-üp gel-geš bičii ool-dar-ĩ-n*  
 sow house-POSS3-DAT run-IP.CV CLOC-GAš.CV small son-PL-POSS3-ACC  
*balyaş-tan dur-γuz-up al-yaš süit iš-tir[i]-geš*  
 mud-ABL stand-CAUS-IP.CV SBEN-GAš.CV milk drink-CAUS-GAš.CV  
*ö-ön-gä giir-geš ool-dar-ĩ-n sawın*  
 house-POSS3-DAT enter:CAUS-GAš.CV son-PL-POSS3-ACC hay  
*döžä-än-gä ĵit-qır-yaš eji-i-n bižix dekpilä-äš bod-u*  
 bed-POSS3-DAT lie-CAUS-GAš.CV door-POSS3-ACC firmly lock-GAš.CV self-POSS3  
*ool-dar-ĩ-nıñ qıdı-ın-ya ĵid-a (w)al-dı.*  
 son-PL-POSS3-GEN side-POSS3-DAT lie-A.CV TAKE.AUX-DI.PST  
 ‘The sow ran to her house, took up her little sons from the mud and nursed them, then entered the house, let the children lie on the dried hay, firmly locked the door, then lay down at the side of her children.’

Such syndetic or conjunctive converbs are perfect for narrative event chaining. In other Turkic languages that do not make use of the *-GAš* converb as do the Sayan

9 Of course, such strings are indeed also reminiscent of same-subject formations in Tofan and Todzhu; see 5.1 below. Jungar Tuvan likely instantiates this pattern here also.

Turkic varieties, such strings are typically marked by forms with the *-p* converb, as in the following examples from Kyrgyz (32) and Middle Chulym (33).

- (32) Kyrgyz (Imart 1981: 600; Johanson 1995: 329)

*Men erteŋ menen tur-up zaryadka žasa-p kiyin-ip*  
 I morning INS stand-IP.CV gymnastics do-IP.CV dress:RFLXV-IP.CV  
*žu:n-up čay išip mektep-ke bar-a-žat-a-m.*  
 wash:RFLXV-IP.CV tea drink-IP.CV school-DAT go-A.CV-LIE.AUX-PRS-1SG  
 'In the morning I get up, exercise, dress and wash myself, eat breakfast (lit. drink tea) and go to school.'

- (33) Middle Chulym (Field Notes)

*Mān pir kanza tarta-p anzondin pičay a-p*  
 I one pipe pull-IP.CV then knife take-IP.CV  
*anī soyu-p eed-i-n käme-zi-m-ge<sup>10</sup> sa-p*  
 3.ACC skin-IP.CV meat-POSS3-ACC boat-POSS3-POSS1SG-DAT put-IP.CV  
*äp-ke čan pa-ya-m.*  
 house-DAT return AUX.TLOC-GAN.PST-1SG  
 'I smoked one pipe (of tobacco), then I took out my knife, skinned it, put its meat in my boat, and returned home.'

## 5. Switch-reference

Given this predilection to conjunctive chaining in a narratively propulsive manner, and that the overwhelming majority of such forms share the subject or first actant between the converb marked clause(s) and the base clause, it is not too difficult to imagine how certain Sayan Turkic lects have conventionalized this as encoding same-subject in a switch-reference system. This appears to be the case in both Tofan and in Todzhu (Bergelson & Kibrik 1987a, 1987b, 1995), two of the taiga lects of Sayan Turkic. Other Sayan Turkic lects show this as well, especially Soyot, Jungar Tuvan and, to a certain extent, Central Tuvan, but not literary Standard Tuvan.

Switch-reference is a formal mechanism used for tracking subject continuity or discontinuity within complex sentences employed in numerous languages around the globe. There are generally speaking two basic categories which may (or may not) both receive formal marking within a given language, namely same-subject marking and different-subject marking; other, finer gradations within same- or different-subject marking or more restricted systems are also found. The former construction marks subject continuity across components of a complex sentence, while the latter conversely marks subject discontinuity. First identified by Jacobsen (1967), switch-

10 As a critically endangered language, Middle Chulym is showing signs of language attrition such as overgeneralization of third singular possessive markers (or possible full re-grammaticalization of these as definite markers) even in forms also taking other personal possessive marking such as this one.

reference can now be said, contra Haiman (1983), not to be especially exotic or unusual. Most noteworthy in the present context is the pioneering work on the switch-reference system in Todzhu in Bergelson & Kibrik (1987a/b, 1995). Other studies on switch-reference generally or in specific languages include non-exhaustively, MacKenzie (2007, 2010), Nichols (2000), Stirling (1993), Rising (1992), Roberts (1988, 1997), Tsujimura (1987), Finer (1985), and Munro (1979). In fact, representative languages with formal switch-reference systems may be found in all corners of the earth, for example in the Trans-New Guinea languages Kewa (34) or Usan (35), the Nilotic language Lango (36), the Muskogean languages Koasati (37) and Chickasaw (38), and in Gtaʔ, an Austroasiatic language belonging to the Munda family spoken in eastern India; see (39). The actual formal means of realizing these functional categories, of course, vary significantly from language to language.

- (34) i. Kewa [Engan; Papua New Guinea]      ii. Kewa (Foley 1986: 184–185)  
*Nipú réke-na áгаа lá-ma.*      *Nipú réke-na áгаа lá-a.*  
 he stand-3.DS talk say-1PL.NR.PST      he stand-3.DS talk say-3.NR.PST  
 ‘He stood up and we talked.’      ‘He<sub>i</sub> stood up and he<sub>j</sub> talked.’
- (35) i. Usan [Crosilles; Papua New Guinea]      ii. Usan (Haiman & Munro 1983: viii)  
*Ye nam su-ab isomei.*      *Ye nam su-ine isorei*  
 I tree cut-SS I.went.down      I tree cut-DS it.went.down  
 ‘I cut the tree and went down.’      ‘I cut the tree and it went down.’
- (36) i. Lango [Nilotic; Uganda]  
*Rwòt ò-pòy-ò ní è-cég-ò dɔlgòlá.*  
 king 3-remember-PRF COMP 3SS-close-PRF door  
 ‘The king<sub>i</sub> remembered that he<sub>j</sub> closed the door.’  
 ii. Lango (Noonan 1992: 199)  
*Rwòt ò-pòy-ò ní ò-cég-ò dɔ̀gòlá.*  
 king 3-remember-PRF COMP 3-close-PRF door  
 ‘The king<sub>i</sub> remembered that s/he<sub>j</sub> closed the door.’
- (37) i. Koasati [Muskogean; USA] (Kimball 1991: 523)  
*Athómma-k yomáhli-n calakkí ho-ká:ha-áhco-k*  
 Indian-SUBJ go.about.PL-DS cherokee DISTR-say-HAB-IV.PST  
 ‘They called the wandering Indians Cherokees.’  
 ii. Koasati (Kimball 1991: 91)  
*Okí-n askáhka-k skalapist-ók anahka-áhci.*  
 water-OBJ exit.PL-SS mosquito-SUBJ become-PROG  
 ‘They come out of the water and turn into mosquitoes.’

- (38) Chickasaw [Muskogean; USA] (Munro 1983: 223)  
 i. *hi-ha-cha talowa.* ii. *hi-ha-na talowa.*  
 dance-SS.CONJ sing dance-DS.CONJ sing  
 'He<sub>i</sub> danced and (he<sub>i</sub>) sang.' 'He<sub>i</sub> danced and he<sub>j</sub> sang.'
- (39) i. Gta? [Munda, Austroasiatic; India] (Mahapatra & Zide, no date)  
*Hɽij oʔɽi=mwa sgwa we-la ɖokra gweʔ=we-ge*  
 later.on how.much=year like go-DS old.man die=AUX-PST  
 'Later on, after like several years passed, the man died.'  
 ii. Gta? (Mahapatra & Zide, no date)  
*ŋku gnag-hwaʔ toʔ-ce ga-ge.*  
 tiger door-rope open-SS enter-PST  
 'The tiger opened the door(-rope) and entered.'

Most languages which have been identified as possessing formal switch-reference systems have basic SOV clausal constituent order; this is not universally the case however.<sup>11</sup> In such switch-reference systems, typically the last verb in a series of verb phrases or clauses bears full inflection, the others bearing some formal indication that the verb has either the same subject or a different subject than the one verb in the sentence that bears full inflection (the fully finite verb).<sup>12</sup> Frequently, same-subject constructions lack the redundant marking of subject; in the instance of 'different-subject' marking, the verb may bear a marker of its own subject, in addition to a morphological 'different-subject' marker.

The languages of the world can be categorized into four broad groups with respect to switch-reference marking (40).<sup>13</sup> Sayan Turkic varieties belong to Type A, where both same-subject and different-subject receive formal indexing.

(40) Formal Switch-reference Systems

	Same-Subject	Different-Subject
A)	+	+
B)	+	–
C)	–	+
D)	–	–

11 For example the Lango example in (36) above, in which the basic clause structure is VSO.

12 In VSO languages, naturally, it is the first verb that bears full finite inflection, and all following verbs that operate within the switch-reference system bear an indication of subject (dis)continuity. That is, in these forms one sees the mirror image or inverse pattern of what is typically found in SOV languages. Note that in Lango, only same-subject receives formal marking; so more properly only subject continuity is formally marked in Lango.

13 Actually more, if one includes languages that use unbound particles as switch-reference markers.

Category (A) Languages with a morphological opposition of same-subject vs. different-subject: Koasati (Muskogean); Gta? (Munda, Austroasiatic); Tofan, Todzhu (Turkic)

Category (B) Languages with only same-subject morphologically marked: Lango (Nilotic)

Category (C) Languages with only different-subject morphologically marked: Chalcatongo Mixtec (Mixtecan); Tairora (Papuan)

Category (D) Languages lacking a formal switch-reference system: Khasi (Austroasiatic), Nkore-Kiga (Niger-Congo)

Note that the morphological elements that are used as same-subject and different-subject markers in a given individual language may have other functions, or overlap with other sub-systems of clause-combining in that language. In particular, elements functioning as different-subject markers in narrative genres not infrequently function as markers of conditional, or causally or temporally subordinate clauses in conversational genres. This is important as it provides internal pressure that is potentially in part responsible for the loss or breakdown of a switch-reference system.

As mentioned above, Todzhu and Tofan belong to the group of languages that formally mark both same-subject (5.1) and different-subject (5.2). Each of these subsystems is presented in turn below.

### 5.1 Same-subject marking

The *-GAš* converb occurs on non-finite verbs in Todzhu that are narratively equal events strung together in contexts of subject or first actant co-reference between the syntactically subordinate converb clause and the finite base clause, as in the following sentence (41).

(41) Todzhu (Bergelson & Kibrik 1995: 376)

*Ava-m inek-ti saap-kaš čan-ïp kel-ir.*  
 mother-POSS1SG cow-ACC milk-GAš.CV return-IP.CV AUX.CLOC-AOR  
 ‘My mother will milk the cow and return home.’

Note that this is a condition on co-reference between the *grammatical* subjects of the two clauses, with one notable exception (see 5.2 below). Thus, because passive subject is a grammatical subject, when active and passive clauses are conjunctively strung together in this manner, the same-subject marker appears, regardless of whether the passive or active sentence is linearly first (42)–(43) as long as the two grammatical subjects are co-referential in the complex sentence.

(42) Todzhu (Bergelson & Kibrik 1995: 382)

*Ool ava-zïn-ga čug-dur-up al-gaš*  
 Boy mother-POSS3-DAT wash-<CAUS=PSV>-IP.CV AUX.SBEN-GAš.CV  
*oyna-p čoru-ur.*  
 play-IP.CV move-AOR  
 ‘After the boy gets washed by his mother, he will go out to play.’

- (43) Todzhu (Bergelson & Kibrik 1995: 382)  
*Kara-Ool<sub>i</sub> akī-zīn-ga<sub>j</sub> užuraž-ī ber-geš ol<sub>i</sub> aṇaa<sub>j</sub>*  
 Kara-Ool brother-POSS3-DAT meet-A.CV AUX.TLOC-GAŠ.CV 3PRON 3PRON:DAT  
*ette-d-ir.*  
 beat-<CAUS=PSV>-AOR  
 ‘When Kara-Ool<sub>i</sub> meets his older brother<sub>j</sub>, he<sub>i</sub> will get beaten by him<sub>j</sub>.’

Tofan shows basically an identical system for encoding subject co-reference in strings of this sort. -*GAš* is used on non-final verbs, with finite marking on only the final verb in the string (44)–(48).

- (44) Tofan (Field Notes)  
*Ol kīs kort-kaš sun-u-gen.*  
 that girl be.scared-GAŠ.CV run-PRF-GAN.PST  
 ‘The girl got scared and started running.’
- (45) Tofan (Field Notes)  
*Oṇ kel-geš mana-v olur*  
 s/he come-GAŠ.CV wait-IP.CV SIT.AUX.PROG  
 ‘He will come and wait.’
- (46) Tofan (Field Notes)  
*Murgula-aš aṇna-an men.*  
 blow.murgu-GAŠ.CV hunt-GAN.PST 1SG  
 ‘I blew the *murgu* (birch bark hunting horn) and hunted.’
- (47) Tofan (Field Notes)  
*Oṇ xūn san epte turadi tur-gaš škool(a)-ya [bar-gaš] išten-dir.*  
 s/he day every early morning get.up-GAŠ.CV school-DAT [go-GAŠ.CV] work-NARR  
 ‘Every day she gets up early, goes to school and works.’
- (48) Tofan (Field Notes)  
*Keṇe arig-da xem xiyi-n-da tur-gaš aza ble*  
 Evening forest-LOC river edge-POSS3-LOC stand-GAŠ.CV devil with  
*kooš ble sootaj-īp tur-ar bol-gan.*  
 evil.spirit with converse-IP.CV STAND.AUX.PROG-AOR BE.AUX-GAN.PST  
 ‘In the evening he would stand by a river edge and converse with devils and evil spirits.’

Long strings of same-subject marked verbs may appear in a single sentence, followed by a single finite verb in Tofan, as in example (49).

## (49) Tofan (Field Notes)

*Ol kiši čüme-si-n bööl-geš boo-si-n al-gaš*  
 that person thing-POSS3-ACC gather-GAŠ.CV gun-POSS3-ACC take-GAŠ.CV  
*čid-i-n čari-siη-ga bag-la kat-kaš čiškīn al-gaš*  
 dog-POSS3-ACC reindeer-POSS3-DAT rope-INS tie-GAŠ.CV food take-GAŠ.CV  
*talīy-gan.*

leave-GAN.PST

‘That person gathered up his things, took his gun ...tied his dog to his reindeer with rope, took food, and left.’

In certain peoples’ speech, the verb bearing the same-subject marker appears in a doubled or reduplicated form, as in (50)–(51). This appears to be used in forms where the action described by the verb is of long duration, thus reflecting an iconic use of reduplication in such formations.

## (50) Tofan (Field Notes)

*Oy-da ol čerlä-äš čerlä-äš ühe-y ver-gen.*  
 spring-LOC that live-GAŠ.CV live-GAŠ.CV fly-A.CV AUX.TLOC-GAN.PST  
 ‘He was living by that spring (for some time) and flew away.’

## (51) Tofan (Field Notes)

*Am oη iyla-aš iyla-aš oη toož ühe-y ver-di*  
 now he cry-GAŠ.CV cry-GAŠ.CV he also fly-A.CV AUX.TLOC-REC.PST  
 ‘Then he cried and cried and also flew away.’

Other Sayan Turkic varieties show similar tendencies for the use of *-GAš* as a same-subject marker, e.g., in Central Tuvan (52)–(53).

## (52) Central Tuvan (Anderson &amp; Harrison 1999: 85)

*Öörenikči čoruk-če kör-geš xarīla-an*  
 student drawing-ALL look-GAŠ.CV answer-GAN.PST  
 ‘The student looked at the drawing and answered.’

## (53) Central Tuvan (Anderson &amp; Harrison 1999: 85)

*Nom-nu nomca-aš ol kiži-niη čurttalga-zī-n*  
 book-ACC read-GAŠ.CV that person-GEN life-POSS3-ACC  
*šuptu-zu-n bil-ip al-dī-m.*  
 all-POSS3-ACC know-IP.CV AUX.SUBJ.VERS-REC.PST-1SG  
 ‘I read the book and found out everything about his life.’

Perhaps unsurprisingly, given the similarity between Soyot and Tofan, the texts in Rassadin (2010) suggest that the critically endangered Sayan Turkic variety Soyot makes/made use of this narrative device as well.



- (54) Soyot (Rassadin 2010: 53)  
*Ašta-aš suysa-aš pišä-ä bılä kəs-ıp kör-gän.*  
 hunger-GAŠ.CV thirsty-GAŠ.CV knife-POSS3 INS cut-IP.CV AUX.ATT-GAN.PST  
 ‘He was hungry and thirsty and tried to cut it with his knife.’
- (55) Soyot (Rassadin 2010: 54)  
*Kır-ıp kel-gäš biyä ayaq-qa šäy-ın iş-käš*  
 enter-IP.CV AUX.CLOC-GAŠ.CV DEM cup-DAT tea-POSS3-ACC drink-GAŠ.CV  
*ışt-ım aarı-y ber-dı dep*  
 stomach-POSS1SG be.ill-A.CV AUX.INCH-DI.PST say-IP.CV  
 ‘Returning, he drank his tea (that had been poured out) into that cup and said “my stomach was hurting”.’
- (56) Soyot (Rassadin 2010: 55)  
*Ol-güdi gir-gäš olır-ıp turı*  
 that-ALL enter-GAŠ.CV sit-IP.CV STAND.AUX.PRS[:AOR]  
 ‘He goes into it and sits down.’
- (57) Soyot (Rassadin 2010: 54)  
*Şey dül-gäš jıme tavaqta-aš sal-ıp*  
 tea put.to.boil-GAŠ.CV thing pour.into.bowl.as.meal-GAŠ.CV put-IP.CV  
*ber-gän.*  
 AUX.BEN-GAN.PST  
 ‘They set tea to boil, laid out something to eat.’
- Jungar Tuvan (58)–(59, the latter repeating 21) also reveals strings of semantically coordinate but syntactically subordinate clauses in narratives using -GAŠ on non-final clauses under conditions of subject co-reference, albeit not as frequently as in Tofan or Todzhu.
- (58) Jungar Tuvan (Geng Shimin 2000: 49)  
*Ooŋ so-on-da uruy dalaš-payın olur-yaš ada-zın-ya*  
 PRON3:GEN after-POSS3-LOC daughter hurry-NEG.CV SIT.AUX-GAŠ.CV father-POSS3-DAT  
*ayt-ıptur.*  
 say-PST.NARR  
 ‘Then the daughter, having no hurry, said to her father.’
- (59) Jungar Tuvan (Mawkanuli 2005: 115: 1)  
*Onu al-gaš žor-up olur-up xanas-ka bar-gaš bis-tiŋ*  
 PRON3.ACC take-GAŠ.CV move-IP.CV SIT.AUX-IP.CV Kanas-DAT go-GAŠ.CV we-GEN  
*xuda-bis bol-ur giži-niŋ urug-lar-ı-niŋ žada-a-ŋga*  
 in-law-POSS.1PL be-INTRATERM.VN person-GEN girl-PL-POSS3-GEN dorm-POSS3-DAT

*bar-dī-m.*

go-DI.PST-1

‘After I took it (the letter) I set off and went towards Kanas, and I went to the dormitory of the girls whose father is one of our in-laws.’

Other Sayan Turkic lects show *-GAš* in a same-subject role as well, stringing together semantically coordinate narratively equal events, but ones marked as syntactically subordinate. Examples of this sort can be found in both Monchak/Altai-Tsengel Tuvan (60) and in Dukhan (61).

- (60) Altai-Tsengel Tuvan (Aydemir 2009: 54)

*Ād-ī-n sood-yaš dōžek yaŋdīr-yaš udu-du.*  
horse-POSS3-ACC tie.up.to.rest-GAš.CV bed prepare-GAš.CV sleep-DI.PST  
‘He tied up his horse to rest and made himself a bed, and then he fell asleep.’

- (61) Dukhan (Ragagnin 2011: 140–141)

*E<sup>h</sup>rten iβā-nā iŋ-kaš o<sup>h</sup>t-kar-əp jorə-y geje*  
morning reindeer-ACC send-GAš.CV grass-V.DER-IP.CV move-A.CV evening  
*akkel-geš pay-la-ar.*  
bring-GAš.CV tie-V.DER-AOR  
‘One sets free the reindeer, and takes them out to pasture, and the one brings them back in the evening, and ties them up again.’

