AN INTERPRETATION OF THE KHÜIS TOLGOI INSCRIPTION[1]

ALEXANDER VOVIN
CRLAO (EHESS), PARIS

I humbly dedicate this article
to the Great Old Man of Central Asian Studies,
Prof. Dr. András Róna-Tas

I am going to present below a tentative reading and interpretation of the Khüis Tolgoi inscription (hereafter KhT inscription). This work would not be possible without the pioneering work by Dieter Maue on the decipherment of its writing system, which was also greatly added by the efforts of the team of specialists in 3-D photography that accompanied us on our 2014 expedition to Mongolia.

Although at this stage it is not possible to provide a complete reading and interpretation of the KhT inscription, slightly going ahead it seems beyond reasonable doubt that it is in some form of Mongolic. It also appears that the language of the inscription, although it can be conditionally termed as a variety of Para-Mongolic, is much closer to the mainstream Mongolic languages, such as Middle Mongolian and modern extant Mongolic languages than to Serbi-Khitan,[2] although there are some features that the language of the KhT inscription shares with the latter. It is necessary to note that the KhT inscription is monolingual, which certainly opens the gates for the guess-work and speculations, some of which certainly will be corrected by the following generations of scholars.

The KhT inscription consists of two stones, one preserved in the basement of the National Archeological Institute, and the other placed near the entrance to the building. Certain circumstances prevented us from reading and photographing the text on the second stone.

LINE 1-1[3]

Open transcription (Maue):

bi̯tu̯$_1$-ña̯r qaɣan di*gi*$_1$n *š*$_1$ñi̯n bodı-satva to̯ro̯ks

Tentative transcription (Vovin):[4]

biti-ńer qaɣan digi-n šińi-n bodi-satva törö-ks[e]

Morphemic analysis:

Inscription-PLUR qaɣan die-CM new-GEN
Bodhi-sattva be.born-PAST.PAR

[1] This article has benefited from comments by many colleagues around the world. Besides my partners in crime, and co-authors in our Khüis Tolgoi enterprise: Dieter Maue, Mehmet Ölmez, and Etienne de la Vaissière, I thank especially Marcel Erdal, Akos Apatoczky, Pavel Rykin, Bayarma Khabtagaeva, Christopher Atwood, Jens Wilkens, Werner Nahm, Tibor Porcio, Dominique Dumas, Juha Janhunen, and many others too numerous to mention here. We regret to say that some individuals thoroughly misunderstood the presentation of unpublished materials at the PIAC 2017 and the subsequent discussion forum on academia.edu as a licence to exploit them for their own purposes without authorization. But above all, I owe a debt of gratitude to András Róna-Tas, who was extremely generous with his time to offer me his suggestions in a series of e-mails over the October 2017. Many of his brilliant solutions that escaped me on the first attempt are incorporated here with the proper accreditation. And for those that I chose to be stubborn about and left them as they are, I can only say: "I am sorry, András!" Although this article grew out of the team work of four of us, there are some differences in interpretation that so far we have tried to bridge, but unsuccessfully. A good example of those is the phonetic interpretation of the special sign #3 that Dieter Maue tends to interpret nowadays as a back velar *g*, the interpretation I cannot accept for a number of phonological, morphological, and typological reasons, the main of them being the fact that I cannot come up with a sensible Mongolic reading if it were indeed *g*. Therefore, I preserved the earlier interpretation suggested by Maue, namely {kš}, that is understood here as /ks/. Both the internal linguistic data as well as cross-linguistic typology support, I believe, the latter interpretation much better.

[2] There was an unprecedented growth of Khitan studies, with four significant monographs published in the recent years: Kane 2009, Wu & Janhunen 2010, Shimunek 2017 and Monumental Chinggeltei, Wu, and Jiruhe 2018.

[3] The first number indicates the number of the stone, and the second the number of a line on it.

[4] Unlike the open transcription by Maue, my tentative transcription does not indicate uncertain readings (Maue's italics), and it does not differentiate between subscripted and non-subscripted versions of vowels, e.g. *i* and *i*$_1$, unless I have no interpretation for the words containing these subscripts.

Journal Asiatique 306.2 (2018): 303-313
doi: 10.2143/JA.306.2.3285621

Notes

1. I finally decided to follow my initial thought that *biti-ńer* is a plural form of *bitig* 'inscription' in spite of the suggestion by Róna-Tas that it is the name of a *qaɣan* (p.c., October 2017). Although it may look good syntactically, a *qaɣan* with such a name is not historically attested. However, I think now that my initial interpretation of *digin* as 'tegin' is incorrect, and I follow a suggestion by Róna-Tas (p.c., October 2017), it is preferable to compare it with Khitan *tege-* or *tige-* 'to die'[5] + converbum modale *-n*.

2. *qaɣan,* and *bodi-satva* seem to be for the most part unproblematic, but they also provide no information on the genetic affiliation of the language.

3. *šińi-n* 'new' is the word that is highly diagnostic, clearly pointing to the Mongolic direction. The form is to be read *šini*, cf. EMM *šini* 失你 'new' (MNT §265), although the majority of attestations indicate *šine*, cf. *šine* 𐼲𐼰𐼰𐼴 (Qub IV: 43), WMM *šine* شينه (Mu 334). Cf. Khitan *šɛn (Langjun 7: 6), thus phonetically KhT form is closer to mainstream Mongolic. The final *-n* is likely to be a genitive.

As far as I can tell, there are no clear-cut cases of the adnominal usage of genitive in MM. Thus, e.g.:

> Tümet irgen-*ü* oki-t
>
> Tümet people-*GEN* girl-PLUR (MNT §241)

can be interpreted as 'girls *who are* Tümet persons' or 'girls *of* Tümet people'. Meanwhile, in Khitan, as noted by Shimunek, a genitive case marker can be used as an adnominal marker, albeit this function is restricted to the cases when it is preceded by a numeral (2017: 260), e.g.

> 𘞎𘞀𘞀 𘞏𘞀𘞄
> tau-*un* u.ur
> five-*GEN* division
> the Five Divisions (lit.: divisions *that are* five) (Yelü Xiang-wen 7: 15-16)

4. Bodi-sattva is either a given name of the Turkic *qaɣan* from the First Khanate, or Buddhist Bodi-Sattva could also be meant here, as the text may be a historical, but with the religious overtones.

5. *tǫrǫ-ks* (in script) is likely to be Mongolic *törö-ks[e]* 'to be born', past participle of the verb *törö-* 'to be born'. Étienne de la Vaissière suggested that *tǫrǫ-g$_1$* recently proposed by Dieter Maue instead of the earlier reading *tǫrǫ-ks* may indicate the tribal name *Tiělè* 鐵勒 < EMC *t^{h}iet lək* < LHC *t^{h}et lək*. I find it difficult to accept this proposal for three reasons. First, the vowels in the Chinese transcription have no indication of the feature [+labial], which is quite clearly present in *tǫrǫ-g$_1$*. Second, if the mysterious sign #3 were really a back *g$_1$*, this would imply *torog$_1$* with back /o/ vowels, and not *törög* with front /ö/ vowels. Meanwhile, Chinese transcription clearly includes [-back] vowels. Third, the end of the sixth or the beginning of the seventh century is still too early for the Chinese to lose final *-t* in EMC *t^{h}iet* altogether, or even to undergo the shift of *-t* > *-r*, which did not happen until the High Tang times more than a century later. Thus, although *bodi-satva tǫrǫg$_1$ qaɣan* interpreted as 'Pusa, qaɣan of Tiělè' (lines 1-2) might be persuasive from the historical point of view, I regret to say that from the linguistic point of view the interpretation of *tǫrǫ-kš/g$_1$* as *Tiělè* 鐵勒 seems to be untenable.

