Prefatory remark

The following article is a **draft version** which was presented on the occasion of the PIAC 2017, Székesfehérvár (Hungary) within a panel on the Khüis/Hüis Tolgoi inscription. The panel included also contributions by

- M. ÖLMEZ, On the discovery, the whereabouts, condition of the stones, and our expedition,
- E. de la VAISSIÈRE, The historical context to the Hüis Tolgoi inscription (uploaded on academia.edu), and
- A. VOVIN, Interpretation of the Hüis Tolgoi inscription (discussed sub: https://www.academia.edu/s/07da0d7bfb/interpretation-of-the-huis-tolgoi-inscription?source=news).

*

Schwarzenborn 16 Sep 2017
© Dieter Maue 2017

The Khüis Tolgoi inscription - signs and sounds

Dieter Maue

1 Generals
2 The sign inventory
3 The sound inventory
4 Apendices
I Signs and sounds
II Transliteration
III Transcription

1 Generals

The inscription is written in Brāhmī, the mother of all later Indian scripts and their derivatives , e. g. the Tibetan script, and the basis of five varieties which were used along the Silk roads for Khotanese, Tocharian A and B, Tumshukese, Sogdian and Uigur. The sixth variety was created for the Brāhmī inscriptions of the Mongolian steppe and their language or languages. As to the language of the KhT inscription, it appeared in the course of deciphering that we are most likely concerned with a (Para-)Mongolian language (: "(Para-)Mongolian hypothesis") as will be shown in the following contribution by Prof. A. Vovin. The (Para-)Mongolian hypothesis exercised some controlling effect on the deciphering in the narrow sense (i.e. on the identification of signs and definition of their value). To name two issues: a very special one: Initial r- should not occur in the inscription; a general one: The

underlying language should show front-back vowel harmony which might also be reflected in the sign system.



col. 11 10 9 8 7 6 5 4 3 2 1

Fig. 1: Khüis Tolgoi inscription (detail)

The Khüis Tolgoi inscription surprises with a couple of features which were unseen before in combination with the Brāhmī script:

(1) The script is written from top to bottom[1] in columns which run from right to left. (Fig. 1)
(2) The text is segmented by horizontal strokes mostly into words; occasionally morphemes are separated (regularly the plural morpheme); once a compound is divided into its constituents (bo dä| sa-t va).
(3) The graphematic entity is the syllable, not the akṣara.

[1] Traces of vertically written Brāhmī are known from coins of the late and post-Kushana period and from coins of the Gupta Empire on which the rulers are depicted in standing position accompanied by their names. While other parts of the legend are circularly arranged around the figure, the names are vertically written which was forced by the available space; for specimina s. under coinindia.com.

2 The sign inventory

The KhT Brāhmī alphabet[2] is composed of a selection of signs taken from the basis alphabet and half a dozen additional signs.

2.1 The signs from the basis alphabet

Conso-nants	Occlusive				Fricative		Semi–vowel	Nasal	Liquid	Vibrant
	unvoiced		voiced		un–voiced	voiced				
	un–asp	asp.	un–asp.	asp.						
Labial	p	p^h	b	b^h			v	m		
Dental	t	t^h	d	d^h	s			n	l	r
Retroflex	ṭ	$ṭ^h$	ḍ	$ḍ^h$	ṣ			ṇ		
Palatal	c	c^h	j	j^h	ś		y	ñ		
Velar	k	k^h	g	g^h				ṅ		
Laryn–geal						h				

Vowels	a	ā	i	ī	u	ū	e	o

Fig. 2: The KhT signs (high-lighted) taken from the basis Brāhmī alphabet

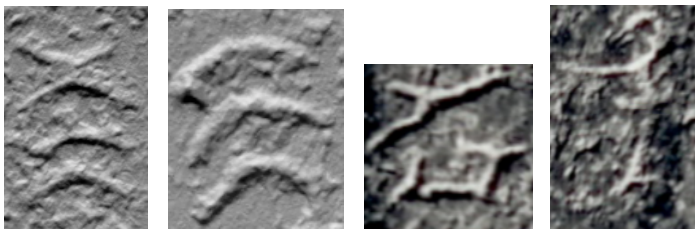
2.2 The consonant signs

The table (Fig. 2) shows the basis alphabet; the elements occuring in KhT are high-lighted (yellow). It is obvious that the KhT alphabet was deliberately and systematically organized. Unservicable signs of the basis alphabet such as retro-flexes and aspirates were dropped. The asymmetry of the velars (g but no k) is strange and most probably by chance as well as the lack of m and perhaps that of ŋ.

[2] I feel free to use the designations 'alphabet' or 'syllabary' for the list of basic signs though both expressions do not adequately describe the character of the Indian script: "Various terms have been suggested for this type of script ... There is as yet no commonly accepted term.", R. Salomon, Indian epigraphy. New York, Oxford 1998: 15.

2.2.1 The special consonant signs

On the other hand, there were obviously some sounds in the KhT language which could not be adequately represented by any of the basic signs. Therefore four special consonant signs were added, s. fig.3.



No. 1 No. 2 No. 3 No.4

Fig. 3: The special consonant signs

2.2.1.1 Special signs nos. 1 and 2



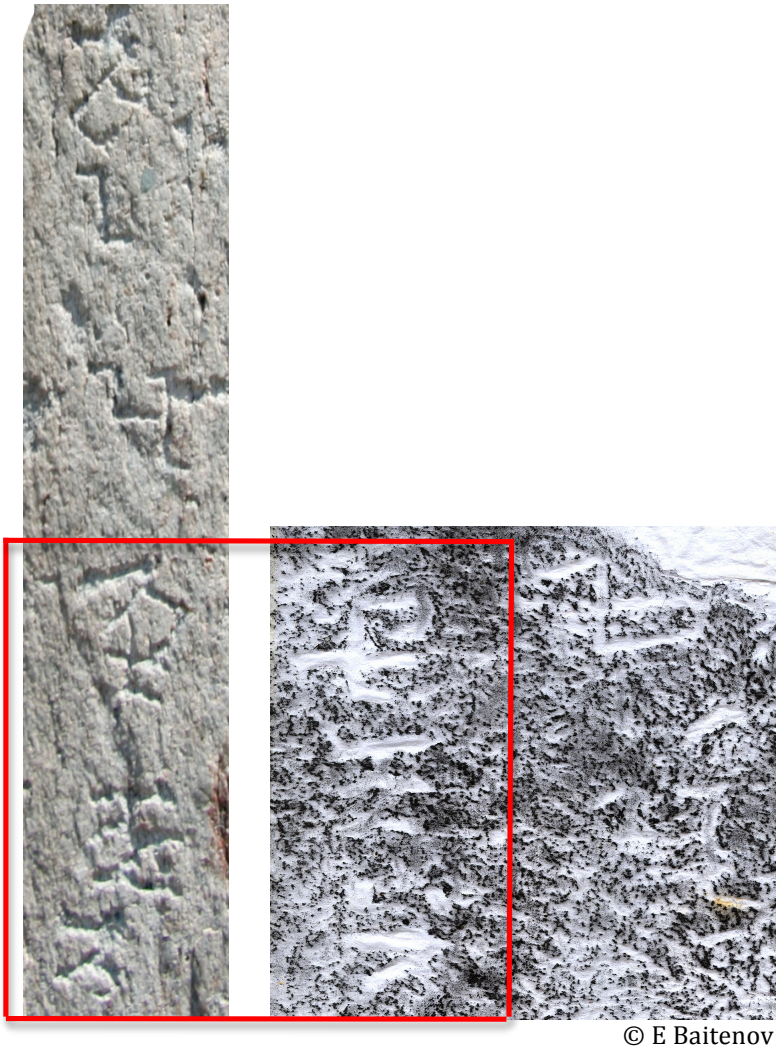
Fig. 4 Special signs nos.1 and 2



Fig. 5: Keregentas and Taraz (© Google Earth)

Signs nos. 1 and 2 form a sign group (Fig. 4) which occurs 12 or 13 times in the inscription. The determination of their sound value was crucial for the decipherment. It succeeded only through Brāhmī inscriptions which were discovered by the Kazakh scholar Eskander Baitenov: two inscriptions together with

a Bactrian fragment hail from the Keregenras valley and a third one from Taraz. (Fig. 5)