## 5.2 Different-subject marking

As stated above, it is not necessary to have a formal contrast between same-subject vs. different-subject in the typology of switch-reference, but Sayan Turkic languages nevertheless show a significant statistical skewing of a particular formation used in a different-subject context as well to express narratively equal events, that *does* appear to formally contrast as a different-subject marker with the same-subject marking in *-GAš*. In Todzhu, different-subject is marked by a *dative* form of the non-past participle or intraterminal verbal noun, i.e., *-Vr-GA*. In third-person forms, no subject index is found in Todzhu, but with first- and second-person subjects, the possessive inflectional series typically appears between the participle or nominalizer and the case index to encode the person and number features of that clause’s subject.

- (62) Todzhu (Bergelson & Kibrik 1995: 376)

*Ava-m inek-ti saap-t-arga Kara-Kīs čan-īp*  
mother-POSS1SG cow-ACC milk-PRF-ARGA.CV Kara-Kīs return-IP.CV  
*kel-ir.*  
AUX.CLOC-AOR  
‘My mother will milk the cow and Kara-Kīs will return home.’

- (63) Todzhu (Bergelson & Kibrik 1987b: 33–34; Anderson & Harrison 1999: 86)  
*Koža aal-ga ba-ar/im/ga,*  
 next village-DAT go-ARGA.CV/POSS1SG/ (=DS),  
*kiži čok bo-orga*  
 person NEG.COP BE.AUX-ARGA.CV  
*udu-vayn=daa čan-ip-kan men.*  
 sleep-NEG.CV=EMPH return-PRF-GAN.PST 1SG  
 ‘I went to the next village, no one was there, so I returned home without spending the night.’

Like the same-subject form, this construction is triggered when the *grammatical* subjects between the converb clause and the base clause are not co-referential. Thus, if the matrix *subject* is coreferent with the converb clause *object* in Todzhu, different-subject marking is found.

- (64) Todzhu (Bergelson & Kibrik 1995: 379)  
*Ool xana-nĩ dozula-arga ol čaraš apar-gan.*  
 boy wall-ACC paint-ARGA.CV PRON3 beautiful become-GAN.PST  
 ‘The boy painted the wall and it became beautiful.’

The reverse holds true as well, when the matrix clause object is co-referent with the converb clause subject, then different-subject marking is found.

- (65) Todzhu (Bergelson & Kibrik 1995: 379)  
*Xooray čoru-y ba-ar/im/ga ava-m*  
 town go-A.CV AUX.TLOC-ARGA.CV/POSS1SG/ mother-POSS1SG  
*men-i kōr-beyn bar-dĩ.*  
 PRON1SG-ACC see-NEG.CV AUX-DI.PST  
 ‘I went to town and my mother did not see me.’

In Todzhu, so-called ‘dative subjects’ likewise do not count as grammatical subjects for this distribution either, and one therefore finds different-subject marking in such contexts. Thus, even when the actants are indeed *co-referential*, if they are not the *grammatical* subjects in the two clauses concerned, different-subject marking is required (66).

- (66) Todzhu (Bergelson & Kibrik 1995: 381)  
*Kara-Ool udu-y beerge/<\*ber-geš> aŋaa sook bol-ur.*  
 Kara-Ool sleep-A.CV AUX.INCH:ARGA.CV /<AUX-GAŠ.CV> PRON3:DAT cold be-AOR  
 ‘If Kara-Ool falls asleep, he will be cold.’

Like *-GAš* forms in same-subject constructions, some different (grammatical) subject constructions can have additional quasi-conditional interpretations in Todzhu as well (67).

- (67) Todzhu (Bergelson & Kibrik 1995: 381)

*Kara-Ool-ga sook boorga <\*bol-gaš> ol igla-y beer.*  
 Kara-Ool-DAT cold be:ARGA.CV/<be-GAš.CV> 3PRON cry-A.CV AUX.INCH:AOR  
 'If Kara-Ool gets cold, he will (start to) cry.'

In Tofan the formal details are slightly different. Here it is typically the *locative* form of the non-past participle or intraterminal verbal noun that is the preferred form in clear different-subject contexts, i.e., *-Arda*. Note that also unlike Todzhu, third person subject may optionally be encoded through the possessive inflectional series in Tofan; compare (68) and (69).

- (68) Tofan (Field Notes)

*Čil ert-e vā-ār/in/de čaa bol-u ver-gen*  
 year pass-A.CV AUX.ARDA.CV/POSS3/ war be-A.CV AUX.INCH-GAN.PST  
 'A year passed and the war started.'

- (69) Tofan (Field Notes)

*Altı hün ert-e vā-ārde toy bol-gan.*  
 six day pass-A.CV AUX.PFV-ARDA.CV wedding be-GAN.PST  
 'Six days passed, and there was a wedding.'

Dative-marked forms of verbal nouns/participles also are skewed significantly to the different-subject context in Tofan, but in such instances the semantic relation between the two clauses is more of temporal adverbial subordination rather than semantic coordination. In other words, dative forms here appear to be both semantically and syntactically subordinate or dependent, whereas the locative-marked ones are semantically coordinate but syntactically subordinate.<sup>14</sup>

14 As a whole, Siberian Turkic languages show variation between dative- and locative-marked non-finite structures and indeed between semantically coordinate vs. subordinate interpretations of the syntactically subordinate forms in question, at least translationally. In other words, either coordinate-anterior or adverbial temporally subordinate interpretations are possible in many instances of dative- or locative- marked participle or converb clauses, albeit generally used within different-subject contexts, i.e., ones that lack subject coreference between the component clauses. So examples (ii) and (iii) could just as easily be translated into English with a syntactically subordinate clause (marked by the locative form of the past or non-past participle) headed by an adverbial subordinator "when" and lacking the conjunction at the beginning of the second finite clause; i.e., (ii) could just as felicitously be translated as 'when he got a little closer, his face became visible', etc.

(ii) Khakas (Anderson 1998: 78)

*sala čayınna-anda sūray+čüz-ı körin-ıbıs-ken*  
 a.little come.near-GANDA.CV face-3 be.visible-PRF-GAN.PST  
 'He got a little closer, and his face became visible.'

- (70) Tofan (Rassadin 1978: 39)  
*Čas bol-ī bā-ärge dajgalee-y bā-är.*  
 spring be-A.CV AUX.INCH-ARGA.CV nomadize.in.taiga-A.CV AUX.INCH-AOR  
 ‘When the spring starts, the nomadizing will begin anew.’
- (71) Tofan (Rassadin 1978: 39)  
*Ol buruḡuu üd-ün-den kö-ör/übüs/ke*  
 that first hole-POSS3-ABL see-ARGA.CV/1PL/  
*čü=te közü-l-bes bol-gan.*  
 what=EMPH be.visible-NEG.AOR BE.AUX-GAN.PST  
 ‘When we looked in that first hole, nothing was visible.’
- (72) Tofan (Rassadin 1971: 242)  
*Kel-ir-ge dilyi ēētir-adiri, čü-nü dile-p*  
 come-ARGA.CV fox ask-NARR/HAB what-ACC search-IP.CV  
*čoru sler.*  
 AUX.PROG 2PL  
 ‘When they came, the fox asked “What are you searching for?”’

In Central Tuvan, one sometimes finds converbs that are etymologically locative-marked forms of the past participle or (post)terminal verbal noun (*-GAn-dA*) functioning in semantically coordinate structures as well, i.e., as a kind of different-subject marker.

- (iii) Altai-kizhi [S. Altai] (Baskakov 1958: 77)  
*Ene-si kol=lo ber-erde, uul-ī*  
 mother-POSS3 hand EMPH give-ARDA.CV son-POSS3  
*kol-ün d’apas-öp kattap sura-dī ada-m kayda?*  
 hand-3.ACC squeeze-IP.CV again ask-DI.PST father-POSS1SG where  
 ‘The mother offered her hand, and the son squeezed her hand and again asked “Where’s father?”’

In (iv) and (v) we see examples of the unaccomplished form in *-GALAG-* marked by the locative in Shor but dative in Middle Chulym, but with an identical meaning.

- (iv) Shor (Efremov 1984: 69)  
*Orta čol-ga čet-keleg-im-de, nagbur čibre-p šik-tī.*  
 middle road-DATreach-UNACMPL-POSS1SG-LOC rain drop-IP.CV AUX.INCH-DI.PST  
 ‘I hadn’t yet reached the middle of the road, and the rain began to fall.’
- (v) Middle Chulym (Dul’zon 1960: 121)  
*Olar orta jol-ga čet-kelek-ke suy ča-ap*  
 they middle road-DAT reach-UNACMPL-DAT water precipitate-IP.CV  
*pa-yan*  
 AUX.INCH-GAN.PST  
 ‘They hadn’t yet reached the middle of the road and it started to rain.’

- (73) Central Tuvan (Anderson & Harrison 1999: 86)

*Xün ün-gen-de, čer čira-an*  
 sun come.out-GANDA.CV land grow.light-GAN.PST  
 ‘The sun came out and the land grew light.’

The locative of the *-Vr* form in Tofan (*-Vr-da*) however, clearly stands in explicit contrast with the *-GAš* form. Compare the sentences (74) and (75) in this regard. In (74) there is no subject co-reference between the two clauses, and the different-subject construction (in *-Vr-da*) is found, while in (75), subject co-reference is maintained, and the same-subject formation (in *-GAš*) is rather required.

- (74) Tofan (Field Notes)

*Oŋ tur-ar/in/da (\*tur-gaš) bis olık taley-gan.*  
 s/he stand-ARDA.CV/POSS3/ we immediately leave-GAN.PST  
 ‘He stood up and we immediately left.’

- (75) Tofan (Field Notes)

*Tur-gaš (\*tur-ar/(in)/da) olık taliy-gan.*  
 stand-GAŠ.CV immediately leave-GAN.PST  
 ‘He stood up and immediately left.’

In certain sentences, both different-subject and same-subject marked forms may be found in the same sentence in Tofan in (76). The (unstated) subject of the first two clauses is Fox, but Bear is the subject of the third clause in the sentence. Therefore the first verb is marked with the same-subject suffix, as its subject is the same as that of the following clause, while the second clause bears the verbal noun plus case marking indicating the different-subject construction, as the following clause has Bear as its subject. The third clause, being the last clause of the sentence, receives full finite marking.

- (76) Tofan (Rassadin 1971: 241–242)

*Iz-in-še dü`h`üp bar-gaš haya hōe-n-da*  
 track-POSS3-PROL lower-IP.CV AUX.TLOC-GAŠ.CV rock foot-POSS3-LOC  
*kō-örde iresaŋ öl-ü ver-gen čī`tīrī.*  
 see-ARDA.CV (= DS) bear die-A.CV AUX.PFV-GAN.PST lie.PROG.PRS  
 ‘(Fox) went along Bear’s tracks and saw: Bear was lying there dead at the foot of the mountain.’

Examples of same- and different-subject marking in the same sentence can be found in other Sayan Turkic lects as well. For example in (77) from Central Tuvan, the subject of the first two clauses is the same (“you”), so the first verb is marked by *-GAš*, but the subject of the third clause has shifted (to “I”) and this triggers the different-subject marking on the second clause.

- (77) Central Tuvan of Tuva (Voinov 2014: 158)

*Öglen-ip al-gaş aalda-p keer/iŋer/ge*  
 marry-IP.CV AUX.SUBJ.VERS-GAŠ.CV visit-IP.CV AUX.CLOC:ARGA.CV/2PL/  
*be-er men šive.*  
 give-AOR 1SG MITIG  
 ‘When you get married and come for a visit, I will give (it to you).’

In the following sentence the reverse holds: the subject of the first clause is not co-referential with that of the second or third, and as a result, different-subject marking appears on the verb of the first clause, and same-subject marking on that of the second.

- (78) Central Tuvan (Voinov 2014: 154)

*Ava-zuu ün-e be-erge Dolaana ün-ü-n*  
 mother-POSS3 go.out-A.CV AUX.TLOC-ARGA.CV Dolaana voice-POSS3-ACC  
*öskert-ip al-gaş baštaktan-gula-an.*  
 change-IP.CV AUX.SBEN-GAŠ.CV joke-INCH-GAN.PST  
 ‘When her mother went out, Dolaana changed her voice and began joking.’

The following Soyot sentence shows a pattern consistent with the last Central Tuvan example: only the set of same-subject-marked, semantically coordinate sentences following the initial subject co-reference shift between the first and second clauses entails a string of four -GAŠ-marked clauses, as they all have the same-subject as that of the finite verb in the final/base clause.

- (79) Soyot (Rassadin 2010: 54)

*Bijä kiši gör-ip olir-ar-ya jaahay jaahay ñaš čon-yaš*  
 DEM man see-IP.CV sit-ARGA.CV beautiful beautiful tree hew/carve-GAŠ.CV  
*qaasta-aš čime bilä jima-äš bürhan murn-i-ŋa*  
 decorate-GAŠ.CV thing INS decorate-GAŠ.CV Buddha front-POSS3-DAT  
*sal-yaš nomna-an oliri.*  
 put-GAŠ.CV read-GAN.PST sit:AOR  
 ‘This person<sub>i</sub> sees, he<sub>j</sub> has carved a very beautiful tree decorated with his pattern, decorated with things that had been put before Buddha, and he<sub>j</sub> is sitting and reading a book.’

While not common in Jungar Tuvan narratives published to date, one does find a system reminiscent of Tofan in the following sentence, as the subject co-reference shift seems to be indicated by a locative marked form, not a dative one. So the subject of the first two clauses is ‘I’ (the narrator) while that of the third is his interlocutor. This triggers same-subject marking on the first verb but different-subject marking on the second.

- (80) Jungar Tuvan (Mawkanuli 2005: 119–120: 19)

*Ooŋ            gǎdǐ-iŋ-ga            žid-ip            al-gaš            sen žiŋxua-nǐ*  
 PRON3.GEN side-POSS3-DAT lie.down-IP.CV AUX.SBEN-GAŠ.CV you žinghua-ACC  
*bil-er-sen            be    de-er/im/de            žok di-di.*  
 know-AOR-2SG Q say-ARDA.CV/POSS1SG-/ no say-DI.PST  
 ‘I lay down next to her and asked her “do you know Zhinghua?” and she said “No”.’

While in these examples subject coreference typically reflects consideration of a co-referential grammatical subject, there is one set of examples where, while grammatical subject co-reference is not strictly speaking maintained, either same-subject or different-subject marking is permitted. This is when there is a possessive relationship between the subjects of the two clauses. In some fixed idiomatic expressions that take genitive-possessive forms syntactically, but referentially indicate the subject, both options are grammatical, as in the following sentences from Todzhu. Again this is regardless of whether the possessed noun in question occurs in the converb clause (81) or the base clause (82).

- (81) Todzhu (Bergelson & Kibrik 1995: 383)

*Ool-duŋ    xöŋn-ü            baksıra-y            be-erge            /baskıra-aš*  
 boy-GEN mood-POSS3 get.spoiled-A.CV AUX.INCH-ARGA.CV /get.spoiled-GAŠ.CV  
*čan-ıp            kel-di.*  
 return-IP.CV AUX.CLOC-DI.PST  
 ‘The boy became sick and returned home.’

- (82) Todzhu (Bergelson & Kibrik 1995: 382)

*Kara-Ool čemnen-ip            aar-ga/            al-gaš*  
 Kara-Ool eat:RFLXV-IP.CV AUX.SBEN-ARGA.CV/ AUX.SBEN-GAŠ.CV  
*xöŋn-ü            bulgan-ı            ber-gen.*  
 mood-POSS3 break.away-A.CV AUX-GAN.PST  
 ‘when Kara-Ool had eaten he felt sick’

In forms with a culturally salient co-identification of referents, for example, someone and their horse, this same observation holds true, and either same-subject or different-subject marked forms are considered grammatical, as in the following examples from Tuvan.

- (83) Tuvan (Bergelson & Kibrik 1995: 383)

*Ād-īm            aar-ı-y            be-erge/            aara-aš            oon*  
 horse-POSS1SG sick-A.CV AUX.INCH-ARGA.CV/ sick-GAŠ.CV PRON3.ABL  
*aŋay    čoru-p            šida-va-dǐ-m.*  
 further move-IP.CV CAP-NEG-DI.PST-1SG  
 ‘My horse got sick and I could not ride further from there.’



- (84) Tuvan (Aydemir 2009: 54; Šamina 1987: 91)

*Ād-īm                      čit-keš                      čaday      qal-dī-m                      ād-īm*  
 horse-POSS1SG disappear-GAŠ.CV on.foot remain-DEF.PST-1SG horse-POSS1SG  
*öl-geš                      čaday      qal-dī-m.*  
 die-GAŠ.CV on.foot remain-DEF.PST-1SG  
 ‘After my horse disappeared, I was left on foot’ ‘after my horse died, I was left on foot.’

### 6. Tail-to-head linkage

Another way in which *-GAš* appears to have been conventionalized in various Sayan Turkic lects is to set off chunks of discourse embedded within a system known as tail-to-head linkage, uniting strings of narratively sequential sequences in a discourse chunk, i.e., across finite-marked sentences, not within them, as occurs with switch-reference.

Tail-to-head linkage is a characteristic feature of narrative event chaining in a number of languages of the world outside of Europe, in particular in Papua New Guinea (de Vries 2005), but also in South America (Guillaume 2011), for example the Bolivian language Cavineña of the Tacanan family. Tail-to-head linkage is a way of stringing backgrounded events and prefacing new information or foregrounded events (Guillaume 2011: 119), and often involves copying of arguments as well as predicates, though the latter form typically is stripped of most TAM marking (but does allow some aspectual marking as in Sayan Turkic). These formations are typically embedded within switch-reference systems as well. Tail-to-head-linkage is useful in narratives to end one event and draw attention to the next (Guillaume 2011: 112). It consists of a copy of the finite form of the preceding sentence in a non-finite form at the beginning of a following sentence (85–86).

- (85) Cavineña [Tacanan; Bolivia] (Guillaume 2011: 110)

*Tume jara-bute-kware                      ike*  
 then lie-GO\_DOWN-REM.PST 1SG  
 ‘...then I lay down (on my raft)’

- (86) Cavineña (Guillaume 2011: 110)

*Jara-bute-tsu                      betsa-kware.*  
 lie-GO\_DOWN-SS swim-REM.PST  
 ‘Having lain down on my raft, I swam.’

In Stirling’s (1993) conceptualization, this tail-to-head linkage recapitulation can serve to “allow the switch-reference marking to be carried over from one sentence to the next” (Stirling 1993: 220-1), as switch-reference *per se* is to track co-reference across semantically coordinate but syntactically subordinate clauses *within* a complex sentence. Such tail-to-head linkage formations typically serve to advance plot lines in the narrative events (Payne 1992).

Like Sayan Turkic, there are no coordinate clauses *per se* formally speaking, only semantically coordinate but syntactically subordinate ones in Cavineña. The different-subject forms of Cavineña, again like Sayan Turkic, often involve temporal sequencing or simultaneity (Guillaume 2011: 117), and are often translationally best rendered in English by an adverbial subordinate clause headed by ‘when’ or ‘while’ (Guillaume 2011: 121).

Within Sayan Turkic one finds instances of this particular type of narrative sequencing, known as tail-to-head linkage, in at least some of the texts published in both Jungar Tuvan (Geng Shimin 2000, Mawkanuli 2005) and in Dukhan (Ragagnin 2011). As mentioned above, these consist of the finite verb (and some of its arguments) that occurs at the end of its own sentence being recapitulated in the *-GAš* form at the start of the immediately following sentence.

Dukhan makes use of the tail-to-head linkage device in some of the texts in Ragagnin (2011). Take as an example the following three sets of tail-to-head linked narrative strings from the same text in Ragagnin (2011). This sequence is numbered sentences 20-22 in Ragagnin’s original

(87) Dukhan (Ragagnin 2011: 234)

- (20)...*tilgə-jək ool tay gür-ə-n o<sup>h</sup>rta ot ot-a-βəd-ar.*  
fox-DIM son mountain limit-POSS3-ACC middle fire fire-V.DER-PRF-AOR
- (21) *Petək tay gür-ə-n o<sup>h</sup>rta ot ot(t)a-aš ja*  
high mountain limit-POSS3-ACC middle fire fire-V.DER-GAŠ.CV yeah  
*aǰa-m siler po o<sup>h</sup>rta jīt-ar siler de-p ol irey*  
father-POSS1SG you.PL this middle lie-AOR 2PL say-IP.CV that bear  
*hayrəhan aǰa-sə-n jīt-kər-ək wa-aš ot*  
merciful father-POSS3-ACC lie-CAUS-GAŠ.CV throw-GAŠ.CV fire  
*ot-a-βəd-ar jime.*  
fire-V.DER-COMP-INTRATERM.VN thing
- (22) *Tilgə-jək ool uləy ot ot-a-βət-kaš....*  
fox-DIM son big fire fire.V.DER-PRF-GAŠ.CV
- ‘...the little fox makes a fire in the middle of the ridge. After setting a fire in the middle of the high mountain range, he says, “yes my bear, please lie down in this place”, and he had that bear (father of his) lie down and he made a fire, having made a big fire...’

