IDEAS BEHIND SYMBOLS – LANGUAGES BEHIND SCRIPTS

**Studia uralo-altaica 52**

**Redigunt**
Katalin Sipőcz
András Róna-Tas
István Zimonyi

# Ideas behind symbols – languages behind scripts

Proceedings of the 60th Meeting of the
Permanent International Altaistic Conference (PIAC)
August 27 – September 1, 2017
Székesfehérvár, Hungary

Edited by Ákos Bertalan Apatóczky

Szeged, 2018

Printed by: Innovariant Ltd., H-6750 Algyő, Ipartelep 4.

ISBN: 978 963 306 663 8 (printed)
ISBN: 978-963-306-664-5 (pdf)
ISSN: 0133 4239

# Contents

Preface ........ 7

Tatiana Anikeeva
Turkic Manuscripts and Old-Printed Books of the Lazarev Institute of Oriental Languages: Exploring the History of Oriental Studies in Russia ........ 9

Chen Hao
*Bark*: A Study on the Spiritual World of the Early Türks ........ 15

Oliver Corff
Nations and Rivers: Their Status and Name in the *Qingshi Gao* Reflections on the *Draft History of Qing* as a Source ........ 21

Balázs Danka
A language behind the script A case study on the Pagan Oɣuz-nāmä ........ 29

Mihály Dobrovits
The Ogur Turks in Chinese records ........ 35

Hsiang-Tai Kao
The field research on the Manchu inscriptions in Beijing ........ 41

Bayarma Khabtagaeva
Some notes on kinship terminology in Yeniseian ........ 47

Kyoko Maezono
Japanese and Mongolian Usages of the Chinese Writing System ........ 57

Dieter Maue – Mehmet Ölmez – Étienne de la Vaissière – Alexander Vovin
The Khüis Tolgoi inscription ........ 73

Rodica Pop
The “five eyes pattern” *tavan nüden hee* ........ 91

Maria Magdolna Tatár
Bortz and Membrok, etymology of two Cuman names from the 13th century ........ 107

Hartmut Walravens
Józef Kowalewski’s Letters to Bernhard Jülg ........ 117

Tana Wu
One Language behind Two Different Scripts* ........ 135

Hülya Yıldız
New Reading Proposal on the Eastern Face, Nineteenth Line of the Bilgä Qaɣan Inscription ........ 147

# A language behind the script
# A case study on the Pagan Oɣuz-nāmä

Balázs Danka
Johannes Gutenberg University
Institute for Slavistics Turkology and circum-Baltic Studies

The Pagan Oɣuz-nāmä[1] (MS, Radloff 1890, 1891, Nour 1928, Pelliot 1930 [1995], Bang–Arat 1932 [1936], Ščerbak 1959, Danka 2016; hereinafter PON) is written in a simplified version of the Uyghur-Mongolian alphabet. The present paper will deal with the problems of reconstruction of the sound system in the language variety PON is written in.

Alphabetic scripts are designed to render sounds. Adaptation of an alphabet to a new language is almost never perfect, because the sound system of the target language the script is adapted to is different from that of the language to which the script had been developed or applied to.

An alphabet encodes the important sound types of a language. The letters used to render sounds are partly based on orthographical conventions on the one hand, and on the intuition of the scribe on the other. Consequently, we know only those characteristics of the sounds which are encoded by the letters. The aim of the present paper is to highlight this problem by a case study on PON.

Sounds consist of one or more distinctive features. The important question to ask here is: What are the important distinctive features of the sounds which are mirrored by the letters?

Vowels in Turkic may consist of the following features: ±front, ±open, ±round.

Only the positive features are marked and thus are relevant, negative features are disregarded, and considered irrelevant. If a vowel consists of none of the relevant features, it is considered neutral. Vowels may consist of more than one positive feature. Therefore, *a* is +open (−front, −round), *i* is +front (−open, −round), *u* is +round (−open, −front), *ä* is +open, +front (−round), *ü* is +front, +round (−open), etc. The most complex vowel in this regard is *ö* with all the three distinctive features being positive +front, +round, +open, and *ï* can be considered as the least complex or the neutral member of the Turkic vowel system, all of its features being negative. I would not go into the details of the question of the so-called closed *e* here. For our present analysis suffice it to say that it consists of the

1 The digital photos of the manuscript are accessible on the webpage of the Bibliothèque Nationale: http://expositions.bnf.fr/islam/gallica/turc2.htm

same distinctive features as *ä*, but it seems that the hierarchy between its features is different, +front being more relevant than +open.