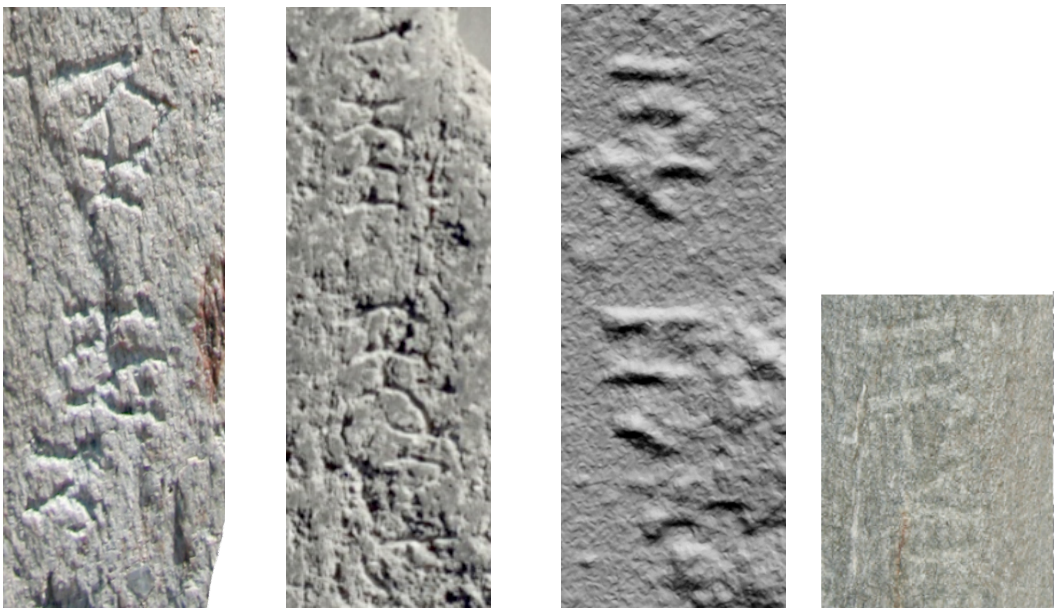


© E Baitenov

Fig, 6: The Keregentas Brāhmī inscriptions

The two Keregentas inscriptions consist of four signs each (Fig. 6): two of them are the same in both inscriptions (in the read box), the other two are different from each other. The stones most probably served as balbals; thus the inscriptions should represent the name of the killed enemy followed by his title[3] as Baitenov suggested. All other possibilities eliminated, the best and only option was the reading "**Kagan"**. Thus the rough value of the two special signs was determined: no. 1 - transliterated as k with stroke above, k̄ - stands for an unvoiced velar k and no. 2 - transliterated through g_1 - for its voiced partner. It appeared that the Keregentas sign group had direct counterparts in KhT and Bugut. The sign g_1a-n is also visible on the Taraz stone. (Fig. 7).

[3] Cf. the Ongi balbal with the inscription: išvara tarkan balbalı, cf. e. g. M. Ölmez 2012: 191.

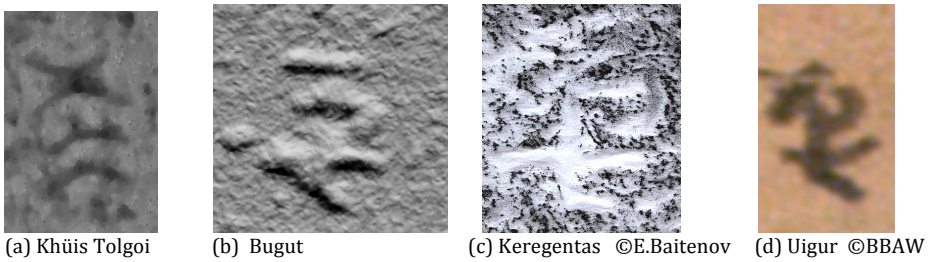


(a)Keregentas © E Baitenov (b) Khüis Tolgoi (c) Bugut (d) Taraz © E Baitenov

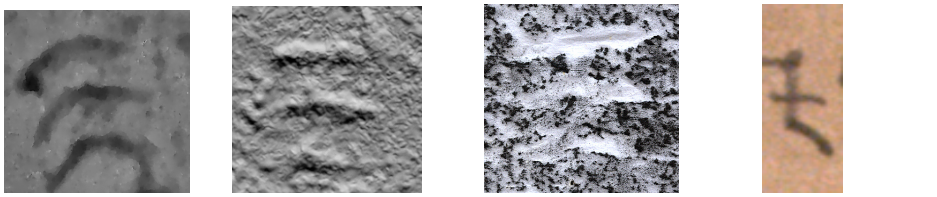
Fig. 7: "Kagan"

The heuristic value of the discovery is considerable:
(1) It is relevant for the further deciphering that words preceding kagan are principally candidates for names.
(2) The KhT inscription is certainly a secular text and the Brāhmī part of the Bugut inscription very likely so; the latter is doubtlessly no Sanskrit sūtra as was generally suspected.
(3) One and the same Brāhmī variety was in use all over the territory of the First Turkic Kaganate.
(4) The facts that (a) in the Bugut bilingual the main inscription was in Brāhmī script and non-Sogdian language and (b) that in Keregentas a Bactrian inscription existed beside the Brāhmī inscriptions - theses facts suggest that the Brāhmī script was used for writing the imperial or official language while the Bactrian and Sogdian scripts and languages had the status of "scripturae francae" and "linguae francae" in the Western and Eastern part of the First Kaganate resp.
(5) The impact on the history of Brāhmī script needs a digression:

(a) A closer look to the signs of Fig. 8 reveals that k̄a and g_1a of KhT differ from those of Bugut and Keregentas. In KhT, the three horizontal elements of g_1a are not connected; k̄a has the headstroke in common with the Bugut and Keregentas forms, but the body is shaped like KhT g_1a. Clearly, the KhT signs show a systematic reshaping; they could continue the Bugut-Keregentas forms, but hardly vice versa. Thus the KhT script must be palaeographically later than that of Bugut-Keregentas.



(a) Khüis Tolgoi (b) Bugut (c) Keregentas ©E.Baitenov (d) Uigur ©BBAW



(a)Khüis Tolgoi (b) Bugut (c) Keregentas ©E.Baitenov (d) Uigur ©BBAW

Fig. 8: The signs <k̄a> (supra) and <g$_1$a> (infra)

(b) Perhaps even more exciting is the second point. Hitherto it was believed that the corresponding Uigur Brāhmī signs (Fig. 8) were created after the Uigurs had left the steppe (A.D. 840 +) and moved to the Qočo area where they adopted the Brāhmī script under the influence of the Tocharian and Tumshukese varieties. However, the striking similarity to the Bugut-Keregentas signs teachs us that Uigur k̄a and g$_1$a were actually inherited. In other words, the Steppe Brāhmī must have survived somewhere for centuries before it reappeared in the Uigur monasteries around the Turfan oasis.

(c) From (a) we suggest that the KhT inscription postdates the Bugut inscription (c. A. D. 582). From (b) we conclude that there is no terminus post quem non derivable from the palaegraphical data. For a more accurate dating of KhT one needs other evidence.

*

We now return to the KhT velars proper. We have seen that the velars of the basic alphabet were adopted; at least g is attested. The question arises: what are the additional velars for? The answer comes from other Mongolian and Turkic writing systems: it is for the differentiation of front and back velars which are/were distinct to such an extent that different signs are/were desirable to encode them appropriately. The sensibility for this difference may be especially high in speakers of languages with front-back vowel harmony. The basic signs were used for the front velars and the new ones for the back velars.

Another question concerns the phonetic value, whether the back velars k and g were stops or fricatives. In Khüis Tolgoi, they occur in some titles: kagan (the g only there), kato(n) = katun(?), darkan with the plural darkad, and in añakay, possibly a name. As a proper name, Añakay is attested elsewhere in Chinese sources and once in Greek[4], while the titles are frequently met with in various languages and

[4] S. E. de la Vaissière, The historical context to the Hüis Tolgoi inscription. Academia.edu. note 5.

scripts so that there was some hope to get relevant phonetic informations from there. In view of the limited time I have to make the long (and rather disappointing) story short: g_1 is almost certainly the representative of the voiced back velar fricative, while for the unvoiced counterpart k both options, stop or fricative, remain open without further evidence which can perhaps be provided in the next section.

2.2.1.2 Special sign no.3



(a) Special sign no. 3 (b) tpa (c) kṣa (d) TS no. 12: χša

Fig. 8: Special sign no. 3 and (alleged) relatives

The sign, s. Fig. 8a, is attested as syllable initial or syllable final (s. Fig. 9) and occurs in the beginning, middle or end of words.

syllable initial:
(a) 7 <3a ni>
(b) 2; 6 <3ä$_1$ rä>
(c) 6 <+ 3a cä >

syllable final:
(d) 3 < ja-3 >
(e) 1; <to ro-3 >
(f) 2 <u qa-3>

Fig. 9: Words with special sign no. 3

At the very beginning of the decipherment it appeared to be simply the ligature tp-, s. Fig. 8b. But that resulted in monstrous forms like ?ukatp, ?tpani and was therefore soon abandoned.

Next, the reading kṣ or χš was guessed, the latter because there seemed to be a genetic connection with the basis ligature kṣa (Fig. 8c) or the Tumšukese special sign no. 12 (Fig. 8d), which is, however, more developed and has got an additional tail. With the equation no.3 = χš two words of the inscription suggested themselves as foreign words:

(a) χšani in the sense of 'regnal year', a wanderwort (wandering word) which is met with in several Central Asian languages[5];
(b) χši$_1$ri 'country', Tumš. χšera-, Khot. kṣīra- id.
Both concepts, 'regnal year' and 'country', being unfamiliar to the steppe people, might have been adopted together with the foreign words. But χši$_1$ri must be questioned in as far as it is limited to the Sakan languages; the Sakans played no such dominant role in the Tarim basin that their concept of 'country, kingdom' might have been attractive for being borrowed.