6. There are two typically Mongolic morphological markers on this line: past participle *-ks* < *ksA, and genitive *-n* after vowel stems (MM *-yin, -īn, -n*)

LINE 1-2

Open transcription (Maue):

> qaɣan bu̯da qaɣanu uqaks
> uqaǰu ksɪ$_1$rɪ añaqay

Tentative transcription (Vovin):

> qaɣan buda qaɣan-u uqa-qs[a]
> uqa-ǰu ksɪrɪ Ańaqay

Morphemic analysis:

> qaɣan Buddha qaɣan-GEN realize-PAST.PART
> realize-CC country Ańaqay

Notes

1. Unless *Buda* is simply 'Buddha', which is the simplest solution, the only other possible explanation I can think about is Khitan *bud* 𘞃𘞄𘞅 'external clan' (Yelü Jue 31: 55), (Hsiao Hui-lian.Can 2: 1), (Yelü Pusuli 7: 18).[6] But this leaves final *-a* stranded: it is unlikely to be a MM dative-locative case suffix *-A*. There are no other fitting attestations either in MM or Pre-Classical WM texts to the best of my knowledge. *Buda qaɣan* is reminiscent of OT *bur-qan* 'lord Buddha'; or archaic OT *pū rkā kāṃ*, equivalent to the later *bur qaɣan*. (Maue 2008: 62-63.)

[5] Khitan examples: *tege'er* 𘞀𘞂𘞇/𘞅 (247.349.269/341) <t.ge.er> 'deceased, died' (masculine past tense in *-er*) (Xiao Dilu 11-8, 23-5, 35-8), *tege'en* 𘞀𘞂𘞈 (247.349.219) <t.ge.én> 'deceased, died, (feminine past tense in *-én*) (Yelü Xiang-wen 32-17). Examples provided by Róna-Tas.

[6] Shimunek glosses this word as 'external clans' (2017: 265), but the plural would probably work only if Khitan *-d -*𘞅 is here a plural suffix.

2. Maue suggested that *ʳuka-* (in script) may reflect either *uruqa* or *uqa*. The first one is improbable, as only EMM *uruqa* 'relatives, descendants' (MNT §11, §23, §24 etc.), a likely loanword from Turkic, comes to mind, and it neither fits the context, nor does agree with apparently verbal nature of the stem. EMM verb *uqa-* 'to realize, to know' (MNT §18, §78, etc.) fits much better. No Khitan cognates.

3. *ksɻrɻ* is probably a borrowing from Tumshuqese *χšera* 'country', as Maue has initially suggested. In this meaning it could fit here and also below on line 1-6.

4. At the first glance, Ańaqay might be a name of the last Ruan-ruan qaɣan, who refused to give his daughter as a wife to Turkic Bumin qaɣan in 546 AD. But this is too early for this inscription, and there is no title *qaɣan*, see line 1-3 note 1.

5. The most important is that this line is full of typical Mongolic morphology: MM idiosyncratic genitive *-u* after stems in *-n* as in *qaɣan-u* 'qaɣan-GEN', past participle *-ks[A]*, converbum contemporale *-ǰU*, and less likely dative-locative *-A*, if the alternative proposed etymology for *bud-a* is correct. Once again we can see that the language of the KhT inscription is much closer to mainstream Mongolic than to Khitan: a) there is no Khitan genitive *-u*, as the Khitan words with final *-n* take *-en* instead (Shimunek 2017: 253), b) the Khitan word *qa ~ qa.ɣa* 'qaɣan' takes the genitive in *-an: qa.ɣa-an* (Kane 2009: 132); c) Khitan has converbum contemporale *-ǰ* corresponding to MM *-ǰU*, and d) no past participle *-ks[A]* or e) dative-locative case suffix *-A*.

LINE 1-3

Open transcription (Maue):

…ɻɻtɻ₁n ǰạ̈ks bọ̈dɻ bigiy-ñär bạ̈yɻ
dọ̈lụ̈ǰạ̈ǰụ̈ hügbü[?]

Tentative transcription (Vovin):

[.?.]-ıte-n ǰa-qs[a] bod-ı beg-ey-ńar bayyı-Ø
dolu-ǰa-ǰu hügbü +?

Morphemic analysis:

[title-iten?] / Vite-CM promise-PAST.PART tribe-ACC
beg-?-PLUR be/stand-NP listen-COOP-CC ? ?

Notes

1. Ańakay is supposed to have a title, and since he cannot be equated with the Ruan-ruan qaɣan Anagui as demonstrated by Étienne de la Vaissière, it is natural to suppose that unreadable beginning of the line 1-3 corresponds to this title, but not to *qaɣan*.

2. I initially thought that *-te-* in the beginning of the line is *te-* 'to say' (quotational verb) was borrowed from OT *tė-* 'id.' into Khitan (Vovin 2013: 622-623). However, Róna-Tas must be correct that *-ıte-* is either an unknown verb ending in converbum modale *-n*, or *-ıten* represents a part of some title (p.c., October 2017).

3. *ǰa-* 'to promise' is probably a cognate to EMM and WMM *ǰa'a-* (HYYY 17b.2), (MNT §82, §84, etc.), (Mu 199) ~ EMM *ǰi'a-* (MNT §121, §145, etc.) 'to report, to promise'.

4. *bod* must be a loan from OT *bod* 'tribe', unless the direction of loan was opposite, cf. OT *bodun* 'population, subjects', apparently with a Mongolic plural *-n*. Accusative marking is strange here, although the similar phenomenon in MM and even modern Khalkha might be comparable.

5. *beg* might be yet another OT loan (OT *beg* 'bek', ultimately from EMC pæk 伯 'elder'), but the *-ey* segment is unaccounted for. The vocalism is also problematic, since *bigiy* in KhT seems to indicate [i], but there is at least one other case of *e* > *i* raising: *šine > *šińi* 'new' in line 1-1.

6. For *bayyı-Ø* 'be/stand-NP' cf. MM *bayyi-* 'to be, to stand, to live' (MNT §90, §108, §131, etc.) and MM nomen praesentis *-(U)yi* in its predicative converbial function (Godziński 1985: 139), which apparently undergoes contraction *bayyi-+-yi > bayyi*. This nomen praesentis *-(U)yi* can be related to Khitan converb *-i* found after various vowels (Kane 2009: 149-150).

7. *dolu* is not likely to be a cognate of MM *dolu'an* 'seven' (without the numerical suffix *-pan (> *-'an* in intervocalic position)), as initially suggested, but constitutes together with the following *jaju* the verbal form *doluǰa-ǰu* 'listen together and',[7] where *-ǰa-* is a cooperative.