Turkic consonants may be orals and nasals. Both categories may be further differentiated by the non-binary features ‘place’, such as ‘labial’ for *p* or *m,* ‘dental’ for *t,* or *n,* etc. Orals sounds can be further differentiated by the non-binary feature ‘manner’, and the binary feature ±fortis. By manner, they may be stops, affricates, fricatives, liquids and glides. Note that place features partly overlap with those observed in vowels.

East Old Turkic is the earliest known variety of Turkic languages. It is well-documented; therefore, it can be used as a basis of reconstruction of historical developments in Turkic languages.

According to the above analysis, the East Old Turkic vowel system can be described as below.

Chart 1.

| | +front | | −front | |
|---|---|---|---|---|
| | −round | +round | −round | +round |
| −open | i | ü | ï | u |
| +open | (e) | ö | | o |
| | ä | | a | |

The most complex element of the system is *ö* which all marked distinctive features ‘+open, +front and +round’. The most underspecified element is back *ï* with all of its features being unmarked. The ‘neutrality’ of *ï* is supported by numerous phonological and morphonological phenomena in Old, Middle- and modern Turkic languages.

The consonant set of East Old Turkic can be summarized as it is in the chart, based on Lars Johanson’s forthcoming work ‘Turkic’.

Chart 2.

| | | Labial | Dental | Palatal | Velar |
|---|---|---|---|---|---|
| Plosive/Affricate | Fortis | *p* | *t* | *č* | *k (q)* |
| | Lenis | *b* | *d* | *(ǰ)* | *(g) (ġ) (ɢ)* |
| Fricative | Fortis | *(φ)* | *s* | *š* | *(χ)* |
| | Lenis | *(β)* | *z (δ)* | | *(γ)* |
| Liquid/Glide | | *(w)* | *l*<br>*r* | *y* | |
| Nasal | | *m* | *n* | *n´* | *ŋ* |

The sounds in parentheses are variants of Proto-Turkic consonants represented in East Old Turkic in different phonological environments. I would like to call attention to the relatively high number of velar sounds in the table to which I shall return later.

Now let us turn to the text I label as PON. It is written in a simplified version of Uyghur-Mongolian script. It is a well-known text in Turkology, yet its value as a historical linguistic source is not fully recognized. Its language is undisputedly a Middle-Turkic variety, but its precise classification has not yet been established. One of the main problems with the text is that we do not exactly know the exact quality of the sounds described by the script.

The grapheme set of PON is established based on the palaeographical analysis of the manuscript, as it is presented in Figure 1.

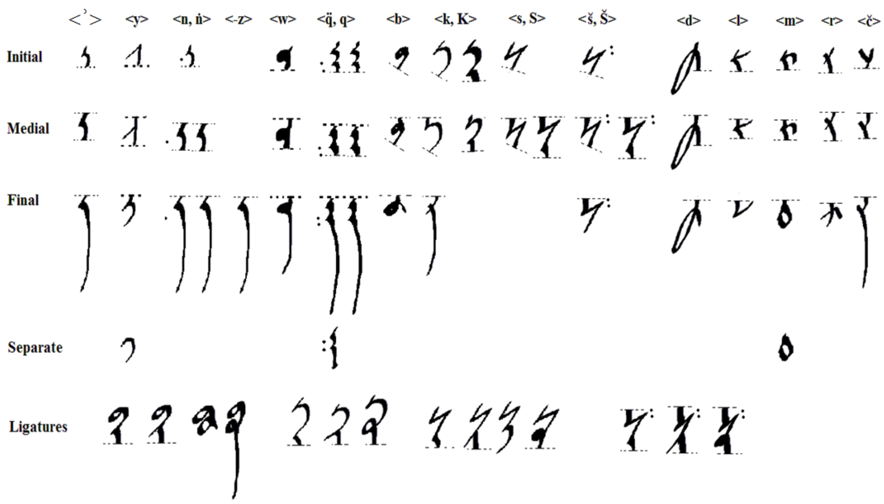


Figure 1. The grapheme set of PON

The graphemes are not listed in the alphabetical order of the Uyghur script, but according to the typological similarities of the letters. Each letter has initial, medial and final forms and only the attested separate forms of the letters are listed here. The letters <n> and <q> have variants distinguished by diacritic dots, but these variants do not distinguish separate sounds, they are used interchangeably. Three letters may be used to render vowels, <ʾ>, <y> and <w>. Some consonant letters have special ligature forms when combined with vowel-letters. Note that letters <k>, <s> and <š> have variants that look like a ligature in combination with <ʾ>, but their reading is a simple consonant. These are transliterated with a capital Latin letter.