In the above set (87), five words in a row including the finite predicate are copied from sentence 20 into the start of sentence 21, the last mentioned word transformed into the non-finite *-GAš* form that typifies this narrative device. Similarly the predicative head in sentence 21 (“set a fire”) is copied in the *-GAš* form at the start of sentence 22.

In the following set (88), representing sentences 26–27 in Ragagnin’s original, a three-word predicative sequence from the end of sentence 26 is copied in the *-GAš* form at the start of sentence 27.<sup>15</sup>

(88) Dukhan (Ragagnin 2011: 234–235)

- (26) ... *šala iṅgay pol-ər/ən/da oh irey hayrəhan petək*  
 bit further become-ARDA.CV-POSS3 ITJ bear merciful high  
*tayga-dan juy-l-əp pa<sup>h</sup>t-a pa-ar jime.* )  
 taiga-ABL roll-V.DER-IP.CV sink-A.CV AUX.TLOC-INTRATERM.VN thing  
 (27) *Juy-l-əp pa<sup>h</sup>t-a par-yaš sösken-den i<sup>h</sup>sər-əp*  
 roll-V.DER-IP.CV sink-A.CV AUX.TLOC-GAš.CV twig-ABL bite-IP.CV  
*a-ar.*  
 AUX.SBEN-AOR

‘He took himself a little further over there, oh, the bear goes rolling down from the high taiga. He rolls down, and he grabs onto a little branch with his teeth.’

The following set of three sentences shows a similar sequencing (representing sentences 30–32 in the original), with the finite verb of sentence 30, *te-er* ‘says’, immediately followed by the *-GAš* form of the same (*te-eš*) at the head of sentence 31, and the finite verb complex of sentence 31 *jinger-te-n-əp pa<sup>h</sup>t-a pa-ar* in turn is also copied in a (genitive-marked form of the) *-GAš* converb at the start of sentence 32; see example (89).

(89) Dukhan (Ragagnin 2011: 235)

- (30) *Meen aja-m pol-sa ah te-er josə-ləy.*  
 I:GEN father-POSS1SG become-COND ITJ say-INTRATERM.VN rule-N.DER  
*Meen aja-m emes tirə te-eš irej hayrəhan ah te-er*  
 I:GEN father-POSS1SG NEG.COP PTC say-GAš.CV bear merciful ITJ say-AOR  
 (31) *te-eš hāāg-ə-n sal-əβət-kaš jinger-te-n-əp*  
 say-GAš.CV willow-POSS3-ACC leave-PRF-GAš.CV tumble-V.DER-MED-IP.CV  
*pa<sup>h</sup>t-a pa-ar.*  
 sink-A.CV AUX.TLOC-AOR  
 (32) *Jinger-te-n-əp pa<sup>h</sup>t-a par-yaš-təŋ*  
 tumble-V.DER-MED-IP.CV sink-A.CV AUX.TLOC-GAš.CV-GEN  
*am in-aarə öl-ü pe-er.*  
 now that-ADV.DER die-A.CV AUX.PRF-AOR

‘If it was my bear father, he should say “ah”, after that, after the little fox said “he is

<sup>15</sup> Note that as both Sentences 21 and 26 show, sometimes the predicative heads in Sayan Turkic languages are formally participles modifying a (nearly) semantically empty copular noun, in both of these examples *jime* ‘thing’. Nevertheless the semantically full predicative elements (lexical and functional) are copied into the following sentence in the *-GAš* form, so this is a peculiarly Sayan Turkic manifestation of the tail-to-head linkage structure.

surely not my bear, the bear says “ah”. After saying it, he just lets go of his willow twig and goes tumbling down. After he goes tumbling down, well he dies there.’

There are actually more tail-to-head linked strings in this discourse chunk, which is the climax of the narration, but these use case-marked forms of the converb (as the last such instance in (89) above is) and are further exemplified below in section 10.

The following two sentences also show this kind of recapitulation of action or tail-to-head linkage that is characteristic of Jungar Tuvan narrative texts.

- (90) Jungar Tuvan (Geng Shimin 2000: 49, lines 6–7)

*Onu diŋna-aš aštīyaq uyan osqun-up qory-up*  
 PRON3.ACC hear-GAŠ.CV old.man wisdom lose-IP.CV fear-IP.CV  
*ö-ön-gä ge-eptur gel-gäš uru-un-ya xan-niŋ*  
 house-POSS3-DAT come-PST.NARR come-GAŠ.CV daughter-POSS3-DAT khan-GEN  
*duž(a)-an jarl-ï-in tözin ayt-iptir.*  
 deliver-POSTTERM.VN order-POSS3-ACC all say-NARR  
 ‘After hearing this, the old man lost his wisdom and being very afraid returned home.  
 After he returned home, he told all the khan’s delivered orders to his daughter.’

In this instance, the finite verb of the first sentence is followed immediately by a non-finite *-GAš* marked form of the same verb. While stylistically odd in English, it appears to be a commonly used device to string together events in narration in this Sayan Turkic variety.

Note that tail-to-head linkage can also be done with the use of a “generic verb linkage” (de Vries 2005: 376–367) or a “summary-head linkage” (Thompson et al 2007: 274), i.e., “do this/that/so”, a patterning that precisely brings to mind the textual frequency of the use of *-GAš* forms based on the pronominal verb root *inja* in Sayan Turkic varieties. This is a highly salient feature of Sayan Turkic narrative structure. Several different forms are attested across the various Sayan Turkic lects enumerated below.

As Ragagnin notes for Dukhan (2011: 141–142), *-GAš* forms of the pronominal verbs “do.like.this”, “do.like.that” are commonly used as adverbials in all the Sayan Turkic varieties, whether steppe or taiga varieties, usually in meanings like “thus”, “as such”, “so”, “therefore” or “in such a manner”. This comes in a number of various forms across different Sayan Turkic varieties: One such is the *-GAš* form of the basic stem, i.e., *injaas*. This variety is found in Dukhan (91) and Soyot (92)–(93):

- (91) Dukhan (Ragagnin 2011: 235)

*in-ja-aš tilgə-jek ool amər sajhanjarya-y pe-er.*  
 that-V.DER-GAŠ.CV fox-DIM son easy nice be.happy-A.CV AUX.INCH-AOR  
 ‘So the little fox is at ease and feels good.’

- (92) Soyot (Rassadin 2010: 53)  
*ĩñja-aš ol čimšä-y ber-gän.*  
 do.thus-GAŠ.CV he move.about-A.CV AUX.INCH-GAN.PST  
 ‘Then he began moving about.’

- (93) Soyot (Rassadin 2010: 56)  
*ĩñja-aš bïr čèšä ay=la èrt-kän.*  
 do.thus-GAŠ.CV one some month=EMPH pass-GAN.PST  
 ‘So several months passed.’

With the middle-reflexive stem derivative *-n-*, in the form *ĩñjangaš/ĩñjaŋyaš*, etc., this clause linker is also found in both Dukhan (94), and Soyot (95), and with elision of the velar in Tofan as well (but triggering place assimilation of the nasal prior to being elided), appearing as *ĩñja:ŋaš* (96).

- (94) Dukhan (Ragagnin 2011: 142)  
*Ĩn-ja-n-gaš gō<sup>h</sup>h<sup>i</sup>-er pol-yan.*  
 that-V.DER-MED-GAŠ.CV nomadize-INTRATERM.VN become-GAN.PST  
 ‘So the time came to move.’

- (95) Soyot (Rassadin 2010: 54)  
*Ĩñjaŋ-yaš bod-ĩ-nñj em-in-den iš-käš*  
 do.thus:RFLXV-GAŠ.CV self-POSS3-GEN medicine-POSS3-ABL drink-GAŠ.CV  
*ekkïre bol-yaš.*  
 healthy become-GAŠ.CV  
 ‘Then having drunk his own medicine, he became healthy.’

- (96) Tofan (Rassadin 1971: 242)  
*Ĩn-ja:-ŋaš dilyi ɛet-tïr-adirï iresaŋ-dan.*  
 do.thus:RFLXV-GAŠ.CV fox ask-CAUS-NARR bear-ABL  
 ‘Thus he asked Bear.’

Dukhan also uses a form of this same construction together with the emphatic clitic *=la* as in the following example (97).

- (97) Dukhan (Ragagnin 2011: 108)  
*On üš gar-ləy-dan so-on-da ol örye tiij*  
 13 snow-N.DER-ABL after-POSS3-LOC that ground.squirrel squirrel  
*ĩn-dəy pičče jime a<sup>h</sup>t-əp ĩn-ja-n-yaš=la ekkə*  
 that-ADJ.DER small thing shoot-IP.CV that-V.DER-MED-GAŠ.CV =EMPH good  
*a<sup>h</sup>t-ar pol-əp gir-er.*  
 shoot-INTRATERM.VN become-IP.CV ENTER.AUX-AOR  
 ‘After becoming 13 years old, they shoot small things like ground squirrels and sables and gradually start to become good shooters.’

Case-marked forms of this same connective element based on the pronominal verb stems are found in both Dukhan and Jungar Tuvan. These are exemplified in section 10 below.

### 7. *-GAš* replacing the *-p* converb in auxiliary verb constructions

In a few instances in certain critically endangered varieties of Sayan Turkic, like Soyot or Tofan, the *-GAš* forms appear to be replacing what from a standard Turkological perspective one would expect to be a *-p* converb form, namely on the lexical verb in a complex predicate or auxiliary verb construction. These represent a later stage of development than the Level 2 that typically characterizes this converb in Johanson's (1995) typology.

In Soyot, this replacement occurs commonly in a large complex predicate of the shape Verb-*GAš* *tur-ar bol-yan* (98)–(100), which appears synchronically to represent a grammaticalized whole in this language.

(98) Soyot (Rassadin 2010: 53)

*Mal-ivīs emii šibiškilä-äš tur-ar bol-yan.*  
cattle-1PL udder-POSS3 covered.in.boils-GAš.CV STAND.AUX-AOR BE.AUX-GAN.PST  
'The udders of our cattle became covered in boils.'

(99) Soyot (Rassadin 2010: 53)

*Žü-dän žü bol-yaš tur-ar bol-yan?*  
what-ABL what be-GAš.CV STAND.AUX-AOR BE.AUX-GAN.PST  
'Why did this happen?'

(100) Soyot (Rassadin 2010: 53)

*Mal-ivīs kenät turīda čòša-aš tur-ar bol-yan.*  
cattle-POSS1PL suddenly stand:ARDA.CV startle-GAš.CV STAND.AUX-AOR BE.AUX-GAN.PST  
'Our cattle were just standing there and suddenly startled.'

In other instances in Soyot *-GAš* is clearly replacing what would be expected to be, and indeed is otherwise attested in other examples in the language as, the *-p* converb form, e.g., before the auxiliary *turī*; compare (101) and (102).

(101) Soyot (Rassadin 2010: 53)

*Kenät čòš=ip algir-yaš turī.*  
suddenly shake.in.fear-IP.CV cry.out-GAš.CV STAND.AUX:AOR  
'That one suddenly shook from fear and cried out.'

- (102) Soyot (Rassadin 2010: 54)

*Dàhìy čoo-yaš kel-ìp turì.*  
 again approach-GAŠ.CV AUX.CLOC-IP.CV STAND.AUX:AOR  
 ‘Again he approached.’

In Tofan, there are also clear forms where the *-GAš* converb is replacing the *-p* converb in auxiliary verb constructions or complex predicate structures (Anderson 2001, 2004).

- (103) Tofan (Anderson 2001: 250)

*Dilyi oluk bar-ìp brää üšpül tüt-kaš al-yan.*  
 fox right.away go-IP.CV one hazel.grouse catch-GAŠ.CV AUX.SUBJ.VERS-GAN.PST  
 ‘The fox went and right away caught himself a hazel grouse.’

- (104) Tofan (Anderson 2001: 249)

*Sen boojalaš-kaš al-ibit-ti-ŋ.*  
 2SG bet:RCP-GAŠ.CV AUX.SUBJ.VERS-PRF-DI.PST-2  
 ‘You have won yourself our bet.’

As pointed out by an anonymous reviewer, Mongolian shows similar instances of the potentially cognate form in *-AAd* replacing the expected converb in *-J* in auxiliary verb constructions. Also, as pointed about by a second anonymous reviewer, the negative forms of the all the basic converb forms (*-p*, *-A/j*, and *-GAš*) share the same negative form and thus are impossible to distinguish in the negative, thereby providing internal pressure for them to be reinterpreted as each other or confused for each other.

## 8. Causal functions

Causal functions of *-GAš* marked clauses are not the most common or typical of the many functions that this element encodes when relating the events of the converb clause to the base clause, but it is nevertheless found in a number of the Sayan Turkic varieties. In Central Tuvan (105) or in Soyot (106), *-GAš* can appear in this causal function with a range of roots, and thus it does not appear to be lexically restricted to *bol-* ‘to be’, as it is in this function in some of the other Sayan Turkic varieties (see below).

- (105) Central Tuvan (Anderson & Harrison 1999: 58)

*Xem uglat-kaš parom čor-basta-an.*  
 river flood-GAŠ.CV ferry go-CES-GAN.PST  
 ‘The ferry stopped running because the river flooded.’

## (106) Soyot (Rassadin 2010: 53)

*Karañh-in-da jü=tä köz-il-bäs bol-i ber-gäš*  
 darkness-POSS3-LOC what=FOC see-PSV-NEG.AOR be-A.CV GIVE.AUX-GAŠ.CV  
*dolyan-diri tüšt-ip kör-gäš ün-är jer*  
 around turn-IP.CV AUX.ATT-GAŠ.CV exit-INTRATERM.VN place  
*joq bol-yan.*  
 NEG.COP BE.AUX-GAN.PST  
 ‘Because nothing was visible in the dark, he tried to turn around, but there was no way out.’

In the unspecified variety of Tuvan described by Tažibaeva (2012), the causal reading of *-GAš* appears restricted to auxiliated uses of the stem *bol-* ‘be’ when following a past-marked, (post)terminal verbal noun or past participle form of a verb (107)–(108), itself often in the role of an auxiliary in a large complex predicate, and with a negative scope operator in the base clause.

## (107) Tuvan, unspecified (Tažibaeva 2012: 207)

*Baž-ım aar-y ber-gen bol-yaš seminar-že bar-ıp*  
 head-1 hurt-A.CV AUX.PRF-POSTTERM.VN BE.AUX-GAŠ.CV seminar-ALL go-IP.CV  
*šida-va-dī-m.*  
 CAP-NEG-DI.PST-1SG  
 ‘Because my head hurt I could not go to the seminar.’

## (108) Tuvan, unspecified (Tažibaeva 2012: 208)

*Kirgışbičii-zin-den aažī+sañ-ī-n küš+ažıl-ya dadıktır-ıp*  
 Kirgischchildhood-POSS3-ABL character-POSS3-ACC labor-DAT train-IP.CV  
*kel-gen bol-yaš am ooñ murn-un-da dirgel-ıp*  
 AUX.CLOC-GAN.PST BE.AUX-GAŠ.CV now PRON3.GEN before-POSS3-LOC appear-IP.CV  
*kel-gen bergedeeškin-ner-den korg-up sürte-ve-en.*  
 AUX.CLOC-GAN.PST difficulty-PL-ABL fear-IP.CV fear-NEG-GAN.PST  
 ‘Because Kırğış had been training his character against difficulties from early childhood, he was not frightened by the problems he now encountered.’

In Todzhu, a similar pattern is attested (109), involving a *-GAš*-marked form of *bol-* in a causal function.

## (109) Todzhu (Bergelson &amp; Kibrik 1995: 399)

*Men korg-a ber-gen bol-gaš börü-nü ölür-üp*  
 I fear-A.CV AUX.PRF-POSTTERM.VN be-GAŠ.CV wolf-ACC kill-IP.CV  
*šida-va-an men.*  
 CAP-NEG-GAN.PST 1SG  
 ‘I could not kill the wolf because I was scared.’



Note however that in this Sayan Turkic lect, the stem *bol-* can also carry causal semantics in different-subject contexts as well, and thus one finds a formal opposition of causal *bol-gaš* in same-subject contexts contrasting with *bo-orga* in different-subject ones in Todzhu (110).

- (110) Todzhu (Bergelson & Kibrik 1995: 399)  
*Börü anıyak bo-orga ölü-be-di-m.*  
 wolf young be-ARGA.CV kill-NEG-DL.PST-1SG  
 ‘Because the wolf was young I did not kill it.’

Note that some of the other semantic functions of *-GAš* in Tofan that can be used as a same-subject marker have analogs in the different-subject marking system as well, maintaining the particular manifestations of the norms of Tofan grammar. Thus, converbs that are etymologically dative-marked past participles/postterminal verbal nouns (*-GAN-gA*) or non-past participles/intraterminal verbal nouns (*-Vr-gA*) can have causal interpretations as well in different-subject contexts (111)–(112). Note as mentioned above that dative-marked forms in Tofan tend to be semantically as well as syntactically dependent or subordinate, unlike the locative-marked ones that are typically semantically coordinate—a system that differs in this respect from that of Todzhu.

- (111) Tofan (Rassadin 1978: 186)  
*Išten-gen/ım/ge möngün ber-gen.*  
 work-GANGA.CV/POSS1SG/ money give-GAN.PST  
 ‘Because I worked, they gave me money.’

- (112) Tofan (Rassadin 1978: 186)  
*Bo hem-nı men üst-ü-n kəš-tı-m keh<sup>i</sup>-er*  
 this river-ACC I above-POSS3-ACC cross-DL.PST-1SG cross-INTRATERM.VN  
*čer čok bol-ırğa.*  
 place NEG.COP BE.AUX-ARGA.CV  
 ‘I crossed this river in the upper reaches because there was no crossing-place.’

Note that a causal meaning can also be found with *bolyaš* in Tsengel Tuvan (113).

- (113) Altai-Tsengel Tuvan (Taube 2008: 117)  
*Ton-um bayay bol-yaš doo-p joru men.*  
 overcoat-POSS1SG old become-GAŠ.CV freeze-IP.CV AUX 1SG  
 ‘Because my *ton* overcoat had become old, I’m always freezing.’

## 9. Purposive functions

Purposive functions of *-GAš* marked forms are also typically lexically restricted and found only in specific grammaticalized combinations, and again frequently are em-

bedded within a same-subject vs. different-subject formal opposition. Thus, in Todzhu one finds the sequence *-Vr deeš* in a same-subject construction grammaticalized to encode the meaning ‘in order to’, while conversely in different-subject formations, that is, when the subject of a converb clause is different from that of the base clause, one rather finds the construction *-ZIn deeš* instead in this same function. Compare (114) and (115)–(116) in this regard.

(114) Todzhu (Bergelson & Kibrik 1995: 400)

*Ava-m-ga użuraž-ır deeš men xooray čor-up-tu-m.*  
 mother-POSS1SG-DAT meet-AOR PURP I town move-PRF-DI.PST-1SG  
 ‘I went to town to meet my mother.’

(115) Todzhu (Bergelson & Kibrik 1995: 400)

*Ava-zı dıšťan-zın deeš uru-u mün-nü*  
 mother-POSS3 rest:RFLXV-3.IMP PURP daughter-POSS3 soup-ACC  
*xayındır-ıp ka-an.*  
 cook-IP.CV AUX.PFV-GAN.PST  
 ‘In order to let her mother rest, her daughter cooked the soup.’

(116) Todzhu (Bergelson & Kibrik 1995: 401)

*Čilig bol-zun deeš ool pečka-nı oda-p-kan.*  
 warm become-3IMP PURP boy stove-ACC heat-PRF-GAN.PST  
 ‘In order to get warm, the boy started up the stove.’

Note that in most Sayan Turkic varieties, the *-GAš* form of the verb *de-* ‘say’, i.e., *deeš*, has been grammaticalized in combination with the *-Vr* form of a verb in the function of a causal or purpose subordinator (or postposition) as well: Dukhan *teeš* (Ragagnin 2011: 167), Tuvan of Tuva *deeš* (Anderson & Harrison 1999: 76), Tofan *dääš* (Rassadin 1978: 179), Tsengel Tuvan *deeš* (Taube 2008: 107, song 35). See examples from Altai-Tsengel Tuvan (117) and from Dukhan (118).

