John Benjamins Publishing Company



This is a contribution from *Diachronica 27:1*
© 2010. John Benjamins Publishing Company

This electronic file may not be altered in any way.
The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.
Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).
Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

*The Khitan Language and Script.* By Daniel Kane. Leiden: Brill, 2009. Pp. xiv, 306.*

**Reviewed by Guillaume Jacques (CNRS — INALCO, Paris)**

The Mongolic language family has a relatively shallow time-depth compared to other language families of Northeast Eurasia such as Tungusic, Turkic and even Japonic: modern Mongolic languages show less than 800 years of divergence (Janhunen 2003a: 1). Following the expansion of the Mongol empire during the thirteenth century, the diversity of the family diminished considerably. Some modern Mongolic languages preserve marginal distinctions that were not represented in any form of Old Mongolian (Janhunen 2003b: 9), but Old Mongolian can be considered to be very close to the common ancestor of all Mongolic languages. Going further back in time than Old Mongolian is only possible through internal reconstruction and comparison with Para-Mongolic languages, which are related to, but distinct from, the Mongolic languages.

Para-Mongolic languages are barely known *Trümmersprachen* (see Vovin 2007). Of these languages, the only one in which continuous texts have been written is Khitan, the national language of the Liáo Empire (907–1125). This language is known from three sources: first, Chinese transcriptions; second, loanwords in Tungusic languages; third, texts written in the two national scripts, the so-called 'Large Script' and 'Small Script'.[1] Although many words are now understood, Khitan is still largely an undeciphered language. The book under review concentrates on the third source of data, especially the Small Script.

As the author (henceforth K) puts it, "This book is probably more sinological than altaistic, more philological than linguistic, more historical than archeological" (p. xii). Indeed, K's main focus is on the decipherment of the Khitan scripts, rather than the linguistic analysis of the reconstructed language. However, K's

* I wish to thank Nathan Hill and Alexander Vovin for their help on this review. I am responsible for any remaining errors.

1. Kara (1987) proposed the terms 'Linear Script' and 'Assembled Script' to refer to these two writing systems, but we keep here K's terminology. The two scripts are entirely different and barely share any common element, although they are used to transcribe the same language. The Small Script has a limited number of characters (about 400) which represent either full words, syllables, or single consonants and can be combined. The set of Large Script characters is much larger, and its structure is still imperfectly understood.

*Diachronica* 27:1 (2010), **157–165**. DOI 10.1075/dia.27.1.05jac
ISSN 0176–4225 / E-ISSN 1569–9714 © John Benjamins Publishing Company

work is highly relevant to specialists of Mongolic languages and Chinese historical linguistics.

K's book is divided into seven chapters. The first chapter presents a synthesis of previous scholarship, and a detailed explanation of how the Khitan small script has been partially deciphered, thanks to the seminal work of Qīnggéěrtài et al. (1985). The second chapter is an annotated list of Khitan small characters, where a reconstruction is given whenever possible, and a list of all the words that have been deciphered in the Khitan Small Script corpus. The third and fourth chapters consist of an English-Khitan glossary and a short study on Khitan morphology, which will be of particular interest to specialists in Mongolic languages. The fifth chapter discusses the Large Script, the other writing system that was used to write Khitan, but whose decipherment is even more limited than for the Small Script. The sixth chapter is perhaps the most important contribution: a detailed analysis of three Khitan texts in Small Script, with tentative translations. Finally, the last chapter presents a reconstruction of Liao Chinese, the dialect that Khitan was in contact with and borrowed from.

In this review, we concentrate on several issues relevant to historical linguistics: Khitan phonology, Khitan and Mongolic comparative historical phonology, some etymological problems, and language contact (in particular with Tibetan and Tangut).

## 1. Khitan phonology

The reconstruction of the Khitan phonological system is of utmost importance if this language is to be used in comparative studies. The exact number of vowels and consonants in Khitan is difficult to ascertain. K does not provide an explicit analysis of Khitan phonology, but his reading of 'small script' characters reflects the way he envisions the reconstruction of the Khitan phonological system (pp. 29–32). We will discuss a few peculiarities of this system that set it apart from Mongolic languages.

First, K proposes a distinction between velar and uvular stops before several vowels. He reconstructs for instance a distinction between ⟨ki⟩ and ⟨qi⟩, ⟨ku⟩ and ⟨qu⟩ for some phonograms,[2] though unfortunately the exact philological reasons for reconstructing uvulars instead of velars is not set out in detail. If K is right

2. We follow throughout K's transcription of Khitan characters for which a phonetic reconstruction is available. The transcription is indicated in cones, ⟨ ⟩. When a word is composed of several characters, they are separated by a dot in the transcription. The acute accents on some forms are just a graphic device in K's transcription to distinguish homophonous characters.

© 2010. John Benjamins Publishing Company
All rights reserved

about Khitan having a phonemic distinction between velars and uvulars, Khitan would considerably differ in this respect from Mongolic languages, where this opposition is not distinctive (uvulars appear before back vowels, and velars before front vowels). Among the languages of the area, a contrast between uvulars and velars is only attested in Sino-Tibetan languages of the Qiangic group such as Rgyalrong (Jacques 2004: 22–23): other languages such as Uighur and Manchu lack a phonemic distinction between these two places of articulation. If indeed Khitan preserves a series of contrastive uvulars that was lost in Mongolic languages, this would have far-reaching consequences, but more research is needed to confirm or disprove this hypothesis.