It is clear for the first glance that the number of graphemes is far smaller than the number of sounds to be described. There are 15 graphemes (not including their positional variants) to describe 36 sounds (27 consonants and 9 vowels, including their allophones). Due to this asymmetry, there are certain graphemes which render several sounds. On the other hand, certain sounds can be rendered by more than one grapheme. The picture is further complicated by orthographic conventions of the Uyghur script, such as a word-initial vowel marked by an <’>, but the letter itself not necessarily renders an actual sound.

Three graphemes, <’>, <y> and <w> are used to render vowels in this script version. There are only a few instances in the whole text, when vowels are rendered by grapheme combinations other than word-initial <’> and either <y> or <w>. However, even these few instances are inconsistent. The attested data of the reflexes of East Old Turkic (EOT) vowels rendered by the letters of the script version of PON, is presented in Chart 3. ordered by the complexity of the vowels, i.e. from the ones having less positive features to the ones having more. The data in bold face show the typical usage of marking a vowel with a letter within PON.

Chart 3.

| | <’> | <y> | <w> |
|---|---|---|---|
| no marked features | ***ï*** | *ï* | *ï* |
| one marked feature | ***a***, *i, u* | ***i*** | ***u*** |
| two marked features | ***ä***, *e, u, ü* | *ä*, ***e*** | ***o, ü*** |
| three marked features | – | – | ***ö*** |

Based on this chart the followings can be determined: The grapheme <’> is used to render almost any vowels except open round vowels *o* and *ö*. To illustrate the phenomenon, the different instances of the EOT word *bäδük* ‘great’ are spelled as <b’dwk>, <bydwk>, <byd’k> and <b’d’k>. The first syllable *e* is marked either by <’> or <y>, while the second syllable *ü* is rendered by <w> or <’>.

If we approach from the direction of sounds, the reflex of the EOT *ï*, for example, can be rendered by any vowel-letters: <q’l’č> *qïlïč* ‘sword’, <’yq̈yr> *ayġïr* ‘stallion’, <qwdwq> *qïδïγ* ‘edge, rim’.

To sum up, we can ascertain that the script fails to render either of the features +open, +front and +round perfectly. The most consistent tendencies are marking round vowels with <w> and the most underspecified vowel *ï* with <’>. This means that the scribe had serious difficulties to render what he heard, most probably due to the reason that what he heard was a different Turkic variety from the one he knew, with quite a different vowel system.

The way of rendering consonants is no less problematic. The letters used to render liquids, glides, nasals, and the fricative *š*, are used in 1:1 correspondence. The combination <nk> is also consequently used to render *ŋ* and, on morpheme-boundaries, *n+g*.

The graphemes used to render stops and affricates are underspecified about the exact quality of the rendered consonants, hiding important developments which are already known from EOT. Thus, <b> may render *p, b* and *β*. <d> may render *t, d* and *δ*. Therefore, nothing can be told about the consonant assimilation processes of the suffix-initial *D*. Similarly, <q> and <q̈> may render *q, ġ,* and *ɣ,* in suffixes, *Ġ;* <k> may render *k, g,* and *G* in suffixes. The graphemes <s> and <-z> are in complementary distribution. In word-initial and word-internal positions only <s> occurs while <-z> occurs only in word-final position. They both may render *s* and *z*.

Let us see for example which sounds can be rendered by the grapheme <q> (freely alternating with <q̈>.

/k/ *q* plosive, velar, fortis <q̈'r'q̈> *qị̈r*$^{o}$*q* 'forty'

/k/ *χ* fricative, velar, fortis <'q> *aχ* 'Oh!'