(117) Altai-Tsengel Tuvan (Aydemir 2009: 124)

*Onu bir yadyırlı-ır deeš bar şülükčü-ler-i-n bar*  
 PRON3.ACC one laugh-AOR PURP all poet-PL-POSS3-ACC all  
*ıraazı-lar-ı-n bar igilči dovşūrču-lar-ı-n jı-ıp*  
 singer-PL-POSS3-ACC all igil-player dovşuur.player-PL-POSS3-ACC gather-IP.CV  
*gelir+jer=le yadyırt-ı al-vas irgin.*  
 AUX.CLOC-INTRATERM.VN+place=INS make.laugh-A.CV CAP-NEG.AOR EVID  
 ‘In order to have her laugh once, he gathered all his poets, all his singers, all his Igil-players and Doshuur-players into one place, but he could not get her to laugh.’

- (118) Dukhan (Ragagnin 2011: 137)

*Gulər al-ər                      deeš    Akköl jora-an.*  
 flour   take-AOR                  PURP   Akköl go-GAN.PST  
 ‘He went to Akköl in order to get flour.’

Negative forms use the *-BAs* non-past negative with *deeš*, as in the following example from Central Tuvan:

- (119) Central Tuvan (Anderson & Harrison 1999: 76)

*Men olar-ga                      šaptik    kat-pas                      deeš    bažiŋ-če*  
 I       PRON3PL-DAT    obstacle   add-NEG.AOR       PURP   house-ALL  
*kir-ip                      kel-di-m.*  
 enter-IP.CV   AUX.CLOC-DI.PST-1SG  
 ‘I came inside in order to not disturb them.’

In Central Tuvan, in addition to the *-Vr* form, this same converb form of “say” can be used with a verbal complement in the conditional to form purposive constructions as well.

- (120) Central Tuvan (Delger-Ool 1960: 123)

*Bis    siler-niŋ=bile    biče    xööreš-se                      deeš    kel-gen                      ulus bis.*  
 we    you.PL-GEN=INS    little    converse-COND       PURP   come-GAN.PST    people 1PL  
 ‘We came in order to talk with you.’

In the early sources on Sayan Turkic, e.g., Katanov (1903), this element, grammaticalized in the construction *-Vr tääš*, is found in a purposive function.

- (121) i. Karagass from Katanov (1903)    ii. Karagass from Katanov (1903)

*küräž-ir                      tääš                      ñämgär-ir    tääš*  
 fight:RCP-AOR       PURP                      feed-AOR       PURP  
 ‘um zu kämpfen’                      ‘um (ihn) zu ernähren’ (Menges 1959–1960: 128)

- (122) Soyot from Katanov (1903) (Menges 1959–1960: 128)

*...qudala-ar    tääš    id-a                      pār-di.*  
 free-AOR       PURP    send-A.CV    AUX.INCH-DI.PST  
 ‘... begann (Leute) auszusenden, um zu freien.’

However in Radloff’s Soyot lect, the purposive construction with *tääš* may take the converb in *-ARGA.CV*, i.e., etymologically a dative-case marked form of the *-Vr* participle.

- (123) Soyon from Radloff's
- Proben*
- (Menges 1959–1960: 128)

*t'a(a)l-arya*                      *tääš*  
 wage.war-ARGA.CV    PURP  
 'in order to wage war' *-arya* = < \*-*ar-ya* <\*-INTRATERM.VN-DAT

In Central Tuvan and some other lects, *deeš* can be used in causal functions as well, but this typically involves past (participle) marked verbal complements, as in the following examples (124)–(125):

- (124) Central Tuvan (Anderson & Harrison 1999: 36)

*Kizil-dan ün-gen deeš öör-üp tur men*  
 Kyzyl-ABL leave-GAN.PST SUBORD be.happy-IP.CV STAND.AUX 1SG  
 'I am happy to be leaving Kyzyl.'

- (125) Tuvan, unspecified (Tažibaeva 2012: 210)

*Eki nom tip ber-gen-iñ deeš sejee*  
 good book find:CV AUX.BEN-POSTTERM.VN-POSS.2SG PURP you:DAT  
*čettir-di-m.*  
 thank-DI.PST-1SG  
 'Thank you for having found a good book for me.'

In this particular function in Central Tuvan, *deeš* has further grammaticalized into a postposition and may now take a nominal complement as well (126).

- (126) Central Tuvan (Anderson & Harrison 1999: 36)

*Sen deeš kel-di-m.*  
 you because come-DI.PST-1SG  
 'I came because of you.'

Note that this *-GAš* form of "say" is lacking in Jungar Tuvan, and instead the *-p* converb form of *de-*, i.e., *de-p*, is used in Jungar Tuvan, based on the complete lack of *deeš* in the texts in Mawkanuli (2005) and confirmed by an anonymous reviewer.<sup>16</sup> In the few instances where one would expect it, based on Tuvan of Tuva norms, one finds *dep* instead.

16 As pointed out about by an anonymous reviewer, this lack likely reflects Kazakh influence in Jungar Tuvan. However, as pointed out by a second anonymous reviewer, it is at least in principle possible that the form in Jungar Tuvan is an archaism showing that it split from the rest of Sayan Turkic before the *deeš* construction was grammaticalized, rather than having replaced an earlier form with a pseudo-archaic form copied from Kazakh. So it is possible that this gap is not innovative or contact triggered, but rather a retention from an earlier state in this peripheral Sayan Turkic variety.

- (127) Jungar Tuvan (Mawkanuli 2005: 135: 98)

*Oyna-ar de-p ol bar-dī.*  
 play-AOR say-IP.CV s/he go-DI.PST  
 ‘She went there in order to play.’

### 10. Case-marked forms of -GAŠ

Case-marked forms of the -GAŠ converb are also attested across different Sayan Turkic varieties in a wide range of functions, most overlapping with unmarked forms of the same converb. Jungar Tuvan stands out as the Sayan Turkic variety that uses ablative marked forms of -GAŠ in the widest range of functions, typically associated with -GAŠ alone, i.e., -GAŠtan. These -GAŠtan forms occur perhaps even more frequently than simple -GAŠ forms do in Jungar Tuvan texts in Geng Shimin (2000) or Mawkanuli (2005). These examples include ones in same-subject/ narrative-propulsion contexts (128)–(131).

- (128) Jungar Tuvan (Mawkanuli 2005: 149: 16)

*Urūy-lar bar-yaš-tan güskö-nūŋ žem-in dažī-ir irgin.*  
 girl-PL go-GAŠ.CV-ABL mouse-GEN food-POSS3.ACC carry-AOR EVID  
 ‘The girls go there and carry away the mouse food.’

- (129) Jungar Tuvan (Mawkanuli 2005: 149: 18)

*Aškījak bar-gaš-tan ool-dar-nī gag-ar irgin.*  
 old.man go-GAŠ.CV-ABL child-PL-ACC hit-AOR EVID  
 ‘The old man goes there and beats the boys.’

- (130) Jungar Tuvan (Mawkanuli 2005: 105: 10)

*Ooŋ bilen arga išd-i-nge apar-ip baza gak-kaš-tan*  
 PRON3:GEN with forest inside-POSS3-DAT take-IP.CV and beat-GAŠ.CV-ABL  
*bir žeže surax-tar sura-p dur irgin.*  
 one some question-PL ask-IP.CV STAND.AUX EVID  
 ‘So having taken them to the middle of the forest and beaten them, (the soldiers) asked them some questions.’

- (131) Jungar Tuvan (Mawkanuli 2005: 150: 19)

*Ool-dar bar-gaš-tan börü-nü ölür-ür irgin.*  
 boy-PL go-GAŠ.CV-ABL wolf-ACC kill-AOR EVID  
 ‘The boys go and kill the wolf.’

The use of the ablative-marked converb functioning in a similar manner to the simplex one in Jungar Tuvan also includes formations expressing a relationship of immediate anteriority between the event of the converb clause and that of the base clause (132).

- (132) Jungar Tuvan (Geng Shimin 2000: 51, 25)

*Unuun dur-yaš-tan ool-dar-ın-ya aytı-ydur.*  
 then stand-GAŠ.CV-ABL son-PL-POSS3-DAT say-NARR/EVID  
 ‘Then she stood up and spoke to her sons.’

Ablative-marked forms of the *-GAš* converb may also encode simultaneity in Jungar Tuvan (133)–(134).

- (133) Jungar Tuvan (Mawkanuli 2005)

*Iyi žibe dur-gaš-tan žok giži gel-be-en de-er-de*  
 two thing stand-GAŠ.CV-ABL no person come-NEG-GAN.PST say-AOR.PRTCPL-LOC  
*üš orus šerig yıştag-nıñ murmu-un-da baalıg-dı*  
 three Russian soldier winter.camp-GEN south-POSS3-LOC mountain.pass-ACC  
*až-ır šab-ıp žoru-y bar-dı irgin.*  
 cross-INTRATERM.VN gallop-IP.CV move-A.CV AUX.TLOC-DI.PST EVID  
 ‘(As) both were standing there and said “No, no one has come”, the three Russian soldiers crossed the mountain pass on the front side of the winter encampment and were apparently galloping away.’

- (134) Jungar Tuvan (Mawkanuli 2005: 151, line 5)

*Žudañ bol-gaš-tan kündölögü emidirel žibe-zı-nıñ ereydep*  
 poor be-GAŠ.CV-ABL daily life thing-POSS3-GEN barely.enough  
*žaža-p gımıyla-p dur-ar irgin.*  
 make-IP.CV live-IP.CV STAND.AUX-AOR EVID  
 ‘Being poor, they eked out a bare living in their daily life.’

Also, in Jungar Tuvan, narrative strings with same-subject co-reference of both ablative-marked and unmarked forms of *-GAš* are found, showing that they largely appear to occupy the same functional domains.

- (135) Jungar Tuvan (Geng Shimin 2000: 49, 9)

*Ooñ so-on-da ada-zın-ya bir qujaq sigen al-dır-yaš*  
 PRON3:GEN after-POSS3-LOC father-POSS3-DAT one armful hay bring-CAUS-GAŠ.CV  
*bod-u arıamjı eš-keš daš qır-ın-ya qaa(y)-yaš-tan*  
 self-POSS3 rope twist-GAŠ.CV stone point-POSS3-DAT stick-GAŠ.CV-ABL  
*örttet-keš xan-ya ada-zı-n dayın sal-ıptur.*  
 burn:CAUS-GAŠ.CV khan-DAT father-POSS3-ACC again send-PST.NARR  
 ‘Then she had her father bring her an armful of hay, and she made a rope and put the rope on a stone and burned it.’

In other Sayan Turkic lects, case-marked forms of *-GAš* are also attested in a range of different functions and even in various different cases in Central Tuvan. Thus, an ablative form can encode anteriority (136)—often in same-subject contexts.

- (136) Central Tuvan of Tuva (Anderson & Harrison 1999: 79)

*Ol bažñ-nĩ tīv-al-gaš-tan dika öörüşkü bol-gan.*  
 that house-ACC find:CV-AUX.SBEN-GAŠ.CV-ABL very happy become-GAN.PST  
 ‘She got really happy after finding the house.’

However, a genitive-marked form of *-GAš*, i.e., *-GAš-tñ*, is also found in Central Tuvan, but conveys a nuance of either immediate anteriority (137) or simple same-subject narrative sequencing (138).

- (137) Central Tuvan of Tuva (Anderson & Harrison 1999: 58)

*Kel-geš-tñ men-če dolga-vīt!*  
 come-GAŠ.CV-GEN I-ALL call-PRF  
 ‘As soon as you return, give me a call!’

- (138) Central Tuvan of Tuva (Anderson & Harrison 1999: 58)

*Düün ol kiži ketti-n-e al-gaš-tñ čemnen-e*  
 yesterday that man dress-RFLXV-A.CV AUX.SBEN-GAŠ.CV-GEN eat:RFLXV-A.CV  
*al-gaš-tñ čuun-a al-gaš-tñ škola-že*  
 AUX.SBEN-GAŠ.CV-GEN wash:RFLXV-A.CV AUX.SBEN-GAŠ.CV-GEN school-ALL  
*čor-up-kan.*  
 go-PRF-GAN.PST  
 ‘Yesterday that man got dressed, ate, washed up and went to the school.’

In Dukhan, the genitive-marked form of *-GAš* may indeed also convey anteriority without same-subject co-reference across the converb and base clauses (139):

- (139) Dukhan (Ragagnin 2011: 111)

*Geje pol-sa jerle tayya gʰhʰə-sə iʰhʰ-er*  
 evening become-COND really taiga person-POSS3 drink-INTRATERM.VN  
*ji-ir je-m-ən ji-p al-yaš-təŋ*  
 eat-INTRATERM.VN eat-N.DER-POSS3.ACC eat-IP.CV take-GAŠ.CV-GEN  
*gil-ər jime jok.*  
 do-INTRATERM.VN thing NEG.COP  
 ‘As for the evening, once they have had their evening meal and drinks, taiga people don’t really have anything to do.’

However, this same formation may also occur in the narrative-propulsive/same-subject context that the unmarked form of the *-GAš* converb also typically occurs in (140).

- (140) Dukhan (Ragagnin 2011: 110)

*Irey hayrəhan-nəŋ pir gara-a-n al-gaš-təŋ pir saahar*  
 bear merciful-GEN one eye-POSS3-ACC take-GAŠ.CV-GEN one sugar

*suu-βad-ar.*

dip-PRF-AOR

‘He takes out one eye of the bear and flicks a piece of sugar (into his mouth).’

As mentioned above, tail-to-head linkage may sometimes be marked in Dukhan narratives by a verb copy in the *-GAš*+genitive form, not just the simple *-GAš* form, as in the following two strings (141)–(142). In (141), representing original text sentences 19–20, the predicative elements of sentence 19 are copied in the genitive-marked form of the converb at the beginning of sentence 20.

(141) Dukhan (Ragagnin 2011: 234)

- (19) *Irey hayrəhan tilgə-jek i<sup>h</sup>xə jettən-əp al-gaš-təŋ*  
 bear merciful fox-DIM two take.by.hand-IP.CV AUX.SBEN-GAŠ.CV-GEN  
*petək day-γədə ün-əp pa-ar jime.*  
 high mountain-DIR exit-IP.CV AUX.TLOC-INTRATERM.VN thing
- (22) *Ün-əp par-γaš-təŋ tilgə-jek ool*  
 exit-IP.CV AUX.TLOC-GAŠ.CV-GEN fox-DIM son  
*tay gür-ə-n o<sup>h</sup>rta ot ot-a-βad-ar.*  
 mountain limit-POSS3-ACC middle fire fire-V.DER-PRF-AOR

‘The bear and the little fox go hand and hand up into the high mountains. Having gone up, little fox makes a fire in the middle of the ridge.’

In (142), on the other hand, representing original sentences 25–26, the whole (end) of sentence 25 is copied in the genitive form of *-GAš* converb, to which the emphatic clitic has been appended.

(142) Dukhan (Ragagnin 2011: 234–235)

- (25) *Pir šala iŋgay al-ər.*  
 one bit further take-AOR
- (26) *Pir šala iŋgay al-gaš-təŋ=na šala*  
 one bit further take-GAŠ.CV-GEN-EMPH bit  
*iŋgay bol-ər/ən/da oh irey hayrəhan petək tayga-dan*  
 further become-ARDA.CV/POSS3/ ITJ bear merciful high taiga-ABL  
*juy-l-əp pa<sup>h</sup>t-a pa-ar jime.*  
 roll-V.DER-IP.CV sink-A.CV AUX.TLOC-INTRATERM.VN thing

‘(And the bear) takes himself a little further over there. After having taking himself a little further over there, oh, the bear goes rolling down from the high taiga.’

Finally, there are also case-marked forms of *-GAš* of the pronominal verb in its connective functions in Sayan Turkic varieties like Dukhan and Jungar Tuvan. In Dukhan, this appears as a genitive-marked form of the pronominal verb stem in its middle-reflexive derivative (143), i.e., *iŋa-n-gaš-təŋ*.



- (143) Dukhan (Ragagnin 2011: 142, 77)

*Īn-ja-n-gaš-təŋ*                      *höngen*                      *te-p*                      *ekkə*                      *je-m.*  
 that-V.DER-MED-GAŠ.CV-GEN    höngen.bread say-IP.CV    good    eat-N.DER  
 ‘And so, *höngen* bread is a good food.’

In Jungar Tuvan, one finds an *ablative*-marked form of the *passive*-marked form of the same, viz., *ī(n)ja-l-yaš-tan* ~ *anja-l-yaš-tan* (144)–(146).<sup>17</sup>

- (144) Jungar Tuvan (Geng Shimin 2000: 49, line 35)

*Ī(n)ja-l-yaš-tan*                      *ol*                      *uruy*                      *ol*                      *χan-niŋ*  
 do.thus-PSV-GAŠ.CV-ABL    that daughter that khan-GEN

<sup>17</sup> Note that in addition to the use of the passive form of the pronominal verb stem in an ablative case-marked form, Jungar Tuvan stands apart morphophonologically from the remaining Sayan Turkic lects in a number of salient ways, most likely due to prolonged contact with Kazakh. Among the atypical features found in this high-contact variety of Tuvan, which has undergone clear restructuring due to the particular sociolinguistic milieu in which it is presently embedded, is (a) that some speakers have bound subject markers in the 1SG.SUBJ *-BIn* with full consonantal assimilation and vowel harmony (vi), not clitic ones like [=men] with neither, as in Central Tuvan of Tuva (vii), and, moreover, as is still found in the speech of other Jungar Tuvan speakers as well. For more on Jungar Tuvan, see (Mawkanuli 1999, Geng Shimin 2000 and especially Rind-Pawłowski 2014, 2016)

- (vi) Jungar Tuvan (Mawkanuli 2005: 116)

*žok men aragu iš-pes-pin*  
 no I vodka drink-NEG.AOR-1SG  
 ‘No, I don’t drink vodka.’

- (vii) Central Tuvan (Field Notes)

*čok men pivo iš-pes men*  
 no I beer drink-NEG.AOR 1SG  
 ‘No, I won’t drink (any) beer.’

This restructuring also includes (b) the presence of high > low round harmony (viii)–(ix), e.g., in the present/future or the *-A/j* converb, and (c) low > low round harmony (ix)–(x) in the past-tense marker *-GAn*, or plural *-LAr* (and sometimes the dative *-GA*, but not the locative *-DA*), all of which are features that likely reflect idiolectally manifested convergence with Kazakh, resulting from the strong Kazakh-contact milieu within which Jungar Tuvan is found today.

- (viii) Jungar Tuvan (Mawkanuli 2005: 135)

*Nak ola-p büd-ör.*  
 exact do.that-IP.CV end-AOR  
 ‘It ends exactly like that.’

- (ix) Jungar Tuvan (Mawkanuli 2005: 114, 108)

*sös-tör-ü öl-gönde*  
 word-PL-POSS3 die-GANDA.CV  
 ‘his words’    ‘when/if (you) were dying’

- (x) Jungar Tuvan

*ät-tan düž-ö žügö*  
 horse-ABL descend-A.CV why  
 ‘getting off from the horse’ ‘why’

*dažir-ar-i-nan ???*

abuse-INTRATERM.VN-POSS3-ABL/INS

‘So it was that the daughter (was freed) from the khan’s humiliation...’

- (145) Jungar Tuvan (Geng Shimin 2000: 50, line 7)

*Anja-l-yaš-tan, ol qıryan qayay bičii ool-dar-i-nen qadi*  
do.thus-PSV-GAŠ.CV-ABL that old sow little son-PL-POSS3-INS together  
*jobalanjoq äme+dzuw-up jörü-y be-er ergin.*  
without.worry live-IP.CV MOVE.AUX-A.CV AUX.PRF-AOR EVID  
‘So that sow lived together with her little sons with no worry.’

- (146) Jungar Tuvan (Geng Shimin 2000: 51, 27)

*Anja-l-yaš-tan suu-ya bar-yaš serin suŋ-ni jilmilen-ip*  
do.thus-PSV-GAŠ.CV-ABL water-DAT go-GAŠ.CV cool water-ACC greedy-IP.CV  
*iž-e bašta-dī.*  
drink-A.CV begin-DI.PST  
‘So arriving at the water, she began to drink the cool water greedily.’