Second, some Khitan words are reconstructed with initial *l-*, such as the adjective ⟨l-iau-qú⟩/⟨l-iau-qu⟩ "red" (p. 113). In Mongolic languages, however, words with initial *l-* are not found in the native vocabulary; Khitan does not apparently have the same phonotactic constraint. K compares ⟨l-iau-qú⟩ to MM *hula'an* "red". If this comparison is valid, the presence of initial *l-* in Khitan can be explained as a secondary evolution, due to loss of some initial syllables.

## 2. Khitan and Mongolic

In spite of the scarcity of Khitan data, it can be said with certainty that this language has remarkable archaisms that are not found in any known Mongolic language. In particular, it preserves initial *p-* in words such as ⟨po⟩ "time" (p. 122) which have become *h-* in Mongolic[3] (Janhunen 2003b: 396) and disappeared altogether in modern dialects such as Khalkha (for instance MM *hon* "year", the cognate of Khitan ⟨po⟩, becomes Khalkha *он* [ɔŋ]). It also preserved a distinction between the palatal nasal initial phoneme *ñ and the dental nasal *n (Shimunek 2007) in words such as ⟨ñi.qo⟩ "dog", MM *noqai*.

Khitan is in some ways more innovative than MM, and some of its phonetic innovations have typological parallels in modern Mongolian dialects.

First, it shows assimilation of *i* by *a* in subsequent syllables. This change (known as 'vowel breaking') is well-known in dialects such as Khalkha, where MM *i* followed by *a* in the next syllable changes to *ja*. For instance, the word MM *kitad* "Khitan, Chinese" becomes *хятад* [xjatʰət]. A parallel evolution is attested in Khitan: ⟨ś.au.a⟩ *ɕawa "falcon" (p. 97) is probably related to MM *šibawu.n* "bird". We can reconstruct the following steps from a Proto-Khitan-Mongolian *siba > *ɕiwa > *ɕawa. It should be noted that in this particular word, vowel assimilation

3. Classical and Middle Mongolian forms are cited from Kara (2009). By Middle Mongolian (MM), we mean vocabulary from the Secret History.

© 2010. John Benjamins Publishing Company
All rights reserved

occurred in a very different way in Khalkha *uyɐyy* [ʃʊwʊ] (vowel fusion in between the second and the third syllable of *šibawu.n* occurred *before* vowel assimilation).

In another example, ⟨êm.a⟩ "goat" (p. 99), related to MM *ima'a.n*, Khalkha *ямаа* [jama] "goat", determining whether vowel assimilation was present or absent in Khitan critically depends on how one reconstructs the Chinese dialect that Khitan was in contact with, and how one analyzes the script. Since this philological issue has important consequences for Khitan historical phonology, the data are worth discussing here in some detail.

The character which K transcribes as ⟨êm⟩[4] (#270) is reconstructed in a slightly different way by other authors. Among recent studies on Khitan phonology, Shimunek (2007: 77) proposes reading *jam and Takeuchi (2008: 32, 45)[5] *em or *æm. It is something of a paradox that these diverging solutions are all based on the same Chinese transcriptional data:[6]

**Table 1.** Character 270 in transcriptions of Chinese.

| | Middle Chinese | 'Phags-pa | Khitan characters | Kane's transcription |
|---|---|---|---|---|
| 點 | temX | dem [tɛm] #677 | 254–270 **𘰾𘱇** | ⟨d.êm⟩ |
| | | | 247-327-270 **𘰗𘰼𘱇** | ⟨t-ie-êm⟩ |
| 檢 | kjemX | gem [kɛm] #673 | 334–270 **𘰴𘱇** | ⟨g.êm⟩ |
| | | | 334-327-270 **𘰴𘰼𘱇** | ⟨g.ie.êm⟩ |
| 兼 | kem | gÿam [kjɛm] #694 | 334-327-270 **𘰴𘰼𘱇** | ⟨g.ie.êm⟩ |
| 監 | kæm | gyam [kjam] #698 | 334-335-270 **𘰴𘰽𘱇** | ⟨g.ia.êm⟩ |
| | | | 334-335-184 **𘰴𘰽𘰜** | ⟨g.ia.am⟩ |

The data in Table 1 do not allow any firm conclusion, but suggest that it is unlikely that character 270 alone could represent the rhyme [jam]. Another piece of evidence against this hypothesis is the Khitan title ⟨l.êm.a⟩ "scribe" (p. 104), which is transcribed in Chinese as 林牙 línyá ('Phagspa *lim ya* [lim ja] #723#S22). In this word, the Chinese transcription rules out a reconstruction *ljama.

The exact conditioning of Khitan vowel assimilation is probably very complex, and certainly did not affect all first-syllable *i*'s as it has in modern dialects such as Khalkha.

Second, Khitan possibly shares with Mongolic the change *ti > či. MM has no syllable such as *ti or *di: dentals are always palatalized before i. Khitan is too

**4.** K mentions that ⟨êm⟩ "is probably [iam] ~[iæm] ~[iɛm]" (p. 29).

**5.** This reference is missing in K's bibliography, which is otherwise quite comprehensive.

**6.** The data are cited here from Takeuchi (2008: 47), who presents more variant Khitan graphs than K. The 'Phags-pa Chinese forms are from Coblin (2007).

© 2010. John Benjamins Publishing Company
All rights reserved

poorly known for us to be able to claim that a similar constraint also exists in that language too, but judging from the data in the book under review, it is striking to notice that no such syllables are attested among the reconstructed Khitan, except for the Chinese loanword ⟨hoŋ.di⟩ (p. 95) from 皇帝 huángdì (*xong di* [ɣɔŋ ti] #155 #161). It would be worthwhile looking through all Khitan texts to see whether the combinations ⟨d.i⟩ or ⟨t.i⟩