/k/ *Ġ* plosive? velar, lenis <'d'q̈y> *aδaĠï* 'his foot'

/g/ *ġ* plosive, velar, lenis <y'lqwz> *yalġuz* 'alone'

/g/ ɣ fricative, velar, lenis <'q̈'z> *aɣị̈z* 'mouth'

vowel length <q̈'q̈'r> *qār* 'snow'

The examples show that practically any variants of the EOT *k* and g sounds can be represented by this single grapheme. The grapheme tells us only that the sound is velar in non-front syllables. Other than that the letter neither tells us anything about the sound being a stop vs. fricative or fortis vs. lenis oppositions. Interestingly enough, in the case of *qār* 'snow', the grapheme itself does not mark a sound, but vowel length. This form clearly shows an influence of the Written Mongolian orthographical practice.

Ultimately, the actual quality of the velar consonants can be presumed only based on East Old Turkic. The same holds true for the graphemes <d>, <k>, <č>, <b>. In the case of the grapheme pair <s>: <-z> we know that the sounds in question are fricatives. The sound types represented by the graphemes are summarized in Chart 4. (cf. Chart 2.)

Chart 4.

<table>
<tr><th></th><th>Labial</th><th>Dental</th><th>Palatal</th><th>Velar</th></tr>
<tr><td>Plosive/Affricate</td><td rowspan="2"><b></td><td rowspan="2"><d></td><td><č></td><td rowspan="3"><k, K><br><q>, <q̈></td></tr>
<tr><td rowspan="2">Fricative</td><td rowspan="2"><š, Š></td></tr>
<tr><td rowspan="2"><w></td><td><s, S>, <-z></td></tr>
<tr><td>Liquid/Glide</td><td><l><br><r></td><td><y></td><td>–</td></tr>
<tr><td>Nasal</td><td><m></td><td><n>, <ṅ></td><td>–</td><td><nk></td></tr>
</table>

Yet the situation is not entirely hopeless. If a consonant is spelled with a different grapheme than expected based on EOT, a phonological development or phenomenon can be attested. A word-initial *y~ǰ* fluctuation can be observed in a set of words: EOT *yaruq* vs. PON *yaruq* <yʾrwq> ~ *ǰaruq* <čʾrwq> ‘light’. A few word-initial *b-* sounds show nasalization, not only if the syllable-coda contains a nasal: EOT *buz* vs. PON *muz* <mwz> ‘ice’. Strong aspiration of word-initial *t-* can be observed in a few words. As the *t-* is marked with the grapheme <č> we cannot exclude the possibility of palatalization: EOT *taŋ* vs. PON *taŋ* <dʾnk> ~ $t^h$*aŋ* <čʾnk> A few words which had presumably word-initial *h-* in Proto-Turkic, are spelled with a word-initial <y>: EOT *är* vs. PON *yer* <yyr> ‘man’. The East Old Turkic word-internal *-δ-* is preserved in intervocalic position: EOT *aδïɣ* vs. PON *aδuɣ* <ʾdwq> ‘bear’. Word-internal *_δ-* is changed to *_y-* before consonants: EOT *qaδɣu* vs. PON *qayɣu* ‘sorrow’.

The complete research material and a new facsimile edition of the text will be soon published under the title ‘The Pagan Oɣuz-nāmä’, along with a philological and linguistic analysis’ in the series ‘Turcologica’.

## References


Bang, W. –Arat, R. R. (1936): *Oğuz Kağan Destanı.* İstanbul.

Bang, W.–Arat, R. R. (1932): *Die Legende von Oghuz Qaghan.* Berlin.

Danka, B. (2016): The ‘Pre-Islamic’ Oġuz-nāmä. A philological and linguistic analysis. Szeged (manuscript).

MS – Bibliothèque Nationale *Supplèment Turc.* No. 1001.

Nour, R. (1928): *Oughuz-namè, èpopèe turque.* Alexandrie.

Pelliot, P. (1930): Sur la légende d’Uġuz-Khan en écriture ouigoure, *T’oung Pao* 27, 247–358

Pelliot, P. (1995): *Uygur yazısıyla yazılmış Uğuz Han Destanı üzerine*, Çev. V. Köken. Ankara.

Radloff, W. (1890): *Kudatku Bilik – Facsimile der uigurischen Handschrift.* St. Petersburg, 291–292.

Radloff, W. (1891): *Das Kudatku Bilik von Jusuf Chass-Hadschib aus Balasagun.* St. Petersburg, x–xiii.

Ščerbak, A. M. (1959): *Oguz-nāme; Muχabbat-nāme.* Moskva.