## 11. On the cognate converb forms of *-GAš* outside of Sayan Turkic

In this section I offer some brief comparative comments regarding the other Turkic languages that instantiate structures using a converb cognate with *-GAš* in Sayan Turkic, and that show only a subset of the domains associated with its use in Sayan Turkic varieties.

According to Efremov (1979: 65; 1980, 1984), the cognate element *-Aat* in Sakha/Yakut also functions in domains where there is first actant co-reference across the converb clause and the basic clause. However, earlier examples in Böhrtlingk (1851 [1964]: 387) suggest that the element includes an extra nuance of immediate anteriority (or quasi-simultaneity) of event actions between the converb clause and basic clause.

- (147) Sakha (Yakut) (Böhrtlingk 1851: 387)

*San-āt ä-tärä.*  
think-GAŠ.CV say-NPST:3  
‘Kaum hatte er nachgedacht, so pflegte er zu sagen [Hardly had he thought (it) when he said].’

- (148) Sakha (Yakut) (Böhrtlingk 1851: 387)

*Uot-u ott-öt čaynik-kä xār-ī simän ū-nu*  
fire-ACC set.to-burn-GAŠ.CV teapot-DAT snow-ACC stuff-AN.CV water-ACC  
*orgup-put-tara.*  
make.cook-PST-3PL

‘Sobald sie das Feuer angezündet hatten, stopften sie Schnee in den Kessel und brachten Wasser zum Kochen’ [As soon as they set the fire, they stuffed snow in the teapot and brought the water to a boil].’

Outside of Sayan Turkic,<sup>18</sup> the *-GAč* form does not appear to show a similarly strong tendency to a skewed distribution within narratively equal events, or to be a type of quasi-coordination in a same-subject like pattern, i.e., in a putative switch-reference system. Although many uses found in the literature of these languages in fact do have subject co-reference across relevant clauses, nevertheless this tendency is not as statistically strong as in the Sayan Turkic languages, and many more counter-examples can be found where there is no such subject (or first actant) co-reference across the converb clause and the basic clause. Therefore, even while clearly describing a narrative sequence and the typical “propulsive” or plot advancing meanings characteristic of this converb (Johanson 1995: 327), these forms nevertheless show no first actant co-reference across the clauses. Thus, in Uzbek for example, one finds examples with the converb *-gač* (149) below, without subject co-reference, but with a semantic relation of anteriority from the first to the second converb clause, and also a nuance of possibly either immediate anteriority or quasi-simultaneity from the second converb clause to the base clause.

(149) Uzbek (Johanson 1995: 326; Kononov 1960: 243)

*Tun ket-gač tâñ âqâr-gač quş-lar sayra-y bâşla-di.*  
 night pass-GAŠ.CV dawn whiten-GAŠ.CV bird-PL sing-A.CV begin-DI.PST  
 ‘The night had passed and it had dawned, the birds began to sing.’

In Modern Uighur, the element *-GAč* encodes temporal simultaneity of the converb clause with the base clause (150), while the dative case-marked derivative of this, *-GAčqa* (151), rather encodes the semantic attenuation of reason/cause (de Jong 2007: 195; Friederich & Yakup 2002: 209).

(150) Modern Uighur (de Jong 2007: 195)

*U tamaq ye-gäč gezit oqu-di*  
 PRON3 meal eat-GAŠ.CV newspaper read-DI.PST  
 ‘He read a newspaper while eating.’

(151) Modern Uighur (de Jong 2007: 195)

*Sughuq bol-yač-qa biz öy-din či-mi-du-q.*  
 cold be-GAŠ.CV-DAT we house-ABL leave-NEG-DI.PST-1PL  
 ‘Because of the cold, we did not leave the house.’

The corresponding cognate element is also not limited to same-subject contexts in Karaim, where an augmented form of the cognate element in the form *-GAçox* appears to mark adverbial subordination (Kocaoğlu & Firkovičius 2006: 23; Musaev

18 Perhaps better stated as outside of Siberian Turkic varieties, if, what Efremov says about contemporary Sakha-Yakut is correct, and the language has changed in this regard or expanded the functional domain of the converb since Böhlingk recorded it.

1964: 301), encoding either anteriority (152) or immediate anteriority/quasi-simultaneity (153) between the event semantics encoded by the converb clause with respect to the base clause.

(152) Trakai Karay [Karaim] (Kocaoğlu & Firkovičius 2006: 23)

<i>kel-geçox</i>	<i>tuy-ğaçox</i>
come-GAÇOX.CV	hear-GAÇOX.CV
‘when one has come’	‘when one has heard’

(153) Karaim (Musaev 1964: 301)

<i>Da</i>	<i>edi</i>	<i>kel-gäçox</i>	<i>Avram</i>	<i>Mitsri-gä</i>	<i>da</i>	<i>kör-dü-lär</i>
EMPH	BE.AUX:DI.PST	come-GAÇOX.CV	Avram	Egypt-DAT	EMPH	see-DI.PST-3PL
<i>mitsri-lar</i>	<i>ol</i>	<i>katün-ni</i>	<i>ki</i>	<i>körklü</i>	<i>ol</i>	<i>asträ.</i>
Egyptian-PL	that	wife-ACC	COMP	beautiful	that	very

‘As soon as Avram went to Egypt, the Egyptians saw that his wife was very beautiful.’

In literary Bashkir, *-GAs* converb forms can also have anteriority or causal semantics (Juldashev 1981: 311). It is also possible that in some instances, as in (154), Bashkir exhibits an example of a quasi-finite use of *-käš* as well.

(154) Bashkir (Johanson 1995: 331)

<i>Bëđ</i>	<i>kěj-ën-ëp</i>	<i>toqsay-ðar-ibið-ðï</i>	<i>aθ-ïp</i>	<i>räxmät</i>	<i>äyt-ëp</i>
we	dress-RFLXV-IP.CV	bag-PL-POSS1PL-ACC	hang-IP.CV	thank	say-IP.CV
<i>küreš-käš</i>	<i>ular</i>	<i>bëð-gä</i>	<i>yul</i>	<i>öjrat-ëp</i>	<i>yëbär-ðë-lär.</i>
say.farewell-GAŠ.CV	they	we-DAT	way	teach-IP.CV	send-DI.PST-PL

‘We dressed, shouldered our bags, thanked them and said goodbye; and they showed us the way.’

Tatar shows a wide range of functional domains of the cognate converb element, but even so this converb is not nearly as multi-functional in Tatar as it typically is in the Sayan Turkic varieties. However, in Tatar this converb element does have a number of the same functions that are seen in Sayan Turkic, including (potentially) an emergent or quasi-same-subject coordinative function, as well as marking anteriority in both same- and switch-subject contexts (155)–(156), causal semantics (157, *bul-gaç*), and even possibly semi-finite uses in Tatar literature at least (158, *tor-gaç*). The semantic domain of this converb in Tatar also appears to include the conditional (159), which *might* be an extension of the adverbial temporally subordinate “when” semantics found in other languages, or express a logical event anteriority condition (Burganova et al. 1993: 224–225, 229).

- (155) Tatar (Burganova et al. 1993: 224)

*kayt-kač bel-ü*  
 return-GAŠ.CV know-VN  
 'to know after having returned'

- (156) Tatar (Burganova et al. 1993: 225)

*Ėti-se šěhër-gě küč-kěč Xösěyeniya mëdrěšë-se-neŋ*  
 father-POSS3 city-DAT move-GAŠ.CV Hussainija madrasa-POSS3-GEN  
*bašlangič klass-lar-ın-a ker-ě.*  
 beginning class-PL-POSS3-DAT enter-AOR  
 'When his father moves to the city, he enters the beginning classes at Hussainija madrasa.'

- (157) Tatar (Burganova et al. 1993: 224)

*Kolxozčï-lar kil-ep ker-gěč keně is-e-m-ě*  
 kolkhoz.member-PL come-IP.CV enter-GAŠ.CV only remember-AOR-1-DAT  
*tös-te yalgız tügel-e-m.*  
 fall-DI.PST alone NEG-AOR-1SG  
 'The kolkhoz members returned and only then do I remember I was not alone.'

- (158) Tatar (Burganova et al. 1993: 229)

*Doşman üt-měs di-p išan-dï-k Armiya batır bul-gač*  
 enemy go-NEG.AOR say-IP.CV believe-DI.PST -1PL army hero be-GAŠ.CV  
*Sovet xalk-ı kartı- yaşe aŋar yardëmdë .*  
 Soviet people-POSS3 old-POSS3 young-POSS3 PRON3:DAT help-LOC  
*tor-gač.*  
 STAND.AUX-GAŠ.CV  
 'Because the army is heroic, the whole of the Soviet people, from the young to the old, we believed that the enemy would not come, and helped.'

- (159) Tatar (Burganova et al. 1993: 229)

*Eşlë-gěč namus belën eşlë-rgë kirek.*  
 work-GAŠ.CV conscience INS work-INF NEC  
 'If/when you work, you have to work with conscience.'

In literary Bashkir, the combination of the *-GAs* converb plus the emphatic *=ta* appears to have conditional semantics (Juldashev 1981: 311), whereas this same formal combination rather has immediate anterior semantics in Tatar (160).

- (160) Tatar (Burganova et al. 1993: 229)

*kayt-kač=ta ęyt-ü*  
 return-GAŠ.CV=EMPH say-VN  
 'to say immediately after returning'

Perhaps the closest analog to the frequency with which the *-GAš* converb occurs in texts and has the range of functional domains typical of Sayan Turkic is to be found in Western Yugur, a.k.a. Saryg Uygur or Yellow Uygur (Malov 1957, 1967; Tenišev 1976, Nugteren & Roos 2006; Roos 2000). The element that is cognate with Sayan Turkic *-GAš* in Western Yugur is called the “coordinative gerund” by Geng Shimin & Clark (1992/1993: 204) or the “лишь только/когда” converb (the “just” or “as soon as/when”) by Malov (1957: 187). It is very frequent in the texts in Malov (1967). The range of its functions includes anteriority and/or narrative propulsion in same-subject contexts (161), immediate anteriority (162–163), and some quasi-finite forms (164)–(165) together with reduced clitic copular forms deriving from *er-* (166). These latter may well, in forms reduced to Ø, underlie the quasi-finite uses possibly found in Tatar and Bashkir, and in the Sayan Turkic varieties.<sup>19</sup>

- (161) Western Yugur [Saryg/Yellow Uygur] (Geng Shimin & Clark 1992/1993: 198)

*Pu kel-geš sav+at-qaš=la tergen-ni alıñ-yaš*  
 this come-GAŠ.CV put.in-GAŠ.CV=EMPH cart-ACC take-GAŠ.CV  
*yus-un-ge yet-ıp ke-ptro.*  
 home-POSS3-DAT reach-IP.CV AUX.CLOC-IP.CV:NARR  
 ‘He came back, and put it in and then he took the cart and came home.’

- (162) Western Yugur [Saryg/Yellow Uygur] (Malov 1957: 187)

*Narī mañ-gaš ul’i-sına.*  
 to.that.side go-GAŠ.CV share-HORT.PL  
 ‘Let’s go to the side and share.’

- (163) Western Yugur [Saryg/Yellow Uygur] (Malov 1957: 187)

*Sür-geš xan-ıptro.*  
 drown-GAŠ.CV sweep.away-NARR  
 ‘They were swept away drowning, having drowned.’

- (164) Western Yugur [Saryg/Yellow Uygur] (Geng Shimin & Clark 1992/1993: 197)

*Bu gızıl biiq qorqa-aš tez-gen.*  
 this red cap fear-GAŠ.CV flee-GAN.PST  
 ‘It is said that these Red Caps were scared and fled.’

- (165) Western Yugur [Saryg/Yellow Uygur] (Geng Shimin & Clark 1992/1993: 197)

*Tikin-ge pay-yaš-du.*  
 thornbush-DAT tie-GAŠ.CV-DI.PST  
 ‘He tied it to a thornbush.’

19 Or these forms may rather reflect a process of non-finite to finite marking shift—a process that has come to be called *insubordination* in the literature (Evans 2007, Evans & Watanabe 2016, Robbeets 2009).

- (166) Western Yugur [Saryg/Yellow Uyghur] (Geng Shimin & Clark 1992/1993: 195)

*Alvatī sōyin-geš-ii.*  
 people happy-GAŠ.CV-COP  
 'The people were happy.'

There are even examples in Western Yugur that suggest that the *-GAš* converb is used in this language in the function of tail-to-head linkage (167)–(168).

- (167) Western Yugur [Saryg/Yellow Uyghur] (Geng Shimin & Clark 1992–1993: 197)

*Bo oy-ya yedik gel-gen-ii.*  
 this steppe-DAT reach:CV CLOC-GAN.PST-COP  
*Bo oy-ya yedik kel-geš...*  
 this steppe-DAT reach:CV CLOC-GAŠ.CV  
 'They reached this steppe. They reached this steppe and...'

- (168) Western Yugur [Saryg/Yellow Uyghur] (Geng Shimin & Clark 1992–1993: 198)

*... pēr kise yoq yer-ge yet-iptro.*  
 one person NEG.COP place-DAT reach-NARR  
*Kise yoq yer-ge jet-keš...*  
 person NEG.COP place-DAT reach-GAŠ.CV  
 'He reached a deserted place. He reached a deserted place and...'

Note that derived or further case-inflected forms of this *-GAš/-GAč* converb are attested not only in Sayan Turkic languages (exemplified in section 10 above), but also to a limited extent in non-Sayan Turkic languages. For example, according to Menges (1958–1959, 1959, 1959–1960, [1960–]1963, 1995), putative cognates to the Sayan Turkic *GAš* + case converb forms are said to be found in Literary Ojrot (Altai Turkic) in the form *-GAžIn*, which ostensibly reflects an old lative-instrumental form of *-GAč* > *-GAš*, also realized as *-GAč-tIn*.<sup>20</sup>

20 Note that Menges also remarks on a further possible cognate seen in the complex converb in Xakas *-AbAAs*, for which various etymologies have been offered: by Bang as *\*-a+bar-yač* but by Malov as *\*-Ip+al-yač* (Menges [1960–]1963: 117). However, neither of these etymological proposals adequately accounts for the long vowel and the lack of the velar/uvular consonant in Xakas without invoking an *ad hoc* eliding of the velar due to something like rapid speech. And while the omission of, for example, the *-p* converb or the aorist *-Vr* forms with these same putative roots does occur (*bar*, *al*, in Xakas), this specifically does *not* occur with these same verb stems when used with suffixes that start with a velar/uvular consonant, as this converb does. Also, according to Menges, in theory all of these *-GAč/-GAš* forms are said to correlate with Mongolic *-yat*, suggesting a possible loan source, as it is lacking in runic Turkic sources, but if in Chuvash the *uś* in *-uśan* really is cognate with the converb element under investigation in the present study, as Menges asserts (albeit in a secondary, case-marked version), this may rather reflect an accidental gap in the runic Turkic corpus, and thus it may possibly be old in Turkic after all.

## 12. Summary

The *-GAš* converb and its various derivatives in the Sayan Turkic varieties perform a wide range of functions. The *-GAš* converb may encode anteriority, including immediate anteriority, or even quasi-simultaneity of action between the event expressed by the converb clause and that of the base clause. The *-GAš* converb may also serve to encode narratively equal events in a type of semantic coordination expressed through a syntactically subordinate form as a means of conveying narrative propulsion. In certain varieties, e.g., Todzhu or Tofan, this element has been grammaticalized as a same-subject marker in a switch-reference system. Further, some Sayan Turkic varieties, especially Jungar Tuvan and Dukhan, instead have conventionalized the use of *-GAš* forms to set off chunks of discourse, by copying sequences of the finite predicative element of a sentence at the beginning of the following sentence in the *-GAš* form in a system known as tail-to-head linkage. In both Tofan and Soyot, *-GAš* has been analogically extended into domains of the *-p* converb in complex predicates or auxiliary verb constructions. Further, both causal and purposive semantics can be associated with specific grammaticalized combinations of *-GAš* forms with particular lexemes like “be” or “say” in Sayan Turkic varieties, with causal semantics often associated with a *-GAš* form of *bol-* (‘be’), and purposive semantics with a grammaticalization of a verb in the *-Vr* non-past participle or intraterminal verbal noun form followed by the *-GAš* form of *de-* (‘say’). Lastly, case-marked forms of *-GAš* can appear in many of the same functions as plain *-GAš* forms, particularly in Jungar Tuvan, Dukhan and Central Tuvan. While analogs to subsets of these functions are found in a number of other Turkic languages, especially in W. Yugur and Tatar, only in Sayan Turkic varieties is such a wide array of functional domains and frequency of *-GAš* forms found in complex sentences.

## Abbreviations

ABL	Ablative	INTRATERM	Intraterminal
ACC	Accusative	IP.CV	Converb in -Ip
A.CV	Converb -A, -I, -y	ITJ	Interjection
ADJ.DER	Adjectival derivation	LOC	Locative
ATT	Attemptive	MED	Medialis
AUX	Auxiliary	MITIG	Mitigation
ALL	Allative	NARR	Narrative
AOR	Aorist	N.DER	Noun Derivation
ARDA.CV	Converb in -Ar.dA, allows subject marking by possessives	NEC	Necessitative
ARGA.CV	Converb in -Ar.gA, allows subject marking by possessives	NEG	Negative
CAUS	Causative	NPST	Non-Past
BEN	Benefactive	NR	Near
CAP	Capabilitive	OBJ	Object
CES	Cessative	PFV	Perfective



CL	Classifier	PL	Plural
CLOC	Cislocative	POL	Polite
COLL	Collective	POSS	Possessive
COMP	Complementizer	POSTTERM	Post-terminal
COND	Conditional	PRF	Perfect
CONJ	Conjunctive	PROG	Progressive
COP	Copula	PROL	Prolative
CV	Converb	PRON	Pronominal
DAT	Dative	PRTCPL	Participle
DEM	Demonstrative	PRS	Present
DI.PST	Past form in -DI	PST	Past
DIM	Diminutive	PSV	Passive
DIR	Directional	PTC	Particle
DISTR	Distributive	PURP	Purposive
DS	Different-Subject	RCP	Reciprocal
DUR	Durative	REC	Recent
EMPH	Emphatic	RFLXV	Reflexive
EVID	Evidential	SBEN	Self-Benefactive
FOC	Focus	SS	Same-Subject
GAN.PST	Past form in -GAN	SUBJ	Subject
GAŠ.CV	Converb in GAŠ	SUBJ.VERS	Subject Version
GEN	Genitive	SUBORD	Subordinator
HAB	Habitual	TLOC	Translocative
HORT	Hortative	UNACMPL	Unaccomplished
IMP I	imperative	V.DER	Verb Derivation
INTRATERM	Intraterminal	VN	Verbal Noun

## References

- Anderson, G. D. S. 1998. *Xakas* [Khakas]. LW/M 251. Munich: Lincom-Europa.
- Anderson, G. D. S. 2001. Subject version and object version in Tofa auxiliary verb constructions. *Turkic Languages* 5: 1, 240–269.
- Anderson, G. D. S. 2004. *Auxiliary verb constructions in Altai-Sayan Turkic*. (Turcologica 51.) Wiesbaden: Harrassowitz.
- Anderson G. D. S. & Harrison, K. D. 1999. *Tyvan*. (Languages of the World Materials 257.) Munich: Lincom Europa.
- Aydemir, I. A. 2009. *Konverbien im Tuwinischen. Eine Untersuchung unter Berücksichtigung des Altai-Dialekts*. (Turcologica 80.) Wiesbaden: Harrassowitz.
- Baskakov, N. A. 1958. *Altajskij jazyk*. [The Altai language]. Moscow: AN ASSR.
- Bergelson, M. B. & Kibrik, A. A. 1987a. Sistema pereključenija referentsii v tuvinskom jazyke [The system of the switch-reference in Tuvan]. *Sovetskaja Tjurkologia* 2, 16–32.
- Bergelson, M. B. & Kibrik, A. A. 1987b. Sistema pereključenija referentsii v tuvinskom jazyke [The system of the switch-reference in Tuvan]. *Sovetskaja Tjurkologia* 4, 30–45.
- Bergelson, M. B. & Kibrik, Andrej A. 1995. The system of switch-reference in Tuva: Converb and masdar-case forms. In: Haspelmath & König (eds.) 1995. 373–413.
- Böhtlingk, O. von 1851. *Über die Sprache der Jakuten. Grammatik, Text und Wörterbuch*. Saint-Petersburg. (Reprinted 1964 in Indiana University Publications Uralic and Altaic Series 35.) The Hague: Mouton.