When the (Para-)Mongolian hypothesis was accepted, the value of the sign had to be modified again, because the native words (Fig. 9d) ja-3 and (Fig.9 f) uka-3 could hardly end on χš. Tentatively, the fricative x was guessed instead: jax and ukax could then be analysed as deverbal derivatives on -k(u) which was developed to -x in final position.[6]

It is questionable whether xani (instead of former χšani) and xi$_1$ri (instead of former χši$_1$ri), dubious as it is (s. above), are still eligible as the above supposed foreign words.

A front allophone is probable though not demonstrable. (Here one might remind of the double realisation of the velar fricative in Standard High German in the words "ich" and "ach".)

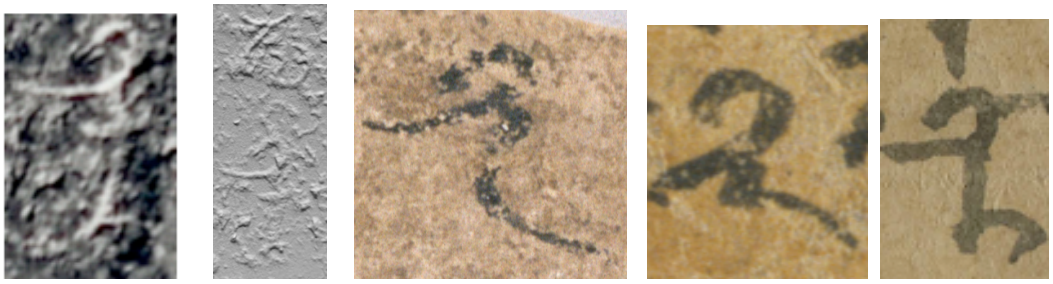
A consequence of the identification of no. 3 as velar fricative would be that the back velar k must have been a stop.

---

5 KS kṣuṇa- 'section, period of time' (DKhS 68a), TS xšana- 'Zyklusjahr' (thus Konow 1935: 50), KrorPkt kṣuna-, KharInscr. kṣun/ṇa-, TochB kṣuṃ 'regnal year' (Adams DTB^2 261: "It is likely that the three Tarim languages [KS, TS, TochB,DM] have borrowed this word, directly or indirectly, from Bactrian, either as a result of Kushan (100 BC to 200 AD) influence or influence of the later Hephthalites (first half of the sixth century AD)") Bactr. (a)χþono "'(calendar) year, (regnal) year', possibly loanword from Gk. χρόνος (A. Thierfelder apud Humbach 1966, p. 24)" (N. Sims-Williams, BactrDoc II).

6 That would not touch its status as phoneme which is proved by the opposition k vs. x in the beginning of the words, e. g. x(ani) vs. k(agan).

2.2.1.3 Special sign no. 4



(Khüis Tolgoi inscr.) (DTA TS 5 r1, © BBAW) (DTA bs7 p.2 l.2 and "Eig. Zähl.1" v3 © BBAW)

(a) Special sign no.4 (b) <tu 4a> (c) < ļ̥> (d) <v₁a> (e)<v₁e->

Fig. 10: Special sign no.4 and its (alleged) relatives

The sign under discussion (Fig.10a) is attested three times in one and the same word <tu 4a>. Being without an exact counterpart in any Brāhmī variety, it was thought to be related to one of the two signs, (1) the long syllabic l< ļ̥> (Fig. 10c) or (2) <v₁> , the bilabial w (Fig. 10d-e).

The first candidate, the long syllabic l, ļ̥,[7] is an invention of the Indian grammarians. It occurs in grammatical treatises or alphabet charts, but never in natural language texts. Excepted are some instances in the Uigur Brāhmī. There, it is used for lı or, exceptionally, for li, both usually spelled as <li, lī>. Thus the long syllabic l, < ļ̥,>, of he Uigur texts turns out to be a superfluous and sophisticated orthographical variant for which there was no other purpose than to show the scribe's intimate knowledge of the subtleties of the Indian script. In a perspicuously and economically organized script like that of the KhT inscription, such extravagances are unthinkable.

The second candidate is a non-Indian sign which occurs in Tumshukese, Sogdian and rarely in Uigur Brāhmī; transliterated by <v₁>, it stands for the bilabial w.[8] The sign shows several variants[9]. Closest to the form of our inscription are two samples from Sogdian mss which are reproduced in Fig. 10d and e. The lower part in KhT could be an attached -r. The reading of the word Fig.10b is therefore tuwa or tuwar.

A second bilabial beside the basic sign <v> in Bodhisatva is not unusual in the non-Indian Brāhmī varieties. The basic sign is nearly exclusively used in Indian (loan) words. If the bilabial fricative w is needed in native or integrated adopted words, the special sign is applied.

---

7 Maue (Simply forgotten? academia.edu) 2014.

8 Maue 1996 (Zu den uigurischen und iranischen Brāhmī-Handschriften der Berliner Turfanfunde. In: R.E. Emmerick et al. (ed.), Turfan, Khotan und Dunhuang ... Berlin 1996. pp. 211-220. Here:) 215 seqq..

9 Maue 1996: 219.

2.3 The vowel signs

The Indian Brāhmī syllabary supplies the following vowel signs (Fig. 11): a, ā, i, ī, u, ū, e, o[10] as independent (syllable initials) or dependent characters; the latter have the form of diacritics added to the preceding consonant or consonant cluster; excepted is -a which is inherent and needs no expression.

Independent	a, ā, i, ī, u, ū, e, o
Dependent	(-a), -ā, -i, -ī, -u, -ū, -e, -o

Fig. 11: Vowel signs of the basis alphabet

Of these vowel signs only very few are used in KhT, Fig. 12.

	front unrounded	central	back rounded
close	(i, -i)		**u?,-u,-ū***
central	(e, -e)		**-o**
open		**a*, -a**	

Legend: a = independent vowel sign; -o = dependent vowel sign; * = occurring once; (i,-i) = lacking

Fig. 12: Basic vowel graphemes in KhT

In KhT, long vowels are not written except for one long u, <-ū>, in the word <tu rū-g> as variant of <tu ru-g> with short u. Apart from that two oddities are observed and must be discussed:
(1) The graphemes for i and e were not adopted, s. 2.3.1.
(2) Initial vowel graphemes are extremely rare, s. 2.3.2.

2.3.1 The special vowel signs

The basic alhabet vowel signs <i> and <e> are absent. But there are met with two diacritics, the special signs nos. 5 and 6, which have no counterparts in the Indian alphabets.

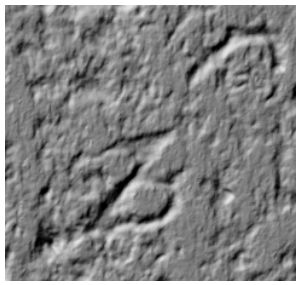
The first one, no. 5, in Fig. 13a attached to d-, consists of two small superscript strokes, looking like a gravis and an acutus (or like a háček split into two parts).

The second one, no. 6, in Fig. 13b over c-, is apparently derived from the first one by adding a loop to the upper end of the acutus.

[10] Not counted the syllabic liquids r̥ and l̥.



(a) KhT 1: <dä>



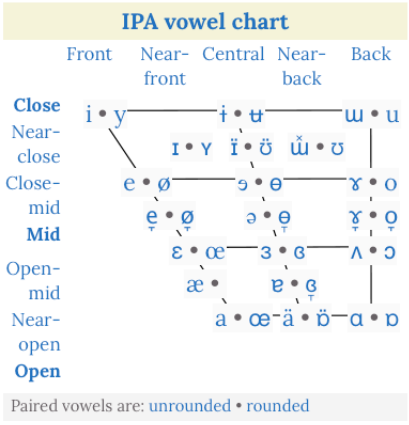
(b) KhT 4: <cä$_1$>



(c) THT 1154 a3 : <yä>

Fig. 13: The special vowel signs nos. 5 and 6

The latter is unknown elsewhere; the former resembles the trema or two dots diacritic, traditionally transliterated through ä[11] (a with trema).[12] It occurs in all non-Sanskrit varieties of the Turkestan Brāhmī, in Fig. 13c with y. In the languages concerned it represents[13] some unrounded central vowel on the scale between mid (ə) and close (ɨ). (s. Fig. 14). The spelling of the foreign word 1;8 <bo dä sa-t va> 'Bodhisat(t)va' removes any suspicion that <ä> could be something completely different in KhT.



© Wikiwand (http://www.wikiwand.com/en/IPA_ vowel_chart_with_audio)

Fig 14: IPA vowel chart

---

11 In imitation of the graphic representation. The transliteration the symbol ä (a with a trema) was introduced by the Tocharianists E. Sieg and W. Siegling and generally adopted by other philologies, though it is easily misunderstood as *a with umlaut* [ɛ].