8. *hügbü* is an interesting case. The language of the KhT inscription has both *p-* and *b-*, so unlike MM, *h-* cannot be a reflex of proto-Mongolic *p-. I suspect that it either denotes initial *h-* as suggested by Maue (but primary *h must be alien to the Mongolic sound system) or *ʔ-*. There are no apparent MM cognates. The last word on this line is unreadable.

9. As in the previous two lines, there is plenty of Mongolic morphology in line 1-3: accusative *-ı* (cf. MM accusative *-i*), converbum modale *-n*, nomen praesentis *-(U)yi* in its predicative function, plural suffix *-ńAr*, past participle *-qs[A]*, and converbum contemporale *-ǰu* and nomen praesentis *-(U)yi* in its predicative function. The last three has been already discussed above. As for the first two, note that there are no accusative *-ı* or *-i* and converbum modale *-n* in Khitan.

[7] This was also proposed by Róna-Tas (p.c., October 2017).

LINE 1-4

Open transcription (Maue):

+ʔ *b*[ɪ]t$_{ɪ_1}$ *ǰ*$_{ɪ_1}$*l*o̱na̱r q(a)ranya*γuñ* *tu̱*wa̱ pu̱ro̱*r*
čɪ̣čɪ$_{1}$ra̱ pügti$_{1}$g ña̱la̱*n*

Tentative transcription:

b[i]ti jilo-nar q[a]ra-n[V]ya-ɣuń tuwa pu̱ro̱-r
čeči-re pügtig ńele-n

Morphemic analysis:

inscription stone-PLUR look-?-NML
tribal.name sin(?)-NML stab-CF saved join-CM

Notes:

1. *jīl̤o* must be related to MM *čila'un* 'stone' (with possible voicing of *č-* in an intervocalic position), and *-nar* is a plural suffix. Note that this time we have *n-*, not *ń-*, possibly under influence of the final *-n* of the preceding word that itself elided (however, cf. also *qato-ńar* 'qaɣan's wives' < *qatun-ńar* in line 1-5).[8] Note that the usage of the plural confirms that there were more than one stone in the inscription.

2. Some remaining individual items on this line can be probably understood, but there is no coherent reading. *kra-nya* looks very 'un-Altaic', but on the basis of *dro* in the line 1-6 that is likely to stand for *doro* 'law', we can suspect the formation of secondary clusters due to the elision of the vowels in the first syllables here as well. I tentatively interpret this as *q[a]ra-* 'to look', cf. MM *qara-* 'to look' (MNT §5, §63, §77, §183), (Mu 292) + *-n[V]ya-*, obscure morphological marker without any obvious parallels in MM or Pre-Classical WM + *ɣuń*, a deverbal nominalizer cf. MM *-'Un* (Godziński 1985: 53) and WM *-ɣun ~ -gün* (Poppe 1964: 46), which forms nouns designating qualities or abstract nouns. An alternative, but much more inferior explanation: *q[a]ra n[o]ya* 'black (i.e., civilian, not military) officials'. Cf. MM *qara* 'black' (MNT §21, §25, etc.), (Mu 227), (IM 444) and *noyan* 'official, lord' (MNT §8, §51, etc.), (Mu 115), (IM 443). *ɣuń* is difficult to explain. It certainly cannot be MM *-wun* or WM *-ɣun*, a deverbal nominalizer, because we have no verb preceding it. The interpretation is really vexing here. Róna-Tas suggested *qara-ńag-un* Tuva 'black-?-GEN Tuva' combining with Tupa discussed below, as the one that may give some sense, though he admits that *-ńag-* is not clear (p.c., October 2017). Thus, since we would be exchanging one unknown to another unknown, I will keep my own interpretation for the time being. In addition, but very importantly, in the KhT inscription, *-ń-* ≠ *-ny-*.

3. *tu̱wa* occurs three times in the KhT inscription: on the lines 1-4, 1-7, and 1-9. It is not clear what it can be, and the problem is further complicated by the fact that *-w-* occurs only in this word. The comparison with the tribal name Tabɣač ~ Taɣbač should not be discarded very easily, but there are difficulties. May be EMM *to'a* 'number' (MNT §229, §265, §278), (HYYY 22b.2), (KMQB I: 9b.2), but note that in other cases MM *-'-* ~ WM *-ɣ-* is reflected mostly as *-ɣ-* in the language of the KhT inscription (except between the identical vowels). WMM *tü'ü-* 'to gather' (Mu 159) is more problematic for phonological and syntactical reasons. Maue suggested tribal name Dūbō ~ Dūbò (都波 ~ 都播 EMC *tuo pwâ ~ tuo pwâ*^{C}).

4. *pu̱ro̱-r* is obscure. *-r* can be a Mongolic deverbal nominalizer *-r*, as in WMM *amu-r* 'peace, patience' (Mu 102) < MM *amu-* 'to rest' (MNT §145, §201, etc.), (Mu 102), (IM 432), but the root is opaque.[9] Róna-Tas suggested MM *bol-* 'to become, to be, to be possible, or allowed' as well as Khitan *po(l)-/bo(l)-* 𘓱/𘐀𘐔 (295/311.186) <p>, <b.o> 'to appoint, promote, become' (Kane 2009: 112), *boor* 𘓱𘐔𘐀 (295.186.255) <p.o.or>, 𘐀𘐔𘐀 (311.186.255) <b.o.or> 'made, caused to be [promoted to a position]' (K 2009: 84, 112), (Yelü Xiangwen 24-37), *jautau boor* <jau tau b.o.or> 'became a punitive comissioner', *boor* is a finite past time masculine form (Xiao Dilu 5-6), (Wu & Janhunen 2010: 67).[10] However, this is problematic for both morphological and phonetic reasons: Khitan past tense suffix *-or* does not reveal itself otherwise in the KhT inscription and the correspondence of KhT *-r-*: MM *-l-*: Khitan *-r-* is also questionable.

[8] *-nar* is probably the Mongolic plural marker *-nAr*, but there is a functional problem with this interpretation. In MM it is always used with either kinship terms or with deities. Khitan plural marker *-ńər ~ -ńəń* is also used for nouns denoting humans (Shimunek 2017: 264). Thus, in both Mongolic and Khitan these suffixes are limited in usage, being reserved only for animate nouns. However, cross-linguistically the functional changes in plural markers are not unknown. Thus, for example, WOJ *-ra* originally used to be a neutral plural marker, used both with animate and inanimate nouns, but virtually acquired pejorative usage with animate nouns in Middle (Classical) Japanese (Vovin 2005: 93). In the similar fashion, WOJ plural marker *-tati* for animate nouns gradually became highly restricted honorific plural marker in MJ (Vovin 2005: 100). The plural marker *-*^{n}*dömö* was restricted to animate nouns designating people in WOJ (Vovin 2005: 98), but in MJ its descendant form *-domo* is used with both animate and animate nouns (Vovin 2003: 41). Also, even if Mong. *čilaɣun* 'stone' (HT *ǰīlo*) is connected etymologically to EOT *taš* 'id.' (it might be not), there is no problem with palatalization here in spite of *biti(g)* 'inscription', where the palatalization *ti* > *či* is still absent. There might be two reasons for it. First, it could be a loan from pre-WOT, which already had a palatalization (cf. Chuvash *čul* 'stone'). Second, it could reflect PT *tyaš 'stone', namely the initial *ty-.. > EOT *t-*, Chuvash *č-*.