- Burganova, N. B. & Valiullina, Z. M. & Gaffarova, F. F. & Ganiev, F. A. & Zakiev, M. Z. & Ibragimov, S. M. & Sagitov, & Salimov, M. A. X. X. & Sepperova, R. I. & Tumaševa, D. G. & Xisameva, F. M. 1993. *Tatarskaja grammatika 2. Morfologija*. [Tatar Grammar 2: Morphology] Kazan: Tatarskoe knižnoe izdatel'stvo.
- Carnie, A. & Jelinek, E. & Willie, M. A. (eds.) 2000. *Papers in honor of Ken Hale. Working Papers on Endangered and Less Familiar Languages*. Cambridge, Massachusetts: Massachusetts Institute of Technology Working Papers in Linguistics.
- Čeremisina, M. I. ed. 1980. *Podčinenie v polipredikativnykh konstruksijakh*. [Subordination in polypredicative constructions]. Novosibirsk: Nauka.
- Deal, A. R. (ed.) 2007. *Proceedings of the 4<sup>th</sup> Conference on the Semantics of Under-represented Languages of the Americas* (SULA 4). (University of Massachusetts Occasional Papers in Linguistics 35.) Amherst, Massachusetts.
- Delger-Ool, A. K. 1960. Osnovnye tipy glagolnyx oborotov v tuvinskom predloženíi [Basic types of verbal phrases in the Tuvan sentence]. *Učenyje Zapiski Tuvinskogo Naučno-Issledovatel'skogo Instituta Jazyka Literatury i Istorii* 7, 118–126.
- Deny, Jean & Grønbech, Kaare & Scheel, Helmuth & Togan, Zeki Velidi (eds.) 1959. *Philologiae turicae fundamenta* 1. Aquis Mattiacis: Steiner.
- Dul'zon, A. P. 1960. Glagol čulymsko-tjurkskogo jazyka [The verb of Chulym Turkic]. *Učenyje Zapiski Xakasskogo Naučno-Issledovatel'skogo Instituta Jazyka Literatury i Istorii*. Abakan: Xakasskij Naučno-Issledovatel'skij Institut Jazyka Literatury i Istorii.
- Efremov, N. N. 1979. O parnosti monosubektnyx i raznosubektnyx temporalnyx polipredikativnyx konstruksii v jakutskom jazyke [On the pairing of same-subject and different-subject temporal polypredicative constructions in Yakut]. In: Čeremisina (ed.) 1980. 59–74.
- Efremov, N. N. 1980. *Složnopolodčinennye predloženiya vremeni v jakutskom jazyke* [Complex temporal subordinate clauses in Yakut]. Moskva: Nauka.
- Efremov, N. N. 1984. *Polipredikativnye predloženiya vremeni v jakutskom jazyke* [Polypredicative temporal sentences in Yakut]. Moskva: Nauka.
- Erdal, M. & Nevskaya, I. (eds.) 2006. *Exploring the Eastern frontiers of Turkic*. (Turcologica 60.) Wiesbaden: Harrassowitz.
- Erdal, M. & Nevskaya, I., & Menz, A. (eds.) 2012. *Areal, historical and typological aspects of South Siberian Turkic*. (Turcologica 94.) Wiesbaden: Harrassowitz.
- Evans, N. 2007. Insubordination and its uses. In: Nikolaeva, I. (ed.) *Finiteness: Theoretical and empirical foundations*. Oxford: Oxford University Press. 366–431.
- Evans, N. & Watanabe, H. 2016. *Insubordination*. Amsterdam: John Benjamins.
- Finer, D. L. 1985. *The formal grammar of switch-reference*. Outstanding Dissertations in Linguistics. New York: Garland Press.
- Foley, W. A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Friederich, M. & Yakup, A. 2002. *Uyghurisches Lehrbuch*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Geng, Shimin 2000. Materials of the Tuvian language in China. *Acta Orientalia Academiae Scientiarum Hungaricae* 53, 47–63.
- Geng Shimin & Clark 1992–1993. Sarig Yugur materials. *Acta Orientalia Academiae Scientiarum Hungaricae* 46, 189–224.

- Guillaume, A. 2011. Subordinate clauses, switch-reference and tail-head linkage in Cavineña narratives. In: van Gijn, R. & Haude, K. & Muysken, P. (eds.) 2011. 109–139.
- Haiman, J. 1983. On some origins of Switch-reference marking. In: Haiman & Munro (eds.) 1983. 105–128.
- Haiman, J. & Munro, P. 1983. *Switch-reference and Universal Grammar*. Amsterdam: Benjamins.
- Haspelmath, M. & König, E. (eds.) 1995. *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms, adverbial participles, gerunds*. (Empirical Approaches to Linguistic Typology 13.) Berlin & New York: Mouton de Gruyter.
- Hymes, D. H. & Bittle, W. E. (eds.) *Southwestern ethnolinguistics*. The Hague: Mouton.
- Isxakov, F. G. & Pal'mbax, A. A. 1961. *Grammatika tuvinskogo jazyka. Fonetika i morfoložija* [A grammar of Tuvan: Phonetics and morphology]. Moskva: Izdatel'stvo vostočnoj literatury.
- Jacobsen, W. H. 1967 Switch-reference in Hokan-Coahuiltecan. In: Hymes, D. H. & Bittle, W. E. (eds.) 1967. 238–263.
- Johanson, L. 1971. *Aspekt im Türkischen. Vorstudien zu einer Beschreibung des türkei-türkischen Aspektsystems*. (Studia Turcica Upsaliensia 1.) Uppsala: Almqvist & Wiksell.
- Johanson, L. 1975. Some remarks on Turkic “hypotaxis”. *Ural-Altaische Jahrbücher* 47, 104–118.
- Johanson, L. 1988. On the renewal and reinterpretation of ‘instrumental’ gerunds in Turkic. *Oriens* 31, 136–153.
- Johanson, L. 1990. Zur Postterminalität türkischer syndetischer Gerundien. *Ural-Altaische Jahrbücher Neue Folge* 9, 137–151.
- Johanson, L. 1991. Zur Typologie türkischer Gerundialsätze. *Türk Dilleri Araştırmaları* 1991, 98–110.
- Johanson, L. 1993. Typen türkischer Kausalsatzverbindungen. *Journal of Turkology* 1, 213–267.
- Johanson, L. 1995. On Turkic converb clauses. In: Haspelmath & König (eds) 1995. 313–347.
- Johanson, L. 1996. Altaische Postterminalia. *Acta Orientalia Academiae Scientiarum Hungaricae* 49, 257–276.
- Johanson, L. 2005. Tuwinische Postverbien und chaldaschische Imperative. In: Oelschlägel, A. et al. (eds.) 2005. 183–186.
- Juldashev, A. A. (ed.) 1981. *Grammatika sovremennogo bashkirskogo jazyka*. Moskva: Nauka.
- de Jong, F. 2007. *Modern Uighur*. Utrecht: Houtsma.
- Katanov, N. F. 1903. *Opyt issledovanija urjangxajskogo jazyk s ukazaniem glavnejshix rodstvennyx otnošenij ego k drugim jazykam tjurkskogo kornja* [An investigation of the Urjangkhaj language with a demonstration of its main genetic relations to other languages of the Turkic stock]. Kazan: Tipo.-lit. Imp. Universiteta.
- Kimball, G. D. 1991. *Koasati grammar*. Lincoln & London: University of Nebraska Press.
- Kocaoğlu, T. & Firkovičius, M. 2006. *Karay. The Trakai dialect*. (Languages of the World Materials 458.) Munich: Lincom Europa.
- Kononov, A. N. 1960. *Grammatika sovremennogo uzbekskogo literaturnogo jazyka*. [Grammar of modern literary Uzbek]. Moskva-Leningrad: Izdatel'stvo Akademii Nauk USSR.
- Li, N. & Lutz, D. (eds.) 2010. *Proceedings of Semantics and Linguistic Theory (SALT 20)*. Ithaca, New York.

- Mackenzie, A. 2007. Non-canonical switch-reference and situation semantics. In: Deal, A. R. (ed.) 2007. 159–170.
- Mackenzie, A. 2010. Subject domain restriction and reference tracking. In: Li, N. & Lutz, D. (eds.) 2010. 269–288.
- Mahapatra, K. & Zide, N. H. no date. *Gta?* texts. Unpublished-ms.
- Malov, S. E. 1957. *Jazyk želtyx ujgurov* [The language of the Yellow Uyghur]. Alma-Ata: Izdatel'stvo Akademii Nauk Kazakskoj SSR.
- Malov, S. E. 1967. *Jazyk želtyx ujgurov. Teksty i perevody* [The language of the Yellow Uyghur. Texts and translations]. Moskva: Izdatel'stvo "Nauka". Glavnaja redaktsija vos-točnoj literatury.
- Mawkanuli, T. 1999. Phonology and morphology of Jungar Tuvan. Indiana University Ph.D. Dissertation.
- Mawkanuli, T. 2005. *Jungar Tuvan texts*. (Indiana University Uralic and Altaic Series 170.) Bloomington: Indiana University Research Institute for Inner Asian Studies.
- Menges, K. H. 1958–1959. Die türkischen Sprachen Süd-Sibiriens, 3. Tuba (Sojon und Karayās). 1. Zur Charakteristik einer einzelnen sibirisch-türkischen Gruppe. *Central Asiatic Journal* 4, 90–129.
- Menges, K. H. 1959. Das Sojonische und Karagassische. In: Deny, J. et al. (eds.) 1959. 640–670.
- Menges, K. H. 1959–1960. Die türkischen Sprachen Süd-Sibiriens, 3. Tuba (Sojon und Karayās). 2. Zur Charakteristik einer einzelnen sibirisch-türkischen Gruppe. *Central Asiatic Journal* 5, 97–150.
- Menges, K. H. [1960–]1963. Die sibirische Türksprachen. In: Spuler, B. (ed) 1963. 72–138.
- Menges, K. H. 1995<sup>2</sup>. *The Turkic languages and peoples: An introduction to Turkic studies*. (Veröffentlichungen der Societas Uralo-Altaica 42.) Wiesbaden: Harrassowitz.
- Munro, P. 1979. *Studies of switch-reference*. (University of California Los Angeles Papers in Syntax 8.) Los Angeles: University of California Los Angeles Department of Linguistics.
- Munro, P. 1983. When 'Same' is not 'Not Different'. In: Haiman & Munro (eds.) 1983. 223–243.
- Musaev, K. M. 1964. *Grammatika karaimskogo jazyka. Fonetika i morfologija*. [Grammar of Karaim. Phonetics and Morphology]. Moskva: Nauka.
- Nichols, L. 2000. Rethinking switch-reference. In: Carnie, A. & Jelinek, E. & Willie, M. A. (eds.) 2000. 5–18.
- Noonan, M. 1992. *Lango*. Berlin & New York: Mouton de Gruyter.
- Nugteren, H. & Roos, M. 2006. Prolegomena to the classification of Western Yugur. In: Erdal, M. & Nevskaya, I. (eds.) 2006. 99–130.
- Oelschlägel, A. et al. (eds.) 2005. *Roter Altai, gib dein Echo! Festschrift für Erika Taube zum 65. Geburtstag*. Leipzig: Leipziger Universitätsverlag.
- Payne, D. 1992. Narrative discontinuity versus continuity in Yagua. *Discourse Processes* 15, 375–394.
- Pawley, A. (ed.) 1997. *Papers in Papuan Linguistics* 3. (Pacific Linguistics A-87.) Canberra: Australian National University.
- Radloff, W. 1866–1904. *Proben der Volksliteratur der türkischen Stämme Süd-Sibiriens* 1–8, 10. Saint Petersburg.
- Ragagnin, E. 2009. A rediscovered lowland Tofan variety in northern Mongolia. *Turkic Languages* 13, 225–245.

- Ragagnin, E. 2011. *Dukhan. A Turkic variety of Northern Mongolia. Description and analysis*. (Turcologica 76.) Wiesbaden: Harrassowitz.
- Rassadin, V. I. 1971. *Fonetika i leksika tofalarskogo jazyka* [Phonetics and lexicon of Tofalar]. Moskva: Akademia Nauk USSR.
- Rassadin, V. I. 1978. *Morfologija tofalarskogo jazyka v sravnitel'nom osveščanii* [The morphology of Tofalar in a comparative light]. Moskva: Nauka.
- Rassadin, V. I. 2010. *Soyotica*. Edited by B. Kempf. (Studia Uralo-Altaica 48.) Szeged: University of Szeged, Department of Altaic Studies.
- Rind-Pawłowski, M. 2014. Text types and evidentiality in Dzungar Tuvan. *Turkic Languages* 18, 159–188.
- Rind-Pawłowski, M. 2016. Nebensatzbildungen in Dzungar-Tuwinischen. Ph. D. Dissertation, Freie Universität-Berlin.
- Rising, D. P. 1992. *Switch-reference in Koasati discourse*. Dallas & Arlington, Texas: Summer Institute of Linguistics Publications.
- Robbeets, M. 2009. Insubordination in Altaic. *Uralo-Altaiskie Issledovanija* 1, 61–79.
- Roberts, J. R. 1988. Switch-reference in Papuan languages: A syntactic or extra-syntactic device? *Australian Journal of Linguistics* 8: 1, 75–117.
- Roberts, J. R. 1997. Switch-reference in Papua New Guinea: A preliminary survey. In: Pawley, A. (ed.) 1997. 101–241.
- Roos, M. 2000. The Western Yugur (Yellow Uyghur) language: Grammar, texts, vocabulary. PhD Dissertation, Leiden University.
- Šamina, L. A. 1987. *Vremennye polipredikativnye konstruksii tuvinskogo jazyka* [Temporal polypredicative constructions of Tuvan]. Novosibirsk: Nauka, Sibirskoe otdelenie.
- Schönig, C. 1997. A new attempt to classify the Turkic languages. *Turkic Languages* 2, 262–277.
- Schönig, C. 1999. The internal division of modern Turkic and its historical implications. *Acta Orientalia Academiae Scientiarum Hungaricae* 52: 1, 63–95.
- Shopen, T. (ed.) *Language typology and syntactic description 3: Complex constructions*. Cambridge: Cambridge University Press.
- Spuler, B. (ed.) 1963. *Handbuch der Orientalistik 5. Turkologie*. Leiden & Cologne: Brill.
- Stirling, L. 1993. *Switch-reference and discourse representation*. (Cambridge Studies in Linguistics.) Cambridge: Cambridge University Press.
- Taube, E. 2008. *Tuwinische Folklortexte aus dem Altai (Cengel, Westmongolei). Kleine Formen*. (Turcologica 71.) Wiesbaden: Harrassowitz Verlag.
- Tažibaeva, S. 2012. Polypredicative constructions of complex consequence in Kazakh compared with Altay Turkic and Tuvan. In: Erdal, M. & Nevskaya, I., & Menz, A. (eds.) 2012. 203–215.
- Tenišev, È. R. 1976. *Stroj saryg-jugurskogo jazyka* [The system of the Saryg-Yugur language]. Moskva: Nauka.
- Thompson, S. A. & Longacre, R. E. & Hwang, Sh. J. J. 2007<sup>2</sup>. Adverbial clauses. In: Shopen, T. (ed.) 2007. 237–300.
- Tsujimura, N. 1987. A comprehensive survey of switch-reference. University of Arizona Ph.D. dissertation.
- van Gijn, R. & Haude, K. & Muysken, P. (eds.) *Subordination in native South American languages*. (Typological Studies in Language 97.) Amsterdam: John Benjamins.

- Voinov, V. 2014. *Politeness devices in the Tuvan language*. (Turcologica 101.) Wiesbaden: Harrassowitz.
- de Vries, L. 2005. Towards a typology of tail-head linkage in Papuan languages. *Studies in Language* 29, 363–384.

# On *kim*, *kimne*:- and *kaña*:-. Three placeholder items in Dolgan

Chris Lasse Däbritz

Däbritz, Chris Lasse 2018. On *kim*, *kimne*:- and *kaña*:-. Three placeholder items in Dolgan. *Turkic Languages* 22, 286–302.

The article provides a corpus-based analysis of three placeholder items, *kim*, *kimne*:- and *kaña*:-, in the North-Siberian Turkic language Dolgan. After a brief discussion of the theoretical state of art in the research on placeholder items (Section 2), the named Dolgan placeholder items are investigated regarding their etymological origin (Section 3) as well as their morphosyntactic properties and their functional domains (Section 4). Finally, the outcomes of this analysis are evaluated against the described theoretical background (Section 5).

Keywords: Dolgan, placeholders, North-Siberian Turkic

Chris Lasse Däbritz, Universität Hamburg, Institut für Finnougristik/Uralistik, Max-Brauer-Allee 60, DE-22765 Hamburg. E-mail: [chris.lasse.daebritz@uni-hamburg.de](mailto:chris.lasse.daebritz@uni-hamburg.de)

## 1. Introduction

So-called placeholders (also known as “lexical fillers” or “oblitive nouns/verbs”) make up a linguistic category which is not yet fully described and often not even recognized in linguistic research, only in the last few years research on placeholders increased (e.g. Amiridze et al. 2010).<sup>1</sup> Roughly speaking, placeholders are understood as lexical items that replace another lexical item in a clause, and are especially important in spoken discourse (Fox 2010: 1–2). A simple example of a placeholder is the following:

- (1) English  
I wanted to know if you found a whatchamacallit, a parking space. (Fox 2010: 2)

The speaker apparently is searching for the right lexical item and fills the arising pause in speech with the lexical item *whatchamacallit*. The placeholder in question can be analyzed here as a lexicalized phrase emerging from *what you may call it* (Fox 2010: 1). English, being a language showing few inflectional morphology, has

1 This publication has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies’ Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies’ Programme is coordinated by the Union of the German Academies of Sciences and Humanities.

no placeholders showing complex morphosyntactic properties. Many languages showing more extensive inflectional morphology, however, exhibit placeholders showing inflectional morphology and more complex morphosyntax, e.g. the Tungusic language Udihe and the Caucasian languages Udi and Georgian (Fox 2010: 3–4). Also Dolgan, a North-Siberian Turkic language,<sup>2</sup> has three lexical elements, *kim*, *kimne:-* and *kaña:-*, which seem to function as placeholders in the above-mentioned sense, as the following examples (2) and (3) may show.

- (2) *Onno kim-ne:-čči-ler buō kim – part'izansk'ij iti atr'ad.*<sup>3</sup>  
 there PH-VBZ-HAB-3PL EMPH PH partisan this troop  
 'There they made this thing—this partisan troop.'  
 (ChVD\_AkEE\_198204\_SoldierInWar\_nar.019)
- (3) *Onto eder ogo-lor-u kel-em-mit ūōret-e-bit, kañ-ĩ:-bīt [...]*  
 then young child-PL-ACC come-CVB.SEQ-1PL teach-PRS-1PL PH-PRS-1PL  
 'Then the young people, after we come, we teach [them] and so on: [...]'  
 (AkNN\_KuNS\_200212\_LifeHandicraft\_conv.077)

In both examples, the speaker apparently inserts some lexical material (*kimne:ččiler*, *kim* and *kañ-ĩ:bīt*, respectively) in order to keep his speech going, either replacing some other lexical item (2) or repeating it (3). Moreover, these two examples show that *kim* and *kimne:-* on the one hand, and *kaña:-* on the other hand, seem to work in functionally different domains.

This article shall describe the named placeholder items in Dolgan, their origin and their function. In Section 2 the theoretical background for this study is provided; Section 3 deals with the origin of the named items and Section 4 deals with their morphosyntax and functional domains. Section 5 finally draws conclusions and evaluates the language-specific data against a theoretical background. The material used in this study comes from the Dolgan corpus of the research project INEL (Grammatical Descriptions, Corpora and Language Technology for Indigenous Northern Eurasian Languages; carried out by the Academy of Sciences in Hamburg),<sup>4</sup> which currently consists of 73 glossed texts (both folklore and narrative texts, as well as conversations) with 7,005 sentences and 45,399 tokens.

- 2 Dolgan, together with its closest relative Yakut (Sakha), forms the North-Siberian subbranch of the Siberian branch of the Turkic language family (Boeschoten 1998: 3). Nowadays, it is spoken by approx. 1,000 people (VPN 2010) on the Taymyr peninsula and in adjacent areas in the extreme north of Siberia (Artem'ev 2013a: 3).
- 3 The transcription follows the one developed in Johanson & Csató (1998: xviii–xxii), the only exception being the marking of vowel length. For typographic reasons the IPA symbol <: > is used instead of the macron.
- 4 [https://inel.corpora.uni-hamburg.de/?page\\_id=920](https://inel.corpora.uni-hamburg.de/?page_id=920), latest access: 23.01.2018.