12 In alphabetic order: KhT 6 + χša cä , 10 + g$_1$a-[] ×*ä$_1$ jä*[-]n , 9 ×*ä yu* , 11 b[]| tä *b*a k̄a , 11 b*ä* tä -g , 3 bä gä-y , 1 b*ä* tä$_1$, 3 ba *yä* , 9 bä tä$_1$ *h*ä$_1$-*ñ* , 1;8 (without daṇḍa) bo dä | sa-t va , 3 bo dä , 4 cä cä$_1$ ra , 1 dä g*ä$_1$* -n , 6 ha(Or: h*ä*)-r gä-n , 10 ja ya bä [?] , 5; 10 (ex coni.) nä$_1$ rä , 8 *pu-g tä*-g cä , 1;8 *ś*ä$_1$ ñä-n , 7 χša nä , 2;6 χšä$_1$ rä.

13 Except for Khotanese, perhaps.

As being modified from the first diacritic we transliterate the second one through < ä$_1$>. It often occurs together the simple <ä> in the same word.[14] The design of the diacritic suggests that its value is related to that of the simple <ä>, too. The question is: how exactly?

Perhaps the word 1 <dä gä$_1$-n> will help to answer the question. It contains both diacritics and represents most likely the word, which is known from Turkic tigin or tegin[15] 'prince, a legitimate son (or grandson) of a ruling kagan'[16] and is widely spread in Asian languages. The relevant material was collected and analyzed by G. Doerfer[17] and V. Rybatzki.[18] The present focus is on its pronunciation around the 7th cent. So we tried to select from the huge mass of material those instances which are (a) not too distant from envisaged time, (b) probably based on first hand information and (c) written in a script which represents and differentiates the vowels. Chinese was included on the base of the reconstructed sound system (here: Pulleyblank's Middle Chinese). Despite of various compromises not more than a dozen or so samples turned out to be suitable for comparison (Fig. 15).

	Language	Script	Orig. script and/ or Translit./Transcr.	
1	Sogdian	Sogdian	tk'y-nt[19]	t(ə?)gin-t
2	KhT-Language	KhT Brāhmī	dä gä$_1$n	
3	Chinese	Chinese	特勤[20] tè qín[21]	EMC dək gɨn

[14] In alphabetic order: KhT 3 + ×ä$_1$ *t*ä$_1$-n , 4 + *b*[] tä$_1$, 10 + g$_1$a-[] ×*ä$_1$ jä*[-]n , 1 b*ä* tä$_1$, 9 bä tä$_1$ *h*ä$_1$-*ñ* , 4 cä cä$_1$ ra , 1 dä g*ä$_1$* -n , 5 du gä$_1$-d , 6 hä$_1$-g bä$_1$-j , 4 *j*ä$_1$ *lo na*-r , 5. 10(ex coni.) nä$_1$ rä , 6 pa<->l χšä$_1$-r , 4 pu-g tä$_1$<->g , 10 u bä$_1$-j , 1;8 śä$_1$ ñä -n , 2;6 χšä$_1$ rä.

[15] Note on the vocalisation of the'prince' in the scientific literature:
To my best knowledge, G. Schlegel (La stèle funéraire du Teghin [sic] Giogh et ses copistes et traducteurs chinois, russe et allemand. Helsinki 1892. (Mém. de la Soc. Finno-Ougr.; 3.)) was the first to see that 特勤 "Tehk'in [PY tè qín, DM] est sans doute la transcription du mot *Tikin*, qui est un titre affecté chez les Turcs aux chefs de Horde." (o.c.6) He adds in a note: "Mon savant collègue à Utrecht, le professeur M. Th. Houtsma, m'écrit que *Tikkin, Tikin* ou *Téghin* est le titre ordinaire de *Prince* dans les dialectes turcs." Without further discussion Schlegel henceforce uses the form "**téghin"** which was adopted by V. Thomsen in his groundbreaking work on the decipherment of the runiform script (Inscriptions de l'Orkhon. Déchiffrées par V. Th. Helsinki 1896.) W. Bang (Zur Erklärung der köktürkischen Inschriften. WZKM 12, 1896) has **"tägin"** while F. W. K. Müller decides for "tegin" (e. g. Ein Doppelblatt aus einem manichäischen Hymnenbuch (Maḫrnāmag). Berlin. 1913. (AKPAW 1912; 5.)). A third spelling, **"tigin"**, was used by others (e. g. A. von Gabain, Alttürkische Grammatik. 3rd ed. Wiesbaden 1974). All three variants are still in use in different spellings. M. Ölmez (What should a new edition of the Old Turkic inscriptions look like? In: 歐亞學刊/International Journal of Eurasian Studies 2, 2015, 80-93. Here p. 85 f.) formulated that, as a rule, "the word which can be written with an i or without a vowel ... should be consistently read with closed ė." According to his spelling "tėgin" he subsumed the 'prince' to this class of words. However, the spelling in the inscriptions is nearly uniformly with <i> except for one or two instances of which I doubt that they are to be taken seriously.

[16] Clauson ED 483a.

[17] G. Doerfer TMEN vol. II § 922.

[18] V. Rybatzki, Die Personennamen und Titel der mittelmongolischen Dokumente. Eine lexikalische Untersuchung. Helsinki 2006. pp. 378-382.

[19] Bugut inscription, Y. Yoshida 2009 (unpubl.) B-2: 11.

[20] Often met with erroneous 勒 lěi, cf. e. g. Chavannes Doc 367.

			提勲[22] tí qín[23]	EMC dɛj gɨn
4	Old Turkic	runiform script	t²ig²n² t²g²n² [24]	tigin tegin [-ɛ-]??[25]
5	Bactrian[26]	Bactrian	τογινο τιγινο[27]	tə$^{?}$gin tigin
6	Sanskrit[28]	Indian	tigina-[29]	*tigin
7	Pahlavi[30]	Pahlavi	tkyn'	
8	Khotanese	Khot. Brāhmī	da gyį nä[31] da gyai na[32] di gyį na[33]	/deǵin, dejin/[34]
9	Middle Persian[35]	Manichean	ṯgyn ṯkyn ṯqyn	tegin [?][36],
10a	Uigur	Manichean	tygytl'r[37]	tigitlär
10b	Uigur	Sogdian	tk'ytk'[38]	t(i)git+kä

---

21 As to 勒 le instead of 勤 qín cf. Chavannes Doc 367 b, Liu Osttürken n. 48, recently Sanping Chen, Son of Heaven and Son of God: Interactions among Ancient Asiatic Cultures regarding Sacral Kingship and Theophoric Names. In: Journal of the Royal Asiatic Society, 12, 2002, p. 296: "Writing 勒 instead of 勤 is a common script error in current editions of almost all dynastic histories." (quoted from n.1 on the lemma "Tegin" of Wikiwand)

22 In various Turfan inscriptions dating from before the mid-7th cent. CE, cf. Y. Kasai (Die alttürkischen Wörter aus Natur und Gesellschaft in chinesischen Quellen (6. und [sic!] 9. Jh.). Der Ausgangsterminus der chinesischen Transkription tū jué 突厥. In: B. Heuer et al. (edd.), "Die Wunder der Schöpfung". Mensch und Natur in der türksprachigen Welt. Würzburg 2012. (Istanbuler Texte und Studien; 9.) pp. 81-140. Here:) 87.

23 On 10th cent. transcripts s. Hamilton 1955: 156.

24 Numerous occurrences, T. Tekin 1968: 380 s.v. tegin, 381 s.v. tigin.

25 T. Tekin o.c. transcribes mostly tigin for t²ig²n² , twice tegin for t²g²n² in KT N8, but - why? - for t²ig²n² in KČ S1.

26 Inscriptions and coin legends (of Barhatekin and his successors; series NumH) of the 8th and 9th centuries, partly together with Sanskrit text.

27 H. Humbach, (Baktrische Sprachdenkmäler. I-II. Wiesbaden) 1966. Here: I 139 a.

28 Occurring together with the Bactrian coin legends, s. Humbach 1966: 141b.

29 Humbach 1966: 60 NumH 208

30 Occurring together with the Bactrian coin legends of no. 5 (Bactrian).

31 Ch 00269, 63

32 P 2024,31;33.

33 P 5538a, 10.

34 Thus E. Hovdhaugen (Turkish words in Khotanese texts. A linguistic analysis. In: Norsk Tidsskrift for Sprogvidenskap 24, 1971, 163-209. Here:) p. 188. (1) "The alternation a – i in the first syllable in Khot. may be an attempt to render [closed, DM] Turkish /ė/." (2) "More difficult to explain is the initial *d*-." Hovdhaugen argues that (a) d- of foreign words was substituted by t- in Old Turkish while (b) "the Chinese transcriptions of this title are consistently written with **d*-." (3) The palatalization -ǵ- > -j- [= ǯ, DM] is a Khotanese development.

35 Early 9th cent., Uigur milieu.

36 References and transcription in D. Durkin-Meisterernst, Dictionary of Manichaean Middle Persian and Parthian. Turnhout 2004, p.323b which are taken from: F. W. K. Müller, Ein Doppelblatt aus einem manichäischen Hymnenbuch (Maḥrnāmag). Berlin. 1913. (AKPAW 1912; 5.)

37 Man I 31: TM 159 v5.

38 Mainz 38 A 9 (korr. 10) in D. Fedakâr (Das Alttürkische in sogdischer Schrift. Textmaterial und Orthographie. (Teil II). In: Ural-Altaische Jahrbücher NF 13, 1994, 133-157. Here:) p.134. Re-

			t*k*[]*ytl*'r[39]	t(i)git+lär
10c	Uigur	Brāhmī	tg$_1$iṃ[40]	t(i)gin
10d	Uigur	Tibetan	ti kyin[41]	tigin

Fig. 15: KhT <dä gä1-n> and parallels

The resulting **cover form** is **D I g I$_1$ n**

The consonants

Final n and the middle g are unquestionable. The initial dental occurs in two variants, t and d, as can also be observed in the case of darkan and perhaps in <dro> if = dörö and comparable with Old Turk. törö 'traditional law'. Hovdhaugen (1971: 188 f.) suggested that intitial t- might be the Turkic substitute for a foreign d- which was applied because (1) initial d- does not occur in indigenous Turkish words and (2) the foreign d- was perhaps closer to the runic t than to d.[42]

The vowels

The representation of the first vowel shows a broad variance. It is often not written at all (syncope). Otherwise there is a tendency to write i, but also Baktrian o[43] and Khotanese a[44] occur alternatively. In contrast to KhT, Khotanese and Uigur Brāhmī do not use the two dots diacritic, though it is part of their alphabets. The minimal information we can get from the presented material is roughly the same as above stated: the first vowel is unrounded and i-like

The representation of the second vowel differs considerably from that of the first vowel. With the exception of the Old Turkic, it is always written with a full vowel sign. The vowel is generally <i> in scripts <i> vs. <e> differentiation.[45] Intuitively, one would say: the second vowel was more i-like and more sonorous

---

Edition: P. Zieme, An unknown story how the people of Andhra were rescued from their sins in a Buddha-less time. Notes on two Old Uigur leaves in Sogdian script. (fc.)

39 Mainz A 3 (korr. 4) s. prec. note.

40 TT VIII D (=Maue Kat I 18) 18 tr. of Skt. rājasuta- 'king's son, prince', TT VIII G (= Maue Kat I 6) 56 and Maue KatI I 91.4.2 ba (tg$_1$i[ ]), both tr. of Skt. kumāra- 'prince, heir-apparent'.

41 Maue KatI 80 b9, interpretation by G. Clauson TM 99. Possibly the same reference is meant by A. von Gabain AtüGr 371a: "tegin (tib. Schr.)".

42 The same substitution could have taken place elsewhere, e. g. in Sogdian, for the same or similar reasons. Hovdhaugen's suggestion does not imply that the (unambiguous) non-Turkic forms with initial t- trace back to Turkic sources.

43 In Bactrian <o> is multivalent standing for w, ŭ, o, ə, Ø, cf. N. Sims-Williams in CLI 233, here most probably for ə.

44 Hovdhaugen (s. note XXXX) holds that "the alternation a – i in the first syllable in Khotanese may be an attempt to render [closed, DM] Turkish /ė/." A similar argument was used in Ölmez's rule for the Old Turkic in runic script: "the word which can be written with an i or without a vowel ... should be consistently read with closed ė." (M. Ölmez ,What should a new edition of the Old Turkic inscriptions look like? In: 歐亞學刊/International Journal of Eurasian Studies 2, 2015, 80-93. Here p. 85 seq.) It is most doubtful whether Ti/egin falls under Ölmez's rule. And if there was no closed e in Tegin, Hovdhaugen's suggestion is empty.

45 The -ai-in the 2nd Khotanese sample represents ɨ, cf. Hovdhaugen 1971: 177.

than the first vowel. Translated into phonetic terms, this could mean: the first vowel <ä> was a sort of shwa indogermanicum [ə], the second, <ä$_1$>, was a kind of the [central ɨ]. Just in this direction points the Chinese transcript 特勤 tè qín; the reconstructed EMC dək gɨn has the shwa in the first and the central ɨ in the second syllable. But 提勲 tí qín, EMC dɛj gɨn from the Turfan inscriptions does not support shwa in the 1st syllable.

With certainty both diacritics under discussion represent i-like vowels. So we assign to < ä> the transcription symbol i and to <ä$_1$> the symbol i$_1$. It is open whether they are two phonemes or allophones of one single phoneme.[46] At any rate, they fill the diagnosed gap in the vowel system of KhT. We assume that they are neutral with respect to backness and frontness,[47] though they are not attested in words which are definitely back vocalic.[48]

Finally, we can guess why the graphemes <i, ī> of the basis alphabet were dropped: they stand for the front vowel i, while, obviously, in KhT signs for the central vowels were required.[49]

2.3.2 The inititial vowels



Fig. 16: KhT 6 ru, or: u

The only manifest instance of an initial vowel is 2 <a ña k̄a-y>. The number increases if the KhT language – as a (Para)Mongolic language – had no initial r. In the this case, the ambivalent sign < ru, u> (Fig. 16) has to be read as u in word initial position[50]: 2 <u k̄a-χš, u k̄a ju>, 7< u k̄a *b*a-r>, 6 <u-*c* (Or: -*n*)>, 9 <u-c>. The clear cases with initial a and u are back vocalic, which is hypothetically generalized. In consequence, no front vocalic words with initial ä-, ü- or i- would be attested. However, we encounter just these vowels preceded by h- throughout: 6 ha-r (Or: h*ä*-r?)

---

46 pügt**i$_1$**g beside 8 *pügt**i***gči would speak for allophones.

47 Cf. Clauson 1962: 194; 204.

48 I. e. words with the back velars k and γ so far. More indicators of backness or frontness are proposed in 2.3.2. If convincingly, 10 ubi$_1$ǰ would be demonstrably backvocalic.

49 In Tocharian /ɨ/ ⇒ <ä> and /i/ ⇐ <i, ī> were different phonemes, cf G.-J. Pinault (Chrestomathie tokharienne. Textes et grammaire. Leuven/Paris 2008.) p. 415.

50 Or - if necessary - as apocopated *uru. An evident case for that did not appear though.

gä-n, 6 hä1-g bä1-j, 3 hu-g *bu* [?]. That no initial h- occurs in back vocalic words might, of course, be owed to the meagre material. But it would make very good sense if h- represented a glottal before front vowels while the glottal preceding back vowels is unmarked (in A/añakay, ukaǰu etc.). This assumption is well compatible with the single internal occurrence of h , 9 <bä tä$_1$ *h*ä$_1$-*ñ*>, there probably in the function of a hiatus bridger between two front vowels.

3 The sound inventory

Though not all problems concerning signs and sounds were solved, some not even touched, a provisional sound inventory can be presented.

3.1 The vowels

The KhT vowel system comprises the back/front vowels a/ä, o/ö, u/ü and i in two variants, i and i_1, both likely central vowels and neutral concerning the opposition back vs. front.[51] One instance of marked vowel length, 5 türūg, does not prove the relevance of long vowels. The vowel inventory of KhT is in accord with the "pre-thirteenth century Mongolian" or the "Proto-Mongolian".[52]

3.2 The consonants

		Occlusive		Fricative		Semi-vowel	Nasal	Li–quid	Vibrant
		unvoiced	voiced	un-voiced	voiced				
Labial		p- [1]	b- , -b- [2]			-w-[3]	– [4]		
Dental		t-, -t- [5]	d-, -d-, -d [6]	-s- [7]			n-, -n-, -n [8]	-l- [9]	-r-, -r [10]
Palatal		č-, -č- ,-č[11]	ǰ- , -ǰ-, -ǰ [12]	š(i)- [13]		-j-, -j [14]	ñ-, -ñ-, -ñ [15]		
Velar	front	–[4bis]	-g-, -g [16]	x-, -x- , -x [17]					
	back	k-, -k-[18]	-g- [19]				–[4ter]		
Glottal		h-, -h-[20]							

Signs and Symbols: n-, -n-, -n = n as word initial, middle, final resp.

Fig. 17: KhT consonants

Evidence

1 4 pügti$_1$g, 4 puro*r*,6 pạlxi$_1$r, 8 *pügti*gči, 10 pạdạ, 11 p*ag*.

2 b-: 1; 8 *bodisatva, 1 b*i*ti$_1$, 2 bụdạ, 3 bạyi, 3 bigiy, 3 bọdi, 4 b[i]ti$_1$, 6 bargo[], 9 biti$_1$*h*i$_1$ñ, 11 b*i*tig. i$_1$
-b-: 3 hüg*b*ü[?], 6 hi$_1$gbi$_1$ǰ. 7 u*k*abar, 7 ǰụlạbạ, 10 ụbi$_1$ǰ, 10 ǰạ*lọ*bạǰ, 10 ǰạyạbi[?] (or: ǰạybi[?]).

3 <v>: 1; 8 *bodisatva.
<v$_1$>: 4 ;7;9 tụv$_1$ạ.

4 Not attested.

5 t-: 1;8;9 tọrọx, 4;7;9 tụwạ, 5; 10 türūg, 6 *ṭạyạ*ǰụ, 7 tügǰü, 7 tügn*y*ä, 7 tụnụ, 8 *to*[+]*gun* (?), 8 *tu*[ ].
-t-: 1 biti$_1$, 1;8 * bodisatva, 3 [...]i$_1$ti$_1$n, 4 *b*[i]ti$_1$, 4 pügti$_1$g, 5 k*at*o, 8 *pügti*gči, 9 biti$_1$*h*i$_1$ñ, 11 b*i*tig.

---

51 S. above.

52 J. Janhunen (in J. Jahunen (ed.), The Mongolic languages. London) 2003: 4 seq. Clauson (1962: 204) adds "closed é" following the authority of the 'phags-pa orthography, s. above, conceding, however, that "there is no other evidence on the subject" (o.c. 205).

6 d-: 1 *digi$_1$n, 3 dọlụǰạǰụ (or: dọlụ ǰạǰụ?),4 drọ (= dọrọ?), 5 dügi$_1$d, 10 *darka*d*, 11 *darkan.
-d-: 1; 8 *bodisatva, 2 *b*ụdạ, 3 bọdi, 10 pạdạ.
-d: 5 dügi$_1$d, 10 *darka*d*.

7 -s-: 1; 8 *bodisatva, 11 ...s*ä*g.

8 n-: 5 *ni$_1$ri, 10 *nạri* (l.: *ni$_1$ri),
-n-: 2 *kaganu, 9 *kaganun, 4 *ǰi$_1$lọnạr, 4 kranyag*u*ñ (= kara°), 4 ñạlạn[...] (or: ñạlạn), 7 xạni, 7 tügn*y*ä, 7 tụnụ, 9 *kaganun,
-n: 1;2;5;5;5;7;8;9; 10;10 *kagan, 1 *digi$_1$n, 1;9 ši$_1$ñin, 3 [...]i$_1$ti$_1$n, 4 ñạlạn (or: ñạlạn[...]),6 härgin, 8 ...*gun* (?), 9 *kaganun, 10 ... *i$_1$ǰi*n, 11 ...*rụn*, 11 *darkan.

9 -l-: 3 dọlụǰạǰụ (or: dọlụ ǰạǰụ?), 4 *ǰi$_1$lọnạr, 4 ñạlạn, 6 pạlxi$_1$r.

10 -r-: 1;8;9 tọrọx, 2; 6 xi$_1$ri, 4 pụrọ*r*, 4 čiči$_1$rạ, 5 *ni$_1$ri, 5 türūg, 4 drọ (= dọrọ?), 6 härgin, 6 bargo[],4 kranyag*u*ñ (= kara°?), 10 *nạri* (l.: *ni$_1$ri), 10 türüg (s. 5 türūg), 10 darka*d*, 11 darkan,
-r: 1;3;5;7 -ñạr (or: ñạr), 4 *ǰi$_1$lọnạr, 4 puro*r*, 6 pạlxi$_1$r, 7 ukabar.
To underline, the suggested absence of initial r- is based on the (-)Mongolian hypothesis. Words with initial <u> could be read with <ru> as well.

11 č-: 4 čiči$_1$rạ.
-č-: 4 čiči$_1$rạ, 6 ...xạči.
-č : 6 u*č* (Or: u*n*?), 9 uč.

12 ǰ-: 3 ǰạǰụ (or part of dọlụǰạǰụ), 10 ǰạlọbạǰ, 3 ǰạx, 10 ǰạyạbi[?] (or: ǰạybi[?]), 4 *ǰi$_1$lọnạr, 7 ǰụlạbạ.
-ǰ-: 2 ukaǰu, 3 dọlụǰạǰụ (or: dọlụ ǰạǰụ?), 6 *ṭạyạǰụ, 7 tügǰü, 10 ... *i$_1$ǰi*n,
-ǰ: 6 hi$_1$gbi$_1$ǰ, 10 10 ụbi$_1$ǰ, 10 ǰạ*lọ*bạǰ.

13 -š-: 1;9 ši$_1$ñin.

14 The Brāhmī sign is <y> . As it is usual practice in Turkological and Mongolistic literature, the semivowel is transcribed by y.
-j-: 3 bạyi, 4 kranyag*u*ñ (= kara°), 6 *ṭạyạǰụ, 7 tügn*y*ä, 9 ..*iẏụ*, 10 ǰạyạbi[?] (or: ǰạybi[?]).
-j: 2 añak̄ay, 3 bigiy.

15 ñ-: 4 ñạlạn, 1;3;5;7 ñạr (or the plural affix though separated from its antecedent by a daṇḍa).
-ñ-: 1;3;5;7 -ñạr as plural affix (otherwise an independent word), 1 *š*i$_1$ñin, 2 añak̄ay.
-ñ: 4 kranyag*u*ñ (= kara°), 9 biti$_1$*h*i$_1$ñ.

16 Written through <g>.
-g-: 1 *digi$_1$n, 3 bigiy, 3 hüg*b*ü[?],4 pügti$_1$g, 8 *pügti*gči, 5 dügi$_1$d, 6 härgin, 6 higbiǰ, 7 tügǰü, 7 tügn*y*ä.
-g: 4 pügti$_1$g, 5 türūg, 10 türüg, 11 b*i*tig, ...säg.

17 x-: 2;6 xi$_1$ri, 7 xạni.
-x-: 6 ...xạči, 6 pạlxi$_1$r, 8 ...xạ.. .
-x: 1;8;9 tọrọx, 2 ukax, 3 ǰạx, 8 ...*gu*x.

18 Written through <k̄>.
k-: 1;2;5;5;5;7;8;9; 10;10 *kagan, 2 *kaganu, 9 *kaganun, 5 k*ato*, 4 kranyag*u*ñ (= kara°).
-k-: 2 ukax, 2 ukaǰu, 2 añak̄ay, 7 ukabar, 10 *darka*d*, 11 darkan.

19 Written through <g$_1$>.

-g-: 1;2;5;5;5;7;8;9; 10;10 kagan, 2 kaganu, 4 kranyag*u*ñ (= kara°), 6 bargo[], 9 kaganun, 8 ...*gun* (?), 8 ...*gu*x.
-g: 11 p*a*g.
h-: 3 hüg*b*ü[?], 6 härgin, 6 hi$_1$gbi$_1$ǰ.
-h-: once intervocalically 9 biti$_1$*h*i$_1$ñ.

Comments

The comparison of the KhT consonant system with the consonant table that Clauson has made for the "pre-thirteenth century Mongolian"[53] leads to the following result, Fig. 18:

		Occlusive		Fricative		Semi-vowel	Nasal	Li–quid	Vibrant
		un-voiced	voiced	un-voiced	voiced				
Labial		p [1]	b			w [2]	(m)		
Dental		t [2a]	d	s , si > ši			n	l	r
Palatal		č	ǰ			j	ñ [2bis]		
Velar	front	(k)	g	x [3]			(ŋ) [4]		
	back	k	?g [3bis]		γ				
Glottal		h ≠ " aspirate h [5] ; intervocalic hiatus ’ "							

Signs and symbols: p = not or not with the same value in Clauson; (m) = accidentally not attested in KhT

Fig. 18: Mongolian consonant system (after Clauson) in comparison with KhT [54]

Notes on Fig. 18

1 *p as such has not survived in Mongolian proper, but Para-Mongolic shows "[p]reservation of the strong labial stop *p as a separate phoneme".[55]

2 The value and/or status of these sounds is not certain. – As to the palatal nasal ñ, the interpretation as palatalized phonetic variant of n is barred by the back vocalism of 2 A/añaā̆kay and 4 kranyag*u*ñ. According to Janhunen, however, the Para-Mongolic had *ny, e. g. in *nyoka 'dog' "as opposed to Proto-Mongolic *noka.i 'dog'."[56] There are no cases where KhT ñ directly corresponds with Para-Mongolic *ny.

---

53 Clauson 1962: 207. Janhunen's consonant tables of Proto-Mongolic (Proto-Mongolic. In: J. Janhunen (ed.), The Mongolic languages. London, New York 2003. pp. 1-29. Here: 6) and Para-Mongolic (Para-Mongolic. In: J. Janhunen (ed.), The Mongolic languages. London, New York 2003. pp. 391-402. Here: 397) are taken into account.

54 Ceteris paribus, the "Proto-Mongolic" presented by J. Janhunen (2003: 6) does not comprise glottals and his velar subsystem counts less elements (*k, *g, *x, *ng): back and front velars are not distinguished; however, a phoneme *x is postulated which is not contained in Clauson's table. It is successor of **p and as such not comparable with the x of KhT where p is still preserved unchanged. For another not unproblematic "primary velar spirant *x" in Para Mongolic s. Janhunen 2003: 397.

55 Junhanen 2003: 396.

56 Janhunen 2003: 397.

2a Unlike in Mongolian, t is also preserved before i, i_1. This feature is shared with Para-Mongolic, being "open the possibility that it reflects the original Pre-Proto-Mongolic, rather than the Proto-Mongolic lineage."[57]

3 The subsystem of velars in KhT seems to comprise unvoiced front k (unattested), front voiced g (medially possibly lost between the same vowels, biti$_1$*h*i$_1$ñ?) , back unvoiced k (occlusive, in word-final position perhaps > fricative x, ukax), voiced back g (realised as fricative, at least intervocally, kagan) and a neutral fricative x (occasionally allophone of the back k). Clauson's picture of the "pre-thirteenth century Mongolian" is different in two points "that there were ... [1] a voiced fricative *ǧ* and possibly a voiced plosive *ǧ*, [2] but not an unvoiced fricative *x*." [58]

4 Occurring only before a velar, no such case in KhT.

5 The KhT h- was tentatively determined as on-glide of front vowels and intervocally as hiatusbridger. At any rate, there was no such restriction in "thirteenth century Mongolian ... [T]here is reasonable doubt whether this was originally a Mongolian sound... [T]he initial *h*- is more likely to be a secondary phenomenon than an original feature of the language." (o.c. 208)

*

Lastly, a few remarks on the syllable structure, consonant clusters, word initials and finals.

The predominant syllable structure is (C)V(C), the consonants[59] in brackets being facultative. If this structure was universally valid that would imply:

(1) #(C)V(C)...#: Only vowels or single consonants can appear as word initials. Here we have two exceptions, both with the liquid r as second element: <dro> and < kra nya gu-ñ > which are probably real or graphematic syncopes of doro/dörö and kara°.

(2) #...(C)V(C): Only vowels or single consonants can appear as word finals. No exceptions. For a comparison with " pre-thirteenth century Mongolic" s. Fig. 19.

57 Janhunen 2003: 397.

58 Clauson 1962: 206.

59 The suggested glottals before vowel (2.3.2) are not counted as separate sounds. Otherwise the structure formula would be CV(C).

Pre-13th cent. Mongol.	-b	-c*	-d	-ğ	-g	-l	-m	-n	-ŋ	-r	-s	-y	– **	– ***	– ***
KhT language		-ǰ	-d	-g	-g			-n		-r		y	-č	-ñ	-x

* restricted to non-native words; not so in KhT.
** expressly excluded by Clauson.
*** not existent in the pre-thirteenth century Mongolic.

Fig. 19: Word final consonants in the pre-thirteenth century Mongolic and in KhT

Notes on Fig. 19
"It is in any event clear that a syllable could not end in CC, so that ... [no] final CC ... could occur in a native word." (Clauson 1962: 207.)
As to the word final consonants Clauson (o. c. 209) writes: "There is no theoretical reason why ... *-b*, ***-c*** [=ǰ], ***-d***, ***-ğ*** [= back g], ***-g*** [= front g], *-l*, *-m*, ***-n***, *-ŋ*, ***-r***, *-s* and ***-y*** [=j] should not occur at the end of a word, but I have not found any native word ... ending on *c* [=ǰ]." [60] The bold-faced finals are also met with in KhT, which has in addition: (1) -č , which is expressly excluded from the end of a Mongolian word of the 13th cent. and (2) the two consonants -x and -ñ which do not belong to the inventory of the "pre-thirteenth century" Mongolian consonants. Furthermore, a special restriction of final -ǰ to non-native words seems not to exist. The three occurrences, 6 hi_1g-$bi_1ǰ$, 10 10 ụ-$bi_1ǰ$, 10 ǰạ*ọ*-bạǰ (error for - $bi_1ǰ$?, are possibly morphologically related. Lastly, KhT admits of unvoiced final consonants, x and č, unlike the Mongolian; the majority of word finals, however, are not unvoiced.

*

(3) #...(C)V(C)+(C)V(C) ...# : No consonant clusters with more than 2 consonants are admissible in the middle of a word. Actually, there is one single triple consonance, 7 tügnyä, which may, however, stand for trisyllabic tüg-ni-yä. For a comparison with " pre-thirteenth century Mongolic" s. Fig. 20.

1↓ 2→	b	ǰ	č	d	t	g_b	k_b	g_f	k_f	y	l	m	s	x
b			+	+	+	+	+	+	+				+	
d							+							
t							+							
g_b			+	+	+								+	
g_f	+	(+)	+		+									
l	+	+	+	+	+	+	+	+						+
m		+		+		+	+		+				+	
n		+		+						+				
ŋ						+	+	+	+				+	

60 In Classical Mongolian, the affricate ǰ "n'est jamais employé à la fin des mots", L. Hambis, Grammaire de la langue mongole écrite. I. Paris 1946, p. XVI. Correspondingly Clauson TM 209: "I have not found any native word word, basic or elongated, ending in c."

r	+		+	+		+	+	+	+			+		
s						+		+						
y											+	+		

Symbols: + = attested in Mongolian; + = attested in KhT; + = attested in both ; ( ) = "the second C is the first sound of a suffix"(Clauson 1962: 207): not included by Clauson; □$_{b/f}$ = back or front velar.

Fig. 20: Consonant clusters in Mongolian 13th and 14th cent. texts and in KhT

Notes on Fig. 20
The KhT clusters -gb-, -gǰ- (in contrast to -gč-), -gt- (in contrast to -γt-) and , not un–expectedly, the foreign -tv- (1;8 bodisatva) are absent from Clauson's list which "is perhaps not quite complete."[61] At any rate, only about half of the KhT instances have counterparts in Clauson's table.
However, an important general observation by Clauson (1962: 207) is also valid for KhT: no "unvoiced first component plosive of a CC" occurs."

*

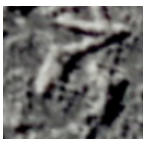
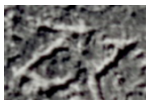
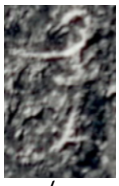
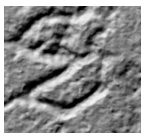
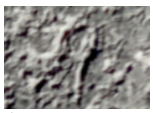
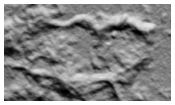
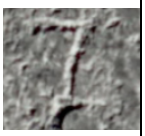
Summing up, the vocalism and consonantism of the KhT are in broad lines, I underline: in broad lines, compatible with those of (Proto-) Mongolian and Para-Mongolic. Future research must show whether and how existing divergences can be explained.

61 Clauson 1962 (= G. ClausonTurkish and Mongolian studies. London 1962. (Prize Publication Fund; XX.)): 207.

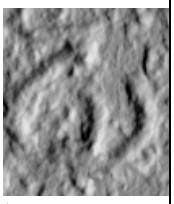
4 Appendices

Appendix I: Signs and sounds

(A) Consonants

	Occlusive		Fricative	Semivowel	Nasal	Liquid	Vibrant
	Voiceless	Voiced	Voiceless				
La-bial	 pu	 ba			-----		
Den-tal	 ta	 da	 sa	 va  v_1a / v_1a-r	 na		
Pala-tal	 ca	 ja	 śa	 ya	 ña	 la	 ra
Front velar		 ga	 <χš> = x				
Back velar	 k̄a	 g_1a					
Glot-tal							

(B) Vowels

	a	<ä>, i	<ä1>, i1	u	ū	o
Inde-pen-dent	 a			 u		
De-pen-dent	(inherent in all conso-nant signs without dia-critic)	 dä	 cä1	 pu	 rū	 bo
				 tu		

Appendix II: Transliteration

Sigla of the sources:

D	photos taken by Diyarbekirli
Dd	D detail (showing the middle part of the inscription)
Dt	D total
Mo	photos provided by Moriaysu
Mor	rubbing provided by Moriyasu
Mz	used when Mzv and Mzn are in accord
Mzv	3D photos taken and processed by FH Mainz, read with rtiViewer
Mzn	3D photos taken and processed by FH Mainz, read with the new viewer
V	photos taken by E. de la Vaissière
Vp	processed photos taken by E. de la Vaissière
ŽS	used when ŽS1 and ŽS2 are in accord
ŽS1	retouched rubbing/photo from *Atlas Orhonskih pamjatnikov.* 2006
ŽS2	variant of ŽS1

N.b.: A varia lectio will be given in the apparatus if and only if the evidence of a source is **not compatible** with the spelling which is recommended by the majority of sources and accepted in the text.

Other signs and symbols

kā	uncertain reading
<a>	(1) a by emendation (2) a as transliteration symbol
[ ]; [a]	loss; a by restoration
[?]	uncertain loss
(o)	normalising addition
|	interpunction marks (without regard to the actual form)
-r	unvocalised r, usually attached to the precedent sign by a small stroke, the so-called virāma stroke
+	equivalent of one syllable
□[?]	something (□) questionable
×	equivalent of a part of a syllable

N.b.: The transliteration symbols follow the accepted transliteration/transcription of the Indian signs; divergent symbols are explained above in §2 "The sign inventory".

(01) b*ä* tä$_1$ | *ñ*a-r | k̄a g$_1$a-n | dä g*ä*$_1$[62]-n | *ś*ä$_1$[63] ñä-n | bo dä | sa-t va | to ro-χš

(02) k̄a g$_1$a-n | *bu* da | k̄a g$_1$a nu[64] | u k̄a-χš | u k̄a ju | χšä$_1$ rä | a ña k̄a-y

(03) + [65] ×ä$_1$ [66] *t*ä$_1$-n | ja-χš bo dä | bä gä-y | ña-r / ba *yä* | do lu ja ju | hu-g *bu* [67] [?]

(04) +[?] *b*[] tä$_1$ | *j*ä$_1$ *l*o *na*-r | k̄ra nya g$_1$*u*-ñ | tu v$_1$a[68] | pu ro-*r* | cä cä$_1$ ra | pu-g tä$_1$<->g[69] | ña la-*n*

(05) × | *k̄*[][70] *g$_1$*a + + k̄*a to* [71] | *ña*[-[?]]r [72] | du gä$_1$-d | nä$_1$ rä | k̄a g$_1$a-n | tu rū-g [73] | k̄a g$_1$a-n

(06) u-*c* [74] | dro | *ta ya* ju [75] | χšä$_1$ rä | ha[76]-r gä-n | ba-r g$_1$o-×[77] | pa<->l[78] *χš*ä$_1$[79] -r | + χša [80] cä | hä$_1$-g bä$_1$-j

---

62 Or: g*ä*? Though the distinctive loop of the diacritic is destroyed the visible part seems to belong rather to <ä$_1$> than to <ä>.

63 <ś> is clear enough here, but better discernible in col. 8.

64 Unusual form.

65 Complex sign, the lower part seems to be (-)h; however, syllable closing h would be strange.

66 Or:×*o*(?);×ä$_1$-*l* (not excluded).

67 <pu> corrected into <bu>?

68 Or: v$_1$a-r.

69 The virāma stroke is not discernible, but cf. *pu-g tä*-g cä in col. 8.

70 Or: g$_1$[]?

71 Or: do?

72 Or: *ña* r[].

73 Spelling with short u in col. 10.

74 Or: -n? Faint or even lacking virāma stroke, but clearly visible in col. 9.

75 S. KT details. If read correctly, the shape of ya is less rounded then in col. 2 and 3.

76 Or: h*ä*? There may be traces of the -ä diacritic.

77 Perhaps g$_1$o-*l*; or, much less probable, instead of g$_1$o-×: g$_1$*r*o.

(07) tu-g *j*u | u *k̄*a *b*a[81] -r | *ña*-r [82] k̄a *g₁*a-n | χša nä | ju la ba | tu nu | tu-g n*ya* [83] | *tu* v₁a

(08) +[84] χša-× *tu*-× *to* × *g₁u-n* [85] | *pu-g tä*-g cä | śä₁ ñä -n | *bo* dä sa-t va | to ro-χš | k̄a g₁a-n

(09) [+] *l*[][86]+ | ×*ä*[87] *yu*[88] | u-c bä tä₁ *h*ä₁-*ñ* | + + *g₁u*-χš | *tu* v₁a[89] | *ña*-r [90] | k̄a g₁a-n | *to* ro-χš | k̄a g₁a nu-n |

(10) +[91] pa[92] da | *na* r*ä* | k̄a g₁a-*n* | tu-[93]ru[94]-g | + g₁a-[] ×*ā̄*[95] *jä*[-]n |[96] u bä₁-j | ja *lo* [97] ba-j | da-r k̄a-*d* | ja ya [98] bä

(11) [?[99]] *ru-n*[100] b*ä* [101] tä [102]-[103]g [104] | + s*a* [105]-g | p*a-g₁* [106] [+ +] j[] [?][107] da-r k̄a-n ba[108] | tä[109] *b*a k̄a [110]

---

78 Or: pla?
79 Or: × ä₁.
80 Or: -χš.
81 Rather than *g₁*a.
82 Less probable: *v₁a.*
83 Or: na, without subscript -y?
84 Space for a complex sign; no intellegible traces.
85 This is what one would guess from ŽS. Though the three curved lines of g₁ are uncertain as well as -u and -n has an inappropriate stroke at the lower end, no better proposal can be made.
86 Or: *lo*?
87 The sign might be a variant of ka (Sander alphabet u); but only the reading nu makes sense and the elsewhere unattested <ka> is excluded from a back vocalic word.
88 The trace above is unclear, perhaps a daṇḍa.
89 Or: v₁a-r.
90 Or: *p*a-r, *ba*-r?
91 Perhaps: hu or h[]-r?
92 Or: *b*a?
93 The virāma stroke is erroneous.
94 Spelled with -ū in col. 5.
95 ×*ā̄* : *g*ā̄ ŽS.
96 *jä*[-]n | : j*ä-x* appears from D instead.
97 The form slightly differs from <lo> in col. 4 and is therefore marked as uncertain.
98 Or: ja[-]y?
99 Probably no loss of script. The upper left rim of the stone is seemingly quite well preserved (mostly smooth-edged, minimal sharp-edged fractures). The stonecutter followed the natural form of the stone which provided not enough space for writing something above ru-n. Probably no lacuna between col. 10 and 11.
100 Or: -*c.* If word initial: *u-n, u-c.*
101 Or: ba. - Closed form of <b>.
102 -ä₁ is not excluded.
103 Virāma stroke is uncertain.
104 <g>, still clearly readable on D, is now partly destroyed.
105 A vowel diacritic, possibly -ä, can not be excluded.
106 Or: pu ×.
107 Probably no loss of script between the two visible akṣaras; there was not enough space for writing.
108 Or: b[ä].

Appendix III: Transcription

Preliminary notes

The transcription is made and to be understood before the background of the (para-)Mongolian hypothesis which includes vowel harmony and absence of word initial r.

Velars and glottals are indicators of backness and frontness:
k ⇐ <k̄>, g⇐<g$_1$> are certainly, (initial) u and a are probably signals for backness;
g=<g> and probably h= <h> signal frontness; x⇐ < χš > is neutral and probably pronounced [ç] or [x] according to the milieu.

The vowel signs represent either front or back vowels; i and i$_1$ are included for practical reasons though they are probably neutral in respect of frontness vs. backness. Their special character is either determined by the described indicators or undecided. Accordingly they are transcribed
ä,ö,ü,i,i$_1$ in words with front indicators;
a,o,u,ı,ı$_1$ in words with back indicators;
a̤,o̤,ṳ,ı̤, ı$_1$̤in words without indicators.

bodı-satva : the hyphen is applied between parts which belong together, but are separated by interpunction mark

(01) bṭtı$_1$-*ñ*a̤r kagan digi*ı$_1$*n[111] *šı$_1$*ñın bodı-satva to̤ro̤x

(02) kagan *bṳ*da̤ kaganu ukax ukaǰu xı$_1$rı̤ añakay

(03) ...*ı$_1$* *t*ı$_1$n ǰa̤x bo̤dı̤ bigiy-ñär ba̤*y*ı̤ do̤lṳǰa̤ǰṳ[112] hüg*bü*[?]

(04) +[?] *b*[ı]tı$_1$ *ǰ*ı$_1$*l*o̤na̤r k(a)ranyag*u*ñ tṳwa̤[113] pṳro̤*r* čı̣čı$_1$ra̤ pügti$_1$g ña̤la̤*n*

(05) × *k*[a]*g*a[nu?] + k*at*o-*ñar* dügi$_1$d nı$_1$rı̤ kagan türū̃g kagan

(06) u*č*[114] d(o̤)ro̤ *ta̤ya̤*ǰṳ xı$_1$rı̤ härgin[115] bargo[l] pa̤lxı$_1$r [+]xa̤čı̣ hi$_1$gbi$_1$ǰ

(07) tüg*ǰ*ü u*ka**b**a*r-*ña*r ka*g*an xa̤nı̤ ǰṳla̤ba̤ tṳnṳ tügn*yä* *tṳ*wa̤[116]

(08) + xa̤[] *tṳ*[] *to*[go[?]]*gun* *pügti*gči śı$_1$ñın *bo*dısatva to̤ro̤x kagan

(09) [+]*l*[] + []*ıyṳ* uč biti$_1$*h*i$_1$*ñ* + + *g*ux *tṳ*wa̤-*ña̤*r kagan *to̤*ro̤x kaganun

(10) + pa̤da̤ *n*<ı$_1$>r*ı̤* kaga*n* türüg [ka]ga[n] []*ı$_1$ǰı̤*n ubı$_1$ǰ ǰa̤*lo̤*ba̤ǰ darka*d* ǰa̤yabı̤[117]

(11) [?]*rṳn*[118] b*i*tig + s*ä*g p*ag* [+ +] j[] [?] darkan ba[119] | tı̤[120] *ba̤* ka

[blanco]

---

109 Or . dä?
110 The rest of the column is blank.
111 Or: dig*i*n.
112 Or: do̤lṳ ǰa̤ǰṳ.
113 Or: tṳwa̤r.
114 Or: ṳ*n*?
115 Or: h*i*?
116 Or: tṳwa̤r.
117 Or: ǰa̤ybı̤?
118 Or: *ruč*? If word initial: *un*, *uč*.
119 Or: b[i].
120 Or . dı̤?