[9] If the rest of this line is deciphered correctly, then *pu̱ro̱-* is likely to mean 'to sin', 'to be evil', or something close: "Tupa [people] stabbed (i.e., exterminated) the sinners (or sins), and joined the saved".

[10] P.c., October 2017.

5. Maue suggested Mongolian *čiči-* 'to stab' + *-rA*, converbum finale. The form *čiči-* appears to be Post-Pre-Classical, as it is not attested in MM or Pre-Classical WM. Cf. EMM *seči-* 'to stab' (MNT §127). Converbum finale *-rA* is well attested in both MM (Godziński 1985: 146) and pre-Classical WM (Poppe 1964: 98). Róna-Tas suggested Mong. *čečere-* 'to tremble' ('to be afraid'), but this is unlikely, because it seems to be a loan of OT *titrä- ~ titirä-* 'id.', and in KhT there is still shift of *t* > *č* before [i] (cf., e.g. the fact that KhT still has *biti(g)* and not *bičig* 'inscription').

6. *pügtig* is obscure. I personally do not have any possible etymologies for it, although Marcel Erdal, Jens Wilkens, Peter Zieme and Róna-Tas independently suggested OT *bögtäg* 'merit, saved', which is itself a loan from Sogdian *βwɣtk* [βuγdē < βuγdak] 'saved' (Gharib 2004: 114). As was pointed by Wilkens in a discussion of 1.1 draft version of this paper on Academia.edu, in this inscription it can be a direct loan from Sogdian.

7. *ńele-n* could potentially be EMM *neyile-* or WMM *neile-* 'to join', followed by the converbum modale *-n*. Róna-Tas suggested that alternatively it may be Mong *yala* 'guilt, crime', although there is no evidence for a final *-n* in thus word (p.c., October 2017). In spite of the fact that KhT {ạ} (in script) denotes [a] or [e], this is highly unlikely, especially that *ń-* > *y-* also remains to be explained. Maue's reading *ñạlụn,* which make possible the derivation from *nalu-* 'to be bend, to bowed, to be sunk' (Schmidt 1835: 82a.) may make sense semantically and syntactically (although the question remains why saved ones are bowed and do not simply bow), but there are two problems: first, KhT {ạ} is a poor equivalent for [u] in the second syllable, and second, *nalu-* is not attested in MM or Pre-Classical WM.

8. There is some Mongolic morphology in this line as well: plural suffix *-nAr*, converbum modale *-n*, converbum finale *-rA*, and possibly a nominalizer *-r*. The first two has been already discussed above. As for the converbum finale *-rA* and nominalizer *-r*, I do not think they are attested in Khitan.

LINE 1-5

Open transcription (Maue):

× *q*[a]*γ*a[nu?] q*ato-ñar* dügi$_1$d n$_{ı1}$rı̣ qaγan türūg qaγan-

Tentative transcription:

[+] q[a]ɣa[n-u?] qato-ńar düge-d nı̄rı̣ qaɣan türüg qaɣan-

Morphemic analysis:

[?] qaɣa[n-GEN?] qatun-PLUR younger.brother-PLUR Niri qaɣan türk qaɣan-

Notes:

1. The first *akṣara* is unreadable.

2. In spite of the poor preservation, the next word must be *qaɣan-u* 'qaɣan-GEN'. It is possible that the next *akṣara* indicated just *-n*, forming therefore, an alternative genitive in *-un* after *qaɣan*, see more on this in the notes to line 1-9.

3. *qato-* can be equated with EMM *qatu* (MNT §54, §55) and MM *qatun* (MNT §64, §104, etc.), (IM 444), (Mu 203) 'qaɣan's wife, lady, noble woman'. *-ńar* is a plural suffix already seen above.

4. *düge-d* is 'younger brothers', metathesized form of MM *de'ü* 'younger brother' (MNT §18, §46, etc.), (Mu 151), followed by a plural suffix *-d*. Róna-Tas came up with an alternative interpretation, suggesting OT *tüke-, tüge-* 'to come to an end, finish', and Mongol *töge-ri-* 'to be lost', meaning that the qaɣan and his wives were executed, finished (p.c., October 2017). Although this explanation saves us from using the metathesis, it leaves the morphology, namely *-d*, stranded without any explanation.

5. Niri qaɣan of the Western Turkic Khaganate (r. 579-603/604 AD) is a grandson of Bumin qaɣan. The etymology of the name Niri cannot be Turkic because of the initial *n-* that is alien to native Turkic. Possibly it comes from Sogdian: *nyrk$^{(')}$* ~ *nyrq* 'male' (Gharib 1995: 252), *nyrk* 'virile' (Benveniste 1929: 226). The segment *-k* is clearly a suffix, at least diachronically, cf. Old Persian *nar* 'man', also Šugnani, Rošori *nīr* 'male', Rušani *nēr*, Yaghnobi *nar* (< ?*nīr, under the influence of Tajiki *nar*) 'id.', also Middle Persian *nērōk ~ nērōg* (*nylwk' ~ nyrwg*), Classical Persian *nērō*, Old Tajiki *nerů ~ nirů* 'strength, might' (Edel'man 2015: 478). 'Viril' or 'mighty' seems to be a fitting name for a *qaɣan*.

6. The rest of the line is quite straightforward: *türūg kagan* is, of course, *türük qaɣan* 'Turkic qaɣan'.

7. We again see typically Mongolic morphology in this line: besides the genitive case suffix *-u* (or may be *-un*) and plural suffix *-ńar* (used this time after an animate noun), we also encounter for the first time Mongolic plural suffix *-d*.

LINE 1-6

Open transcription (Maue):

u*n* d(o̧)ro̧ *tạyạǰụ* ksı̣$_1$rı̣ härgin barγo[l] pạl*k*sı̣$_1$r [+]ksạčı̣ hi$_1$gbi$_1$ǰ

Tentative transcription (Vovin)*:*

-un d[ö]rö taya-ǰu ksı̣$_1$rı̣ hergin bar-ɣo[l] pạlksı̣$_1$-r [+]ksạ-či hi$_1$gbi$_1$-ǰ

Morphemic analysis:

-GEN law worship-CC country title/people
take-NML ?-NML ?-NA ?-PAST

Notes:

1. Among two variants for the first akṣara *un* and *uč*, *uč* is not clear. A reduction from EMM *učir* 'chance, reason, cause' (KMQB I: 21a.2, 21b.1, 25b.4, 26a.3, 28a.3) does not seem very probable. Even less probable would be a connection with WMM *ūča* 'back' (Mu 96, 154, 371), (Ist 71), *uča* (L 1255). I opt for *-un*, genitive.

2. *dörö* is 'law, rule, principle'. Cf. EMM *töre* (MNT §208), *törö* (MNT §216). For the discrepancy between KhT *d-* and EMM *t-* cf. also KhT *digin* vs. *tegin* on line 1-1 as well as Manchu *doro* 'id.', Jurchen *doron* 'imperial seal', borrowed from Mongolic with initial *d-*, not *t-*.

3. *taya-* probably can be explained as EMM *tayi-* 'to worship' (MNT §189). *-ǰu* is a converbum contemporale that we have already seen several times.

4. For *härgin* Maue suggested OT *erkin* ~ *irkin* 'title' (p.c.), discussed by Clauson (1972: 225) and in other sources. It might be further etymologically connected to OT *erk* 'authority, free-will, independence' (Clauson 1972: 22). This may indicate a further connection with EMM *erkin* (sing.) ~ *erkid* (plur.) 'important, excellent, masterly' (de Rachewiltz 2004.1: 415), attested in MNT §105, §109, §153, etc., but not used as a title in Mongolic. It is not inconceivable, on the other hand, that we have here MM *irgen* [irgən] 'people' (with a metathesis of vocalism) (MNT §5, §8, §11, etc.), (HYYY 14b.1), (Mu 119, 123, 136, etc.). Both explanations have their pluses and minuses. Maue's theory does not involve metathesis, but leaves *-g-* instead of *-k-* unexplained (Róna-Tas also informed us that *ir.g.en* in Khitan may be used as a title). Also, in the KhT inscription titles have consistent plural marking when the plurality is meant. May be only one *erkin* is meant here, it is difficult to say, even if the next word is explained as 'taker, collector' (see below). Vovin's explanation has a metathesis, but does not create problem with the consonant *-g-*. Also, *irgen* is never marked with plural.

5. *bar-ɣo[l]* is probably a cognate of MM *bari-* 'to take, to collect' (MNT §19, §25, etc.), (Mu 199, 103, etc.) + MM *-'Ul* (Godziński 1985: 53-54), Pre-Classical WM *-ɣul* ~ *-gül*, a deverbal nominalizer designating occupations, possibly a 'collector'. Alternatively, this form can be interpreted as *bari-* + a cognate of MM *-'Ul* (Godziński 1985: 53-54), Pre-Classical WM *-ɣul* ~ *-gül*, a causative suffix, followed by a zero-marked imperative.

6. The rest of this line is opaque, although *-r* in *paḷksɪ$_1$-r* is likely to be a nominalizer (already discussed in the note 4 to line 1-4), *-či* in *[+]ksạ-či* a Turco-Mongolian nomen actoris, and *-ǰ* in *hi$_1$gbi$_1$-j* can probably be equated with MM past tense in *-ǰi found in distant past *-ǰU'U(y)* and its feminine form *-ǰi'ay* (Godziński 1985: 131-132).

7. Again we have some clearly Mongolic morphology here: *-ǰu, -ǰ, -r* and *-ɣol*.

LINE 1-7

Open transcription (Maue):

tügǰü *uqab*ar-*ñar* qaɣan ksạnɪ ǰụḷạbạ tụnụ
tụšn[] *tụ*wạ

Tentative transcription (Vovin):

tüg-ǰü uqa-ba-r-ńar qaɣan ksan-ɪ ǰula-ba tün-ü
tüš(i)-n[] tuwa

Morphemic analysis:

be.enough-CC realize-PST-NML-PLUR qaɣan
regnal.year-ACC shine(?)-PAST he-GEN
entrust/support-ADN(?) tribal.name

Notes:

1. *tüg-* can be compared either with EMM *tüge-* 'be enough' (§92, §187) (cf. the similar elision of the second syllable vowel reflected by MM *bari-*: KhT *bar-* on the previous line) or with MM *tüge'e-* 'to give out, to distribute' (MNT §213, §232), (Mu 358), possibly a derivative of *tüge-*. Converb *-ǰU* has previously appeared on lines 1-2, 1-3, and 1-6. Róna-Tas suggested Mong. *tug*, Khitan *tuġ(u)t* 令久令 'flags' (247.172.247) <t.uh.t> (Xiao Dilu 26-23) < Mandarin Chinese *dù* 纛, 'yak-tail banner', EMCh *tuo^C*,[11] LHC *ta^C* (p. c., October 2027), but it does not agree with the following *-ǰü*. But the main problem, of course, is that *tug* 'standard' has a back vowel, while *tüg-* has the front one.

2. *uqa-* 'to realize, to know' was discussed previously in conjunction with line 1-2. *-bA* is MM past tense suffix that appears twice on line 1-7. *-r* could be a nominalizer previously mentioned in the note 4 to line 1-4, although there might be a problem with this interpretation, since MM form in *-bA* is a final verbal form (Godziński 1985: 128-129), which has no adnominal function. But since the language of the KhT inscription is not MM, I am going to propose this explanation especially that adnominal forms in East and Central Asia tend to acquire easily final predication function. Note also that Khitan past tense suffix *-bəń*[12] can be used in the adnominal function

[11] I use Schuessler (2009) or Baxter/Sagart (2014) reconstructions (the latter are specifically marked), while Róna-Tas relies on Pulleyblank (1991).

[12] Kane reconstructs *-bo.ń ~ -b.uń ~ -bun* (2009: 147).

(Shimunek 2017: 271). *-ńar* is plural suffix, which we have already seen on more than one occasion. Thus, *uqaba-r-ńar* is 'realizations' or 'those who realized'.

3. *qaɣan* does not require any comments and *ksạn-* according to Maue is a 'regnal year' < Tumshuquese *xšana-* 'Zyklusjahr' + accusative *-ı*.

4. *ǰula-* can be possibly connected with MM *ǰula* 'light, candle, torch' (HYYY 10b.4), (HS 10.20), (Mu 169, 210, 325, 336), (IM 439), (Ist 39) (< WOT *ǰula, cf. EOT *yula* 'light, torch'),[13] which is a nominal root, but it is possible to suspect conversion here, so I tentatively interpret it as *ǰula-* 'to shine'. Cf. MM *to'o ~ to'a* 'number' --> *to'o-* 'to count'.

5. *tün-ü*, *tün-* must be cognate to EMM *te'ün-* (MNT §12, §116, etc.) ~ WMM *tn̄-* (Mu 94, 97, etc.), the oblique stem of *tere* 'that', that can also be used as a third person pronoun, especially anaphorically. *-ü*, is, of course, a genitive form after stems in *-n* (see the note 5 to line 1-2).

6. *tüš(i)-n[]* consists of *tüš-*, a most likely cognate to EMM *tüši-* 'to entrust', to rely upon', possibly also 'to support'. The last word on this line is the enigmatic *tu̱wạ̱* (see note 3 to line 1-4). Unless the loss indicated by Maue's [] really occurred, the only possibility to explain the remaining *-n̲* is a comparison with a Khitan adnominal form in *-n*, that Kane, I believe, incorrectly calls a verbal noun (2009: 155).

7. As before, we see a number of typically Mongolic morphological markers: converb contemporale *-ǰU*, past tense suffix *-bA*, nominalizer *-r*, plural suffix *-ńAr*, and genitive suffix *-U* after *-n* stems.

LINE 1-8

Open transcription (Maue):

[]ıksạ[] *tu̱ toɣoɣun p*[ü]*gtig*či śı̨ñın *bo*dısatva
tǫ̱rǫ̱ks qaɣan

Tentative transcription (Vovin):

[]ıksạ[] tu̱[] to[ɣo]-ɣun pügtig-či šińi-n bodi-satva
törö-ks[A] qaɣan

Morphemic analysis:

? ? count-NML salvation-NA new-GEN
Bodhi-sattva be.born-PAST.PART qaɣan

Notes:

1. The first two words on this line resist any interpretation.

2. *to[ɣo]-* can be compared with EMM *to'o-* 'to count' (MNT §161), WMM *tō-* (Mu 115). *-ɣun* must be cognate to MM deverbal nominalizer *-'Un* (Godziński 1985: 53), and Pre-Classical WM *-ɣun ~ -gün* (Poppe 1964: 46), which forms nouns designating qualities or abstract nouns.

3. Enigmatic *pügtig* previously appeared on line 1-4 as *pügti$_1$g*. Here it is probably followed by nomen actoris suffix *-či*. The only information it provides us is that *pügti$_1$g ~ pügtig* must be a noun. Róna-Tas suggested that $^{p}/_{b}$*ügtägči* is 'the person who got salvation' (p. c., October 2017).

4. The rest of this line is identical to the second part of line 1-1 and the first word of line 1-2.

5. Again we see several typically Mongolic morphological markers: nominalizer *-ɣUn*, genitive case marker *-n* after vowel stems, past participle *-ks[A]*, and nomen actoris *-či* (which is, of course, a common Turco-Mongolic marker, but it remains unclear who borrowed it from whom).[14]

LINE 1-9

Open transcription (Maue):

[+]*l*[] + k*iyü* u*n* bi*ti*$_1$*h*i$_1$*ñ* *puɣan* *tu̱*wạ-*ñ*ạr
qaɣan *tǫ̱*rǫ̱ks qaɣanun

Tentative transcription (Vovin):

[+]*l*[][+] ki-yü un bitig-ińˊ puɣan tuwa-ńar
qaɣan törö-ks[e] qaɣan-un

Morphemic analysis:

? do-PRES ? inscription-GEN ? ? tribal.name-PLUR
qaɣan be.born-PAST.PART qaɣan-GEN

Notes:

1. This line is poorly preserved. The first, fifth, and sixth words are unreadable.

2. *ki-yü* is probably the deductive present in *-yU* for the verb *ki-* 'to do, to make'. Cf. MM *ki-* (MNT §24, §56, §134, etc.). On MM *-yU* see Godziński (1985: 127), and on WM *-yU* see Poppe (1964: 92).

3. The third word *un* is unclear. It cannot be a genitive, because the deductive form in *-yU* is a finite verbal form and also because it contradicts vowel harmony.

4. *bitig-ińˊ* is a genitive form of *bitig* 'inscription', which previously appeared as *biti[g]* on lines 1-1 and 1-4.

[13] Possibly from Turkic *yul- 'to catch fire, to be ignited, be kindled' (Róna-Tas & Berta 2011: 409).

[14] Since the KhT inscription is more than one century older than the earliest OT texts, it is likely that the directionality of borrowing might have been Mongolic → Turkic, and not the vice versa.

5. It is unclear what *puɣan* might mean. Judging by its syntactic position it must be a noun, and not a verb *puɣa-* + converbum modale *-n*.

6. On the enigmatic *tu̱wa̱* see note 3 to line 1-4. This word also appears on line 1-7.

7. *qaɣan* does not require commentary and *törö-ks[e]* has been treated above in the note 5 to line 1-1. It also appears on line 1-8. After *tuwa* 'Tupa' we again see the plural marker *-ńAr*.

8. The language of the KhT inscription has an alternative genitive *-un* for *qaɣan: qaɣan-un,* which is probably a result of secondary paradigmatic reanalysis, although it also could be a metathesis of a form similar to EMM alternative genitive *qaɣan-nu* (MNT §1, §48, §51, etc.).

9. Again several clearly Mongolic morphological markers are present: genitives *-un* and *-iń*, plural *-ńAr*, deductive present *-yU*, and past participle *-ks[e]*.

LINE 1-10

Open transcription (Maue):

+ pa̱da̱ *n*<ɪ$_1$>rɪ̱ qaɣa*n* türüg *qa*ɣa[n]
[..]ǰɪ̱n ubɪ$_1$ǰ ǰa̱ɩ̱o̱ba̱ǰ darqad ǰa̱ɣa̱bɪ̱

Tentative transcription (Vovin):

[sina]pa-da Niri qaɣan türüg qaɣa[n]
[k/g]iǰi-n ubi-ǰ ǰalo-ba-ǰ darqa-d ǰay bi-

Morphemic analysis:

domain-LOC Niri qaɣan türk qaɣan
approach/follow-CM ?-PST direct-PAST-PAST
free.man-PLUR happy be-

Notes:

1. There is a lacuna before *pada-*. KhT *-pa* can correspond only to MM *-'a*. Given the preceding genitive form *qaɣan-un* of *qaɣan* on line 1-9, I believe we have only one option for the restitution of the lacuna, namely *sina*, which gives us EMM *sina'a* 'land, country, domain, province' (MNT §141). Therefore, I reconstruct KhT *[sina]pa* 'id.' Final syllable *-da* can be equated with a locative MM case marker *-dA* (Godziński 1985: 85-87).

2. We have already met Turkic qaɣan Niri in line 1-5, so Niri qaɣa*n* türüg [qa]ɣa[n] does not require any commentary.

3. []*iǰi*-n looks like a verbal form ending in converbum modale *-n*, but establishing the identity of the root might be difficult, although MM *[g]iǰi-* 'to follow' (MNT §253), *[k]iǰi-* 'to approach' (MNT §195) seem to be the only candidates.

4. The root *ubi-* does not have any parallels, but on distant past *-ǰ* see the note 6 to the line 1-6.

5. *ǰalo-ba-ǰ* may consist of the root *ǰalo-* (cf. WM *ǰalu-* 'to direct', 'to make straight'[15]), past tense *-bA*, and distant past *-ǰi*. Róna-Tas suggested *jalobaj*, perhaps to be related to Mong. *jalbari-* ' to pray' <*jalba- and to be read as *jalbaj-* (p.c., October 2017), but it leaves the verbal form without inflection, and besides I can see no phonetic justification or motivation for the loss of a whole syllable *-ri-*.

6. *darqa-d* is a plural form of *darka-n* 'free man, official'.

7. *ǰa̱ɣ* looks like a cognate to EMM *ǰaya'an* 'happy, blessed' (MNT §66, §194, §248), where *-'an* is likely to be a suffix.

8. *bi* is a cognate to MM *büi* 'to exist' (MNT §18, §35, §44, etc.), (Mu 105, 110, 139, etc.).

9. There are several typically Mongolic morphological markers on this line as well: converbum modale *-n*, distant past *-ǰ*, past *-bA*, locative case marker *-dA*, and plural suffix *-d*.

LINE 1-11

Open transcription (Maue):

[?]*ru̱n* b*i*tig + s*a̱*[] p*aɣ* [+ +] ǰ[] [?] darqan
b[ɪ]tɪ*ba̱* qa

Tentative transcription (Vovin):

-rün bitig [+]sA[] paɣ [+ +] ǰ[] [?] darqa-n
b[i]ti-be qa

Morphemic analysis:

-CP inscription ? ? ? ? ? official-SING
write-PST ?

Notes:

1. The last line is in a very poor state of preservation, but it looks like a signature of a scribe.

2. Since [?] in the beginning of this line represents an uncertain loss, I choose to ignore it, and treat *rUn* as *-rün*, a marker of converbum praeparativum, linked to *bi-* 'to be', 'to exist', therefore *bi-rün* 'as it was'.

3. *bitig* is 'inscription', see also lines 1-4 and 1-9

4. *darqa-n* is a singular form of *darqa-* 'official, free man' that we have already seen as plural *darqa-d* on line 1-10.

5. *biti-be* is the past tense form of *biti-* 'to write'. The rest of the line is undecipherable at the present moment.

6. Three Mongolic morphological markers appear on this line: singular *-n,* converbum praeparativum *-rUn*, and past tense suffix *-bA*.

[15] Not attested in Pre-Classical WM (Tumurtogoo 2006) or in MM.

TENTATIVE PARTIAL TRANSLATION

1-3. Qaɣan [of] the inscriptions died and when a new qaɣan, who was [re]born as a new Bodhisattva, knows lord Buddha knowledge, and promises … the country's Aṅaqay [title], begs and tribes, stand, and listen together… **4.** Looking at the inscription stones, Tupa [people] exterminated [their] sins and joined the saved **5.** … qaɣan's wives [and] younger brothers, [and] Niri qaɣan, qaɣan [of] Türks **6.** worshiped the Law, and country's erkins and collectors … **7.** are enough and those who realized that qaɣan's regnal years were shining, and Tupa whom he supported/entrusted **8.** counting … those who attained salvation … qaɣan who was [re] born as a new Bodhisattva **9-11.** do… of the inscription …the qaɣan of Tupa was [re]born. In the qaɣan's domain, [they] followed Niri qaɣan, qaɣan of Türks and … [He] directed [them]. As the free men were happy, inscription … official wrote …

COMPARATIVE CHART OF KHÜIS TOLGOI AND MONGOLIC/KHITAN MORPHOLOGY

marker	**Khüis Tolgoi**	**MM**	**Pre-Classical WM**	**Khitan**
genitive after *-n* stems	-U ~ -Un	-U ~ -nU	-U	-en
genitive after consonantal stems	-Un ~ -iń	-Un	-Un	-un, -en,
genitive after vowel stems	-n	-yin, WMM -īn, -n[16]	-yin	-n, -on, -un
locative	-dA	-dA	--	-de, -do, -du
accusative	-ı ~ -i	-i ~ -yi	-i ~ -yi	-Ø
plural suffix	-ńAr	-nAr	-nAr	-ńer ~ -ńeń
plural suffix	-d	-d	-d	-d
singular suffix	-n	-n	-n	--
nomen actoris	-či	-či	-či	--
nomen praesentis (with converbial function)	-yi > -Ø (after -yi)	-(U)yi	-(U)yi	-Vi
converbum modale	-n	-n	-n	--
converbum contemporale	-ǰU	-ǰU	-ǰU	-ǰ ~ -č
converbum finale	-rA	-rA	-rA	--
converbum praeparativum	-rUn	-rUn	-rUn	--
adnominal	-n[]	--	--	-n
past	-bA	-bA(i)	-bA(i)	-beń
distant past	-ǰ	*-ǰi	-ǰuqui ~ -ǰüküi	--
deductive present	-yU	-yU	-yU	--
nominalizer	-ɣuń ~ -ɣun	-'Un	-ɣun ~ -gün	--
nominalizer	-r	-r	-r	--
nominalizer	-ɣol	-'Ul	-ɣul ~ -gül	--
functionally unclear verbal suffix	-n[V]yA-	--	--	--

I believe that the chart above clearly demonstrates that the language of the KhT inscription shows much more features common with mainstream Mongolic than with Khitan. This also makes sense geographically, since the KhT inscription comes from the territory much more to the North than the home territories of Serbi and Khitan were.

Certainly, we have sufficient lexical difference between the language of KhT and the first monuments in MM. But this also should come as no surprise, given the fact that more than half of a millennium separates these texts, as well as the hard-core reality (sorely missed by Nostraticists and other long-rangers) that the lexicon represents the most unstable part of any language.

[16] After stems ending in *-ai*.

Some notes on the language of the non-Sogdian part of the Bugut Incription

I agree with Maue that in spite being a bilingual text, this is a much harder nut to crack than the KhT inscription. It is written in the same Brāhmī script, and the preliminary and very tentative conclusion is that it is in the same language. The Bugut inscription is even in the worse state of preservation than the KhT inscription, so that a reconstruction of any single word is highly hypothetical and provisional, and in addition it seems that the contents of the Sogdian and Brāhmī inscriptions only partially coincide, so it is not a Rosetta stone by any standards. So far I have managed to decipher the first line, which appears to be better preserved that the other lines. This is still the work in progress, but we hope to report our preliminary findings at the 61st meeting of PIAC in Bishkek at the end of August 2018.

Conclusion

I trust that the most innovative and revolutionary linguistic part of our discovery is that the first 'Altaic'-type language attested in the steppe by a continuous text is not Old Turkic, but Mongolic. Apart from the fact that it now places Mongolic as a textually attested language family in the late sixth or early seventh century and makes it a contender among 'Altaic' languages with Koreanic that has the first continuous text attributed to 594 or 596 AD. For the sake of the comparison, the first Old Japanese text is dated only by 697 AD, and the first OT text by 713 AD. Tungusic Jurchen texts are not attested before the twelfth century.

This discovery can also potentially reverse, at least in some cases, the traditional point of view that the directionality of borrowing was always from Turkic to Mongolic, as some of them will have now more certain Mongolic origin.

Finally, it appears that the same language as in the KhT inscription appears in the same script, and more importantly in the same language in much shorter Keregentas inscriptions from Eastern Kazakhstan. What polity or what tribe/nation might have possibly spoken this language is the question that should be answered by historians. As a very speculative proposal, based entirely on linguistic grounds, I would suggest that KhT inscription is in the Tabɣač/Taɣbač language. It is quite clear that the specimens of the Tabɣač/Taɣbač language in the Chinese transcription that were long the subject of scholarly study,[17] are much closer to Khitan than to mainstream Mongolic. But then Li Yong-Sŏng has recently and very persuasively demonstrated that there were two distinct Tabɣač polities: Northern Qi and Northern Zhou (2017) that quite well could represent two different 'Tabɣač' languages, because it is well known how easily ethnonyms can migrate across ethnic and state boundaries.

[17] To mention Boodberg (1936), Bazin (1950), Ligeti (1970), Chen (2005), Vovin (2007), and most recently Shimunek (2017) just as a few representative ones.

Abbreviations of languages and grammatical terms

CC	Converbum contemporale
CF	Converbum finale
CM	Converbum modale
CP	Converbum praeparativum
COOP	Cooperative
EMC	Early Middle Chinese
EMM	Eastern Middle Mongolian
EOT	Eastern Old Turkic
GEN	Genitive
IMP	Imperative
LHC	Late Han Chinese
NA	Nomen actoris
NML	Nominalizer
NP	Nomen praesentis
OT	Old Turkic
PART	Participle
PAST	Past tense
PLUR	Plural
PRES	Present
SING	Singular
WM	Written Mongolian
WMM	Western Middle Mongolian
WOT	Western Old Turkic

Primary sources:

Khitan

Hsiao Hui-lian.Can	Inscription on the canopy of Hsiao Hui-lian muzhi ming, 1080 AD
Langjun	Langjun inscription, 1134 AD
Xiao Dilu	Xiao Dilu Fushi muzhi inscription, 1114 AD
Yelü Jue	Yelü Jue muzhi ming inscription, 1071 AD
Yelü Pusuli	Yelü Pusuli muzhi inscription, 1105 AD
Yelü Xiang-wen	Yelü Xiangwen muzhi inscription, 1091 AD

Middle Mongolian

HS	The edict of Hai-shan, 1305 AD
HYYY	Hua-yi yi-yu, 1389 AD
IM	Ibn Muhannā vocabulary, 14th c. AD
Ist	Istanbul vocabulary, 14th or 15th century AD
KMQB	Kitad Mongɣol Qarilčaɣan-u Bičig (Sino-Mongolian Documents), 14th c.
L	Leiden vocabulary, 1343 AD

MNT	Mongγol niuča tobča'an, ca. 1240 AD
Mu	Mukaddimat 'al-'Adab, 14th century AD
Qub IV[18]	Fourth edict of Qubilai qaγan, 1280 AD

Secondary sources:

Базылхан, Н., 2005: Қазақстан тарихы туралы түркі деректемелері, II том, көнетүрік бітіктастары мен ескерткіштері (Орхон, енисей, талас), Алматы.

Baxter, William & Laurent Sagart 2014. *Old Chinese: A New Reconstruction*. Oxford & New York: Oxford University Press.

Bazin, Louis 1950. Recherches sur les parlers T'opa (5e siècle après J.C.). *T'oung Pao* 39: 228-329.

Benveniste, Émile 1929. *Essai de grammaire sogdienne. Deuxième partie: morphologie, syntaxe et glossaire*. Paris: Librarie orientaliste Paul Geuthner.

Boodberg, Peter A. 1936. The language of T'opa Wei. *Harvard Journal of Asiatic Studies* 1: 167-185.

Chen, Sanping 2005. Turkic or Proto-Mongolian? A Note on the Tuoba Language. *Central Asiatic Journal* 49.2: 161-173.

Chinggeltei, Wu Yingzhe, Jiruhe 2018. *Qidan xiaozi zai yanjiu*. Vol. 1-3. Hohhot: Nei Menggu daxue chubanshe.

Clauson, sir Gerald 1972. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford: Clarendon Press.

Diyarbekirli, Nejat 1979. "Orhun'dan Geliyorum", Türk Kültürü, Sayı 198-199, cilt XVII, Nisan-Mayıs 1979: 321-384 (1-64).

Edel'man, Joy I. 2015. *Этимологический словарь иранских языков*, Vol. 5: l-n. Moscow: Nauka – Vostochnaia literatura.

Gharib, B. 2004. *Sogdian Dictionary (Sogdian-Persian-English)*. Teheran: Farhangan Publications.

Godziński, Stanisław 1985. *Język Średniomongolski*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Жолдасбеков, Мырзатай – Қаржаубай Сартқожаұлы, 2006: Орхон ескерткіштерінің толық атласы, Астана / Жолдасбеков, Мырзатай - Каржаубай Сарткожаулы 2006: Атлас орхонских памятников, Астана.

Kane, Daniel 2009. *The Kitan Language and Script*. Leiden & Boston: Brill.

Ligeti, Lajos 1970. Le tabghach, un dialect de la langue sien-pi. In: *Mongolian Studies*, ed. by L. Ligeti, Budapest: Akadémiai Kiadó, pp. 265-308.

Maue, Dieter 2008. Three languages on one leaf: on IOL Toch 81with special regard to the Turkic part. *BSOAS* 71: 59-73.

Ōsawa Takashi, Suzuki Kōsetsu, R. Munhutoruga 2009. モンゴル国現存遺跡・突厥碑文調査報告 – Бичээс II / Biçeesu II. 2006 оноос 2008 оны Монгол улс дахь Түргиын бичээс ба эртний дурсгалыг судлах экспедицийн илтгэл, Улаанбаатар хот 2009.

Poppe, Nicholas 1957. *The Mongolian Documents in ḥP'ags-pa Script*. Wiesbaden: Otto Harrassowitz.

Poppe, Nicholas 1964. *Grammar of Written Mongolian*. Wiesbaden: Otto Harrassowitz.

Pulleyblank, Edwin G. 1991. *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: UBC Press.

de Rachewiltz, Igor 2004-2013. *The Secret History of Mongols. Translated with a Historical and Philological Commentary*. Vols. 1 & 2 (2004), vol. 3 (2013). Leiden & Boston: Brill.

Róna-Tas, András & Árpád Berta 2011. *West Old Turkic. Turkic Loanwords in Hungarian*. 2 vols. Wiesbaden: Harrassowitz.

Сартқожаұлы, Қаржаубай, 2003. *Орхон мұралары*, Астана.

Shimunek, Andrew 2017. *Languages of Ancient South Mongolia and North China*. Wiesbaden: Harrassowitz Verlag.

Schuessler, Axel 2009. *Minimal Old Chinese and Later Han Chinese. A Companion to Grammata Serica Recensa*. Honolulu: University of Hawai'i Press.

Tumurtogoo, D. 2006. *Mongolian Monuments in Uighur-Mongolian Script (XIII-XVI Centuries)*. Taipei: Academia Sinica.

Tumurtogoo, D. 2010. *Mongolian Monuments in 'Phags-pa Script*. Taipei: Academia Sinica.

Vovin, Alexander 2003. *A Reference Grammar of Classical Japanese Prose*. London: Routledge/Curzon.

Vovin, Alexander 2005. *A Descriptive and Comparative Grammar of Western Old Japanese. Part 1: Phonology, Script, Lexicon, and Nominals*. Folkestone: Global Oriental.

Vovin, Alexander 2007. Once Again on the Tabgač Language. *Mongolian Studies* XXIX: 191-206.

Vovin, Alexander 2013. 'Old Turkic Loanwords in the Khitan Language'. In: *Yalım Kaya Bitigi*. Festschrift for Prof. Osman Sertkaya. Ed. by Hatice Şirin User and Bülent Gül, pp. 621-25.

Wu, Yingzhe & Juha Janhunen 2010. *New Materials on the Khitan Small Script*. Folkestone: Global Oriental.

[18] Numeration is given according to the edition by Tumurtogoo 2010, not Poppe's edition of 1957, because many more documents in 'Phags-pa were discovered after Poppe's edition.