## 2. Theoretical background

As was mentioned before, linguistic research often has ignored placeholders and similar phenomena. This may be because English and other western European languages often use items like English *uhmm* or German *ähm*, which are probably better analyzed as fillers than as placeholders (Fox 2010: 1) and lack morphological and syntactic marking, or even other strategies that operate in this functional domain. This does not, however, hold true for a number of other languages (see e.g. Fox 2010, Podlesskaya 2010). In spite of the apparently wide occurrence of placeholders in the languages of the world, Hayashi & Yoon (2006) and the volume on placeholder items (Amiridze et al. 2010) are by now the only available theoretical anchoring points for the cross-linguistic investigation of this phenomenon.

According to Hayashi & Yoon (2010: 37) a placeholder is a linguistic unit which, firstly, is a referential expression used by the speaker to replace another lexical item and, secondly, occupies the same syntactic slot as the omitted original lexical item would have done. Placeholder items often are used when the speaker has difficulty finding a lexical item; a central function of placeholders is, hence, to keep the speech going on and signal to the hearer that the speaker is still formulating his utterance (Fox 2010: 5–6). Placeholder items following the relevant lexical item, as in example (3), obviously behave somewhat differently, as the lexical item in question was already mentioned and can, therefore, not be sought for by the speaker. Whether placeholder items following the relevant lexical item nevertheless can be analyzed similarly to those preceding it remains an open question which can only be answered by investigating a wider range of languages (Podlesskaya 2010: 25). Therefore, in this study, I will use the terms “placeholder” and “placeholder item” in a broad sense, i.e. for a linguistic unit serving as a referential expression which substitutes for another lexical item either preceding or following it.

The etymological origins of placeholder items seem to be quite diverse. Nevertheless, some frequent sources can be observed (Podlesskaya 2010: 12–13): pronouns like Russian *eto* ‘this’, semantically bleached nouns like Armenian *ban* ‘thing’, a combination of both like Mandarin *zhe-ge* ‘this-CL’, and lexicalized constructions like English *whatchamacallit* (< *what you may call it*).

Another theoretical issue is how placeholder items are integrated morphologically and syntactically into the utterance in question. English *whatchamacallit*, on the one hand, shows no morphological and syntactic properties, whereas Dolgan *kim*, *kimne:-* and *kaña:-* clearly do. Taking a closer look at placeholders that do have morphological and syntactic properties, one can distinguish between nominal and verbal placeholders. According to Podlesskaya (2010: 13), nominal placeholders, i.e. placeholders which substitute for nominal constituents, are more common than verbal placeholders, i.e. placeholders which substitute for verbal constituents. However, this frequency hierarchy (which also implies that a language showing verbal placeholders obligatorily also has nominal placeholders) is far from proven, though it fits well with the data at present. The next question to arise is whether placeholders—

both nominal and verbal—copy the morphology and the syntactic properties of the substituted lexical item and if so, to what degree. Three examples given by Podlesskaya may give a tentative answer to that:

- (4) Russian<sup>5</sup>  
*Ja xoč-u et-o ubežat' ot nego.*  
 1SG want-1SG PH-NOM/ACC.SG.N run.away.INF from he.GEN  
 'I want... uhhh... to run away from him.' (Podlesskaya 2010: 19)
- (5) Russian  
*Ja zanimaj-u-s' et-oj kak=eë tipologi-ej.*  
 1SG deal.with-1SG-REFL PH<sub>1</sub>-INSTR.SG.F PH<sub>2</sub> typology-INSTR  
 'I deal with this, whatchamacallit, typology.' (Podlesskaya 2010: 21)
- (6) Udihe  
*Ñaŋga-da sin-e-we aŋi-nde-ze-mi kese-li-nde-ze-mi.*  
 little-FOC you-0-ACC PH-SEM-SUBJ-1SG torture-INC-SEM-SUBJ-1SG  
 'I will PH [whatchamacallit]... torture you a little.' (Podlesskaya 2010: 15; from Nikolaeva & Tolskaya 2001)

These examples show the three theoretically possible degrees of morphological adaptation of the placeholder to the replaced item: In example (4) the nominal placeholder *eto* (originally a proximal demonstrative pronoun) replaces the infinitive complement clause *ubežat' ot nego*; i.e., it does not mirror the morphosyntax of the replaced item at all—regardless of whether *eto* is analyzed as nominative or accusative. In example (5) the placeholder *et-oj* (instrumental singular feminine form of *eto*) copies the morphology of the replaced item *tipologiej* (instrumental singular of *tipologija*)<sup>6</sup> exactly. Example (6) finally shows an intermediate degree of copying of the morphology of the replaced lexical item. Whereas person and number marking *-ze-mi*, as well as the semelfactive marker *-nde*, are copied, the inchoative/inceptive marker *-li* is not copied. All this implies that there is no clear-cut pattern for the morphosyntactic behavior of placeholders. However, this also has to be proven by investigating a wider range of languages.

To sum up, one can say that placeholders from quite different etymological sources can occur both preceding and following the substituted lexical item. Furthermore, and according to the linguistic data currently available, nominal placeholders seem to be more prominent than verbal placeholders and the morphosyntactic behavior of both nominal and verbal placeholders varies significantly across languages and even within languages (see examples (4) and (5) from Russian).

<sup>5</sup> The original glossing and translation are slightly adapted.

<sup>6</sup> The gender of the noun is not expressed by the case ending in Russian, but by the lexical item itself. As *tipologija* is a feminine noun, adjectives and demonstratives pronouns have to agree with it.

### 3. Origin of the placeholder items

*Kim* is originally an interrogative pronoun meaning ‘who’ (Stachowski 1993: 147–148), which has etymological parallels all over the Turkic languages; cf. Turkish *kim* ‘id.’ (Lewis 1991: 72), Tatar *kem* ‘id.’ (Poppe 1963: 81) and Yakut *kim* ‘id.’ (Xaritonov et al. 1982: 197).

*Kimne:-* is a verb which is derived from the interrogative pronoun *kim* with the frequent verbalizer *-LA:-*,<sup>7</sup> cf. *albīnna:-* ‘to betray’ < *albīn* ‘lie, deception’ and *ha:majda:-* ‘to speak Nganasan’ < *ha:maj* ‘Nganasan’ (Stachowski 1997: 60–61).

The case of *kaña:-* is more complicated. Dolgan *kaña:-* has the Yakut cognate *χaya:-* ~ *χāyā:-* which is translated as ‘what do, what be, for what do’ (Pekarskij 1959: 3243). The correspondence of Dolgan *ñ* and Yakut *γ* ~ *ȳ* seems to be regular; cf. the lexemes Dolgan *meñi:* ‘head’ (Stachowski 1993: 178) and *añi:* ‘sin’ (Stachowski 1993: 34) vs. Yakut *meγi:* ‘brain; head’ (Slepcov 1972: 247) and *ayī:* ~ *aȳī:* ‘sin’ (Slepcov 1972: 35–36). According to Pekarskij (1959: 3243) the Yakut form can be analyzed as the interrogative stem *\*χay-* and the verbalizing element *-a:-*. The former most probably traces back to the Proto-Turkic interrogative stem *\*qa:ñ-*, as Proto-Turkic *\*ñ* regularly yields Yakut *γ* ~ *ȳ* (Johanson 1998: 106), and the existence of a nasal in the stem can also be shown by forms like Uyghur *kañu* ~ *kayu* ‘which’ (Räsänen 1969: 232) and Khalaj *qa:ñi* ‘id.’ (Johanson 1998: 106). Hence, the Dolgan form of the placeholder item *kaña:-* has to be regarded as older than the Yakut equivalent *χaya:-* ~ *χāyā:-*, which is not surprising, as Dolgan is often treated as more conservative than Yakut (e.g. Stachowski 1993: 17). Nevertheless, this is astonishing in so far, as interrogatives themselves, like Dolgan *kaya* ‘what kind of; which’ and *kaydak* ~ *kaytak* ‘how’ (Stachowski 1993: 133), look like the corresponding Yakut forms *χaya* and *χaydaχ* (Slepcov 1972: 474). But bearing in mind that the Dolgan placeholder *kimne:-* (see above) and the Yakut placeholder *tuoyta:-* (Slepcov 1972: 405) also stem from interrogative pronouns, it seems quite plausible that the placeholders *kaña:-* and *χaya:-* ~ *χāyā:-* in both Dolgan and Yakut also trace back to the combination of an interrogative stem and a verbalizer.

### 4. Morphosyntax and function of the placeholders

#### 4.1 *kim*

As *kim* is originally an interrogative pronoun and shows no further derivational morphology, it can be assumed that it behaves morphosyntactically like a nominal element. Dolgan nouns are inflected for number, case and possession; pronouns are inflected for number and case, rarely also for possession (Artem’ev 2013b: 137–138). In what follows, it shall be investigated to which extent the placeholder *kim*

7 Due to palatal-velar and labial-illabial vowel harmony as well as consonant assimilations the suffix has got 16 allomorphs; in this case *-LA:-* has the surface structure *-ne:-*, because it is attached to a stem with palatal-illabial vocalism and a stem-final nasal.

takes nominal morphosyntactic properties, and whether it copies the morphological patterns of the substituted item. Examples (7)–(9) illustrate this.

- (7) *Onton ol kim, Bal'šakov, d-i:-bin, [...]*  
 then that PH Bal'šakov say-PRS-1SG  
 'Then that, what's his name, Bol'šakov, I say, [the old Nikolaj, he was a very old man.]' (EIBK\_KuNS\_2004\_StorytellersUstAvam\_conv.016)
- (8) *[...] illa:-ččĩ e-ti-bit kim-ŋe iē, onton*  
 sing-PTCP.HAB be-PST1-1PL PH-DAT/LOC PTCL then  
*kolkhoz-ka, kulu:b-ka munn'-us-t-a:ččĩ e-ti-bit.*  
 kolkhoz-DAT/LOC club-DAT/LOC gather-RECP/COLL-REFL-PTCP.HAB be-PST1-1PL  
 '[...] we sang, in whatchamacallit, then in the kolkhoz, in the club we gathered.'
- (9) *[...] min bier-bit-im kim-mi-n, hantap-pi-n*  
 1SG give-PST2-1SG PH-POSS1SG-ACC parka-POSS1SG-ACC  
 'I gave my whatchamacallit, my parka [to the house of culture then].'  
 (AkNN\_KuNS\_200212\_LifeHandicraft\_conv.126)

In example (7) both the placeholder (*kim*) and the replaced item (*Bal'šakov*) behave morphosyntactically like nominal; i.e. they are inflected for case and number (zero-marked nominative singular),<sup>8</sup> syntactically they both function as the element expressing the subject. In examples (8) and (9) the placeholder *kim* copies the morphosyntactic properties of the substituted lexical item completely. According to the analyzed material, this pattern seems by far to be the most frequent one. To put it differently: If a nominal lexical item in Dolgan is inflected for number, case and/or possession, and is replaced by the placeholder *kim*, then the placeholder mostly reflects the morphological properties of the substituted item exactly. This also holds true for derivational morphology, as example (10) shows.

- (10) *Kim-ne:k e-ti-bit, prijomnik-ta:k, [...]*  
 PH-PROPR be-PST1-1PL receiver-PROPR  
 'We had, whatchamacallit, a receiver [...]'  
 (EIBK\_KuNS\_2004\_StorytellersUstAvam\_conv.142)

While the examples (7)–(10) are quite straightforward to analyze, the following three examples show some peculiarities which are not that easy to analyze.

8 It should be mentioned here that the nominative singular is not indicated in glossing. Nominal elements that are not glossed for case and/or number are therefore to be regarded as zero-marked for nominative singular.

- (11) *Ba:gĩ sar buollagĩna, kim buollagĩna, kuka:ki:*  
 that rough.legged.buzzard however PH however, Siberian.jay  
*buollagĩna [...]*  
 however  
 ‘That rough-legged buzzard, however, the whatchamacallit, the Siberian jay [...].’  
 (UKET\_2002\_FoxJayBuzzard\_flk.064)
- (12) [...] *kim-iŋ na:da, iāld'-im-iāk-kĩ-n*  
 PH-POSS2SG have.to be.sick-NEG-PTCP.FUT-POSS2SG-ACC  
*na:da.*  
 have.to  
 ‘[The most important thing is] one has to whatchamacallit, one must not be sick.’  
 (AkNN\_KuNS\_200212\_LifeHandicraft\_conv.116)
- (13) *Ontoniti kim, kak=ego Jarockij Konstantin [...]*  
 then that PH PH Jarockij Konstantin  
 ‘Then that, whatchamacallit, what’s his name, Jarockij Konstantin [...].’  
 (EIBK\_KuNS\_2004\_StorytellersOfUstAvam\_conv.036)

On first glance, example (11) looks like example (7), i.e. consists of a bare noun phrase *kuka:ki:* with the preceding placeholder *kim* replacing it. However, this case is complicated by the fact that *ba:gĩ sar buollagĩna* is a false start by the speaker: The story is about the rough-legged buzzard (*sar*) who is fooled by the fox. The Siberian jay (*kuka:ki:*) laughs about that, but now the jay himself is fooled by the fox. So the speaker starts accidentally talking about the buzzard here, remarks on this, and starts the sentence again, using *kim* as a placeholder for the sought-for lexical item *kuka:ki:*. Thus, the placeholder *kim* apparently has not only the function to keep the speech going in case of word-finding trouble, but also to repair a false start, which is, obviously, a very similar functional domain. Example (12) shows, how complex the syntactic contexts can be where the placeholder *kim* can be used. The particle *na:da* ‘need to; have to’ is a loan from Russian, and can have both a noun phrase and a non-finite construction as complement. In the second clause of the example, *na:da* takes a non-finite complement clause as complement. This is not, however, mirrored by the form of the placeholder, which forms a noun phrase here. Nevertheless, in both clauses the person-number reference coincides: Both complements of *na:da* show the possessive suffix of the second person singular.<sup>9</sup> So, in the one case the placeholder copies the morphosyntactic properties of the substituted item here (person-number reference), but in the other case it does not (noun phrase vs. complement clause), although the choice of a nominal placeholder could maybe

9 It should be remarked here that the possessive suffix of the second person singular does not necessarily refer to a second person singular in Dolgan. It can also be used as a device marking given and/or accessible referents (cf. Siegl 2015).

be explained by the existence of a participle taking nominal categories in the complement clause. Thus, the placeholder *kim* does not necessarily behave uniformly within one and the same occurrence. Finally, in example (13) *kim* interacts with the Russian placeholder *kak=ego* (cf. also example (5)), which shows its deep anchoring in the structure of language. Even in the context of massive Russian interference (another Russian placeholder used, Russian name) the Dolgan placeholder is kept, and is not fully replaced by a Russian one.

All the examples discussed until now showed both a form of *kim* and the replaced lexical item in the sentence. However, there are some examples where only a form of *kim* can be found:

- (14) *Kim, kim kula:k-tar di: di:-l-ler [...]*  
 PH PH kulak-PL EMPH say-PRS-3PL  
 ‘Uhhh, “they are kulaks”, they said [and took them away].’  
 (BeES\_1997\_HistoryOfKatyryk\_nar.053)

At least the first occurrence of *kim* cannot be connected to other lexical material in the sentence. Hence, *kim* can also be used as a placeholder, when the replaced lexical item is not part of the sentence anymore. This in turn implies the question whether *kim* can be inflected for number, case and possession in these environments too. The following examples show instances of such environments:

- (15) [...] *horok kihi-ler-in küren-ette:-bit-ter iti kimie-ke.*<sup>10</sup>  
 some human-PL-POSS2SG flee-FREQ-PST2-3PL that PH-DAT/LOC  
 ‘[...] some people<sup>11</sup> fled to that whatchamacallit...’  
 (BeES\_1997\_HistoryOfKatyryk\_nar.070)

- (16) “*Öl-lü-büt agaj, kim-ŋi-n kiste:*”, [...].  
 die-PST1-1PL just PH-POSS2SG-ACC hide.IMP.2SG  
 ‘“We are dying, hide the whatchamacallit”, [she said to me].’  
 (LaVN\_KuNS\_1999\_MusicRepressions\_conv.056)

In both (15) and (16), there is apparently an inflected form of the placeholder *kim* without the replaced lexical item. Hence, it seems not to be obligatory in Dolgan that the replaced lexical item occurs in the same sentence as the placeholder *kim*. Finally, examples (15) and (16), as well as all the other examples, clearly show that *kim*, originally meaning ‘who’ and, thus, referring to animate referents, does not differentiate between animate and inanimate referents anymore. Neither are there instances of the interrogative pronoun *tuök* ‘what’, i.e., the inanimate counterpart of

<sup>10</sup> *Kimie-* is an allomorph of *kim* and is regularly used in the dative-locative case.

<sup>11</sup> As it was mentioned above, the possessive suffix of the second person singular can be used as a referential device; this is the case here, too.

*kim*, operating in this functional domain. Hence, it can be concluded that *kim* loses its animacy-related markedness when occurring as a placeholder item

All in all, one can say that the nominal placeholder *kim* can occur in sentences both with and without the lexical item which it replaces. Regardless of whether the replaced lexical item is overtly present in the sentence, *kim* nearly always mirrors its morphosyntactic properties; however, in the latter case it remains questionable whether this can be called mirroring, as there is no overt material to mirror. Moreover, *kim* occurs even in quite complex syntactic contexts (example (11) and (12)) and also in interaction with massive Russian interference (example (13)).

#### 4.2 *kimne:-*

Regarding the fact that *kimne:-* is formed by *kim* and the verbalizer *-LA:* (see above), it seems obvious to see *kimne:-* as the verbal counterpart of *kim*. As expected, *kimne:-* indeed behaves morphosyntactically like a verb:

- (17) *Onton buöllagīna kimn-i:r iti üčügej bagaj-dīk huruj-ar.*  
 then however PH-PRS.3SG that good very-ADVZ write-PRS.3SG  
 ‘Then, however, he, whatchamacallit, writes very well.’  
 (ErTS\_AkPG\_1994\_AAPopov\_nar.087)

So, *kimne:-* shows person-number and tense marking here (third person singular, present tense) and it mirrors the morphosyntax of the replaced lexical item *hurujar*. In what follows it shall be investigated to what extent this morphosyntactic copying of the respective lexical item is obligatory and, if not, which categories are copied and which are not. There are quite many examples where a complete copying can be observed, even in very complex morphosyntactic environments like the negated habitual mood in the past tense:

- (18) [...] *kim-i da kimne:-ččī-te huok e-ti-lerē, huruj-a:ččī-ta*  
 who-ACC NEG PH-HAB-3SG NEG be-PST1-3PL write-HAB-3SG  
*huok e-ti-lerē urut.*  
 NEG be-PST1-3PL before  
 ‘[No, I did not hear that somebody should sing in radio], they didn’t whatchamacallit anybody, they didn’t record.’ (EIBK\_KuNS\_2004\_StorytellersUstAvam\_conv.098)

Even in complex discourse contexts, as in example (19a)–(19c), this holds true:

- a. PoPD: *A bihigi ogonñor-bu-n gītta užē*  
 and.R 1PL husband-POSS1SG-ACC with already  
*texnikum-tan kimne:-bit...*  
 technical.school-ABL PH-PTCP.PST  
 ‘And my husband and I, already from the technical school we whatchamacallit...’

- b. KuNS: *Bil-s-ibik-kit?*  
 get.to.know-RECP/COLL-PST2-2PL  
 ‘You got to know each other?’
- c. PoPD: *Bul-s-ubut* *e-ti-bit, [...].*  
 find-RECP/COLL-PTCP.PST be-PST1-1PL  
 ‘We have found each other, [...]’  
 (PoPD\_KuNS\_2004\_Life\_conv.047–049)

The first speaker (PoPD) apparently has trouble finding the words, and uses the non-finite form *kimne:bit*. The second speaker (KuNS) interrupts and proposes the finite form *bilsibikkit*, but PoPD continues his sentence with a complex verb form where the form *bulsubut* has the same morphosyntactical shape as *kimne:bit*.

However, there are many instances where the morphosyntax of the replaced lexical items is not completely, or even not at all, mirrored. As it was observed for the nominal placeholder *kim*, there are examples where the replaced lexical item is not visible at all:

- Ol ihin gini kepset-i-n kimne:-bit-tere buolla itigirdik.*  
 that because.of 3SG story-POSS3SG-ACC PH-PST2-3PL EMPH so  
 ‘Therefore they whatchamacallit his stories like this.’  
 (PoNA\_AkPG\_1994\_MPXarlampiev\_nar.060)

For this example (20) the context is the following: the speaker tells about a radio journalist who came from Yakutia to the Dolgans and brought Yakut material to broadcast. In the sentence at hand, he is about to explain what was done with this material, but apparently does not find a proper expression for it. Neither in the following sentences does it become clear what was meant, as the person conducting the interview asks something different. The next examples (21)–(22) show various complex instances of partial copying of the morphosyntax of the replaced item:

- (19) *Ilim ütt-e kimne:-tek-ke, kömölös-tök-püne, [...].*  
 net place-CVB.SIM PH-PTCP.COND-DAT/LOC help-TEMP-1SG  
 ‘When I whatchamacallit help them placing the nets, [I am helping the old men].’  
 (KiMN\_19900417\_Milkmaid\_flk.019)
- (20) *Imit-e:ri gïn-nak-tarïna, de kimn-i:l-ler, kïh*  
 soften-CVB.PURP make-TEMP-3PL well PH-PRS-3PL girl  
*ogo-loru-n ïgïr-a:ččï-lar togo ere.*  
 child-POSS3PL-ACC call-HAB-3PL why INDEF  
 ‘When they are about to soften [it], they whatchamacallit, they call their daughters for some reason.’ (BeES\_2010\_HidePreparation\_nar.033)



Example (21) can be interpreted in two different ways. Either *ilim ütte kimne:tekke* and *kömölöstökpüne* are two different, independent clauses, in which case the placeholder would function like in example (20); i.e., the lexical item which it replaces would not be overtly realized. Or it is possible to regard *kömölöstökpüne* as the lexical item which is replaced by *kimne:tekke*. From my point of view, the latter alternative is more likely, as there is clearly no pause between the two word forms in the audio file. That means that the morphology of the replaced item is not completely mirrored by the placeholder; *kömölöstökpüne* is a finite form (temporal mood, first person singular) which is built up of the (deep) morphemes *-TAK* and *-BInA*. *kimne:tekke* in turn is a non-finite form (conditional participle, dative-locative case) which is made up of the (deep) morphemes *-TAK* and *-GA*. Indeed, *-TAK* is etymologically the same item, and the two constructions have a similar reading, namely a temporal and/or conditional one. In either case, it can be stated that the placeholder does not copy the morphosyntax of the replaced item completely here. In example (22), finally, the person-number markings of the replaced lexical item and the placeholder are identical. However, the placeholder does not show the mood marking, which is seen at the replaced item.

To sum up, it can be stated that the placeholder *kimne:-* behaves very similarly to the placeholder *kim*. It can occur in contexts both with and without the corresponding lexical item being overtly realized and it often, but not necessarily mirrors the morphosyntactic properties of the replaced item. Hence, *kimne:-* can be both formally and functionally regarded as the verbal counterpart of the nominal placeholder *kim*.

#### 4.3 *kaña:-*

First of all, it must be stated that *kaña:-* occurs far less often in the corpus (22 tokens) than *kim* or *kimne:-* (244 and 123 tokens). *kaña:-* is a verb and, hence, should be expected to show verbal morphology. The following examples show the contexts in which *kaña:-* occurs.

- (21) [...] *prav'it'elstva d'e ehîe-ke üčügej-dik kömölöh-üö, kañ-îe.*  
           government well 2PL-DAT/LOC good-ADVZ help-FUT.3SG PH-FUT.3SG  
           ‘[...] the government, well, will help you and so on.’  
           (BeES\_1997\_HistoryOfKatryk\_nar.081)

- (22) *Taŋn-a-güt, kañ-i:-güt.*  
       dress-PRS-2PL PH-PRS-2PL  
       ‘You dress yourselves and so on.’  
       (ChSA\_KuNS\_2004\_ReindeerHerding\_conv.077)

- (23) *Ba ogo-lor-u utu-t-um-a:ri-lar, kaña:-m-a:ri-lar [...]*  
 this child-PL-ACC sleep-CAUS-NEG-CVB.PURP-3PL PH-NEG-CVB.PURP-3PL  
 ‘In order not to let their children sleep and so on, [our fathers and their fathers had this tale, a tale especially for children.]’ (UkET\_2002\_FoxJayBuzzard\_flk.003)

Apparently *kaña:-* behaves in all three examples uniformly. It occurs directly after the lexical item to which it belongs, and it reflects the morphosyntactic properties of the latter completely, even in complex verb forms like in example (25). According to the translations given by native speakers, *kaña:-* has the rough meaning of ‘and so on’ (the original translations given by the speakers are in Russian, hence its equivalent *i tak dalee*, *i tak prochee* etc.). In almost all other instances in the corpus *kaña:-* behaves exactly the same. One example stands out, but can be explained anyway.

- (24) a. *I tam ja kömölöḥ-ö:ččü-bün [...] razdelivaj-d-ī:*  
 and.R there.R 1SG.R help-HAB-1SG fillet.R-VBZ-CVB.SIM  
*kañ-ī:*  
 PH-CVB.SIM  
 ‘And there I help [with the cutting of the fish] filleting and so on.’
- b. *Kañ-ī:*  
 PH-CVB.SIM  
 ‘And so on.’ (BeSN\_2009\_Family\_nar.051-052)

So in (26a) *kaña:-* shows the same behavior as in the previous examples. In (26b) the form of *kaña:-* seems to stand alone, however, it is probably a simple repetition of the same form in (26a). Hence, the form *kañī:* in (26b) refers to the lexical item *razdelivajdi:* in (26a), and remains as an anaphoric instance of the former element.

At this point, so-called co-compounds (cf. Wälchli 2005) should be mentioned, as some Turkic languages exhibit them and some of them seem to work similarly to *kaña:-*; cf. e.g. Turkish *doktor-moktor* ‘doctor or the like’ (Wälchli 2005: 168). In Dolgan, co-compounds are formed by combining two semantically similar lexical items which then take on a more general meaning (Stachowski 1997: 86). An example of the usage of these co-compounds is the following.

- (25) *Hogotogun üöške:-bit, ine-ti-n-aga-ti-n*  
 lonely arise-PTCP.PST mother-POSS3SG-ACC-father-POSS3SG-ACC  
*öjdö:-böṭ La:jku die-n it olor-but*  
 remember-PTCP.NEG Laajku say-CVB.SEQ dog live-PST2.3SG  
 ‘There lived a dog named Laajku, which grew up lonely and did not remember its parents.’ (BeVP\_1970\_Laajku\_flk.001)

In this example, the combination of the lexemes *ine* ‘mother’ and *aga* ‘father’ apparently means ‘parents’ in a more general sense. From a formal point of view, these

co-compounds do indeed behave like the placeholder *kaña:-*, as they are also built up of two lexical items which show the same morphosyntactic properties. Functionally, however, they seem to fulfill different purposes than *kaña:-* does. Most often two lexemes are combined in order to express a semantically more general lexeme (*mother* and *father* for *parents* or *brother* and *sister* for *siblings*). Moreover, both components—at least in Dolgan—are always full lexical items, whereas *kaña:-* is lexically empty. Constructions like the Turkish one mentioned above do not occur. Therefore, it can be said that co-compounding is a phenomenon in Dolgan which is formally similar to the use of the placeholder *kaña:-*, but operates in different functional domains.

To sum up, *kaña:-* occurs far less often than *kim* or *kimne:-*; it always stands after the lexical item to which it belongs and shows the morphosyntactical properties of the latter completely. Though *kaña:-* is formally a verb, its meaning is hereby close to the English phrases *and so on*, *et cetera*.

#### 4.4 Similarities and differences in usage

From a formal point of view, it has to be mentioned that *kim* is a nominal placeholder, whereas *kimne:-* and *kaña:-* are verbal placeholders; i.e., the former one replaces nominal lexical items and the latter ones replace verbal lexical items. The most striking difference between *kim*, *kimne:-*, on the one hand, and *kaña:-*, on the other hand, is apparently their position in the clause. The former ones precede the lexical item they belong to, while the latter one follows it directly. (This does not, of course, hold true for the instances where *kim* and *kimne:-* occur without belonging to any lexical item.) Furthermore, *kim* and *kimne:-* often reflect the morphosyntactic properties of the replaced lexical item, but not obligatorily. (Partial mirroring, or no mirroring at all—which mostly can be explained then—is also observed.) *kaña:-*, however, does it consistently in every instance. Moreover, *kaña:-* occurs strikingly less often (21 tokens) than *kim* and *kimne:-* (244 and 123 tokens).

Regarding the functional domains of the placeholders under investigation, it also has to be stated that *kim* and *kimne:-* behave quite uniformly, and differ from the usage of *kaña:-*. The former are used in contexts where the speaker has trouble finding words but apparently wants to use lexical material in order to signal to the hearer that s/he is working on his/her utterance. *kaña:-*, in contrast, does not occur in those contexts, but always follows a lexical item with full meaning and is equivalent to English ‘and so on’. That means that the speaker makes up an open amount of lexical items which can be connected to the already named item.

The following table sums up the formal and functional characteristics of the three placeholders.

Table 1. Formal and functional characteristics of the placeholders

	<i>kim</i>	<i>kimne:-</i>	<i>kaña:-</i>
Nominal/verbal	nominal	verbal	verbal
Place in the clause (with respect to the relevant lexical item)	before	before	after
Possibility of standing alone	yes	yes	no
Copying/mirroring of morpho-syntax	yes, but not obligatorily	yes, but not obligatorily	yes
Context	trouble finding words	trouble finding words	adding possible similar lexical items

### 5. Conclusion and further discussion

First of all, it can be stated that the Dolgan material supports the impression that placeholders in many languages are more complex in terms of both form and function than, e.g., English *whatchamacallit*. The investigation of the Dolgan material supports the definition of a placeholder made by Hayashi & Yoon (2010), as all three placeholders are referring entities and are integrated into the syntactic structure of the relevant clause like the substituted lexical item is, or would have been (see also discussion below). Three theoretical issues were pointed out in Section 2 which need more research in a wider range of languages: position of placeholders wrt. the lexical item, possible hierarchy of nominal and verbal placeholders, morphosyntactic behavior of placeholders. In what follows, these issues shall be discussed on the basis of the investigated Dolgan material.

In Dolgan there are two placeholders, *kim* and *kimne:-*, preceding the respective lexical item and one placeholder, *kaña:-*, following it. The former ones obviously satisfy Hayashi & Yoon's definition of a placeholder (see above); even if they do not reflect the morphosyntactic properties of the substituted lexical item, they take the same syntactic slot, a good example being (12). *kaña:-* behaves in the same way, however, it remains to be determined whether it really is a referential expression. Obviously, *kaña:-* does not refer to the same entity as the relevant lexical item does, as it does not replace the latter; nevertheless, it makes up an open number of similar items; i.e., it is a referring entity. Therefore, also *kaña:-*, as a placeholder following the lexical item, is to be analyzed as a placeholder in *sensu stricto* according to Hayashi & Yoon's definition. Nevertheless, it obviously functions differently than *kim* and *kimne:-*. Whether this is a language-specific feature or a universal tendency remains an open question.

Regarding a possible hierarchy of nominal and verbal placeholders, the investigated Dolgan data is not unambiguous. On the one hand, the verbal placeholder *kimne:-* is derived from the nominal placeholder *kim*, which strongly supports Podlesskaya's (2010: 3) claim in favor of a hierarchy. On the other hand, the verbal placeholder *kaña:-* lacks—at least synchronically—a clear nominal counterpart. Whether the interrogative stem *\*qa:ñ-* can be interpreted diachronically as such a counterpart remains an open question, and needs more cross-linguistic investigation.

Hence, on the basis of the investigated material, no final conclusion can be drawn about whether there is a hierarchy of nominal and verbal placeholders. *kimne:-*, being derived from *kim*, and *kim*, occurring twice as often as *kimne:-*, however, point in the direction that Podlesskaya is right in claiming that nominal placeholders are more common and, hence, higher within a frequency hierarchy than verbal placeholders.

The morphosyntactic behavior of the three placeholders in Dolgan obviously displays great variation. For all of them the tendency towards complete copying of the morphosyntax of the substituted lexical item has been observed. This holds especially true for *kaña:-*, where all instances show complete copying. *kim* and *kimne:-* also often show complete copying, but obviously not obligatorily (cf. examples (14) and (22) respectively). When the latter ones mirror the morphosyntax of the substituted item only partly, this seems to be independent from certain grammatical categories; e.g., *kimne:-* can be marked for the habitual mood (example (18)) or not (example (22)) without any obvious reason for choosing one of the options. Altogether, that implies that the morphosyntactical behavior of placeholders is not necessarily consistent in one language, but can show considerable variation.

Finally, I hope this article increases the understanding of this phenomenon from both a language-specific and a theoretical-typological point of view. Language-specifically it may clarify some peculiar features of the complex Dolgan morphosyntax, and from a typological and theoretical point of view it may shed some light on the nature and status of placeholder items in general. Nevertheless, it has to be mentioned that the research on placeholder items calls for more cross-linguistic investigation—this study may have done its bit for Dolgan.

### Acknowledgements

I would like to thank the anonymous reviewers and—in alphabetical order—Alexandre Arkhipov, Florian Jark and Beáta Wagner-Nagy for their valuable comments on an earlier draft of this article. Without them the article would not be the same as it is now. It goes without saying that any remaining mistakes and obscurities are my sole responsibility.

### Abbreviations

1	first person	COLL	collective
2	second person	COND	conditional
3	third person	CVB.PURP	converb of purpose
ABL	ablative	CVB.SEQ	sequential converb
ACC	accusative	CVB.SIM	converb of simultaneous action
ADVZ	adverbializer		
CAUS	causative	DAT	dative
CL	clitic	EMPH	emphatic

F	feminine	PL	plural
FOC	focus	POSS	possessive
FREQ	frequentative	PROPR	propriative
FUT	future	PRS	present
GEN	genitive	PST	past
HAB	habitual	PTCL	particle
IMP	imperative	PTCP	participle
INC	inceptive ~ inchoative	R	Russian item/code-switch in(Dolgan) discourse
INDEF	indefinite	RECP	reciprocal
INF	infinitive	REFL	reflexive
INSTR	instrumental	SEM	semelfactive
ITER	iterative	SG	singular
LOC	locative	SUBJ	subjunctive
N	neuter	TEMP	temporal (mood)
NEG	negative	VBZ	verbalizer
PH	placeholder		

## References

- Amiridze, Nino & Davis, Boyd H. & MacLagan, Margaret (eds.) 2010. *Fillers, pauses and placeholders*. (Typological studies in languages 93.) Amsterdam: Benjamins.
- Artem'ev, Nikolaj M. 2013a. *Dolganskij jazyk. Učebnoe posobie dlja obščeeobrazovatel'nyx učreždenij. 1. Vvedenie, obščie voprosy, fonetika i grafika, leksika* [Dolgan language. Textbook for secondary schools. 1. Introduction, general questions, phonetics and graphemics, lexicon]. Saint-Petersburg: Almaz-Graf.
- Artem'ev, Nikolaj M. 2013b. *Dolganskij jazyk. Učebnoe posobie dlja obščeeobrazovatel'nyx učreždenij 2. Morfologija* [Dolgan language. Textbook for secondary schools. 2. Morphology]. Saint-Petersburg: Almaz-Graf.
- Boeschoten, Henrik 1998. The speakers of Turkic languages. In: Johanson & Csató (eds.) 1998, 1–15.
- Fox, Barbara. 2010. Introduction. In: Amiridze et al. (eds.) 2010, 1–9.
- Hayashi, Makoto & Yoon, Kyung-eun 2006. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formation trouble. *Studies in Language* 30: 3, 485–540.
- Hayashi, Makoto & Yoon, Kyung-eun 2010. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formation trouble. In: Amiridze et al. (eds.) 2010, 33–66.<sup>12</sup>
- Johanson, Lars. 1998. The history of Turkic. In: Johanson & Csató (eds.) 1998, 81–125.
- Johanson, Lars & Csató, Éva Á. (eds.) 1998. *The Turkic languages*. London & New York: Routledge.
- Johanson, Lars & Csató, Éva Á. 1998. Notes on transcriptions and symbols. In: Johanson & Csató (eds.) 1998, xviii–xxiii.
- Lewis, Geoffrey 1991. *Turkish grammar*. Oxford & New York: Oxford University Press.
- Nikolaeva, Irina & Tolskaya, Maria 2001. *A grammar of Udihe*. Berlin: Mouton de Gruyter.

12 Hayashi & Yoon 2010 is a shorter reprint of Hayashi & Yoon 2006.

- Pekarskij, Eduard K. 1959. *Slovar' jakutskogo jazyka 3* [Dictionary of the Yakut language 3]. Reprint of the original edition 1907–1930. Moscow: Akademija Nauk SSSR.
- Podlesskaya, Vera 2010. Parameters for typological variation of placeholders. In: Amiridze et al. (eds.) 2010, 11–32.
- Poppe, Nicholas 1963. *Tatar manual: Descriptive grammar and texts with a Tatar-English glossary*. (Indiana University publications / Uralic and Altaic series 25.) Bloomington: Indiana University.
- Räsänen, Martti 1969. *Versuch eines etymologischen Wörterbuchs der Türksprachen*. (Lexica Societas Fenno-Ugricae 17, 1.) Helsinki: Suomalais-Ugrilainen Seura.
- Siegl, Florian 2015. The non-possessive use of px.2p in Nganasan and Dolgan—a reappraisal. *Finnisch-Ugrische Mitteilungen* 39, 67–100.
- Slepcov, Pjotr A. 1972. *Jakutsko-russkij slovar'. 25300 slov* [Yakut-Russian dictionary. 25300 words]. Moscow: Sovetskaja énciklopedija.
- Stachowski, Marek 1993. *Dolganischer Wortschatz*. (Prace językoznawcze 114.) Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Stachowski, Marek 1997. *Dolganische Wortbildung*. Kraków: Księgarnia Akademicka.
- Wälchli, Bernhard 2005. *Co-compounds and natural coordination*. (Oxford Studies in Typology in Linguistic Theory) Oxford: Oxford University Press.
- Xaritonov, Luka N. & Korkina, Evdokija I. & Petrov, Nikolaj E. & Ubrjatova, Elizaveta I. 1982. *Grammatika sovremennogo jakutskogo literaturnogo jazyka 1. Fonetika i morfologija* [Grammar of the modern Yakut literary language]. Moscow: Nauka.
- VPN 2010 = *Vserossijskij perepis' naselenija. 4. Nacional'nyj sostav i vladenie jazykami* [All-Russian census of the population. 4 National composition and command of languages]. [//www.gks.ru/free\\_doc/new\\_site/perepis2010/croc/Documents/Vol4/pub-04-05.pdf](http://www.gks.ru/free_doc/new_site/perepis2010/croc/Documents/Vol4/pub-04-05.pdf), latest access: 29.01.2018.

## TURKIC LANGUAGES

### Information and Guidelines for Contributors

**Contents.** TURKIC LANGUAGES contains articles, review articles, reviews, contributions to discussions (remarks and replies), reports, surveys of publications, lists of publications received, and addresses of contributors.

**Original contributions.** As a rule, only original contributions are accepted. Submission of a manuscript implies that it has not been previously published and that it is not under review elsewhere.

**Language.** The preferred language of publication is English. Authors not writing in their mother-tongue are asked to have their manuscripts checked with respect to language and style.

**Length of contributions.** The suggested length of articles and review articles is 10–30 pages.

**Submission of manuscripts.** Authors are requested to send their manuscripts in electronic form, preferably as a doc-file together with a pdf version. Manuscripts should have a separate title page with the author's name and full mailing address. Please submit your contribution electronically, by e-mail or on disk. Manuscripts will not be returned.

**Style sheet.** Contributors are requested to comply with the style sheet which will be sent on request.

**Permissions.** It is the author's responsibility to request any permission required for the use of material owned by others.

**Abstracts.** Authors submitting articles and review articles are requested to include an English abstract of about 100 words.

**Referee system.** The journal uses a referee system in order to ensure a high standard. The anonymous reviewers will only read copies of the contributions without any identification of the author. While they may give suggestions for improvements, they are not obliged to specify the motivations for their decision to recommend or not to recommend publication in each case.

**Proofs.** Authors will receive one copy of the first proof, which, duly corrected, should be returned normally within two weeks. Changes against the original manuscript must be avoided or kept to a minimum.

**Free author copies.** Authors will receive a free copy of the volume in which their article appears. Additionally they will receive the print-ready pdf-file version of their article.

Prof. Dr. Dr. h.c. Lars Johanson, Turkic Languages,  
Institute of Slavic, Turkic, and Circum-Baltic Studies,  
Johannes Gutenberg-University Mainz, D-55099 Mainz, Germany  
E-mail: johanson@uni-mainz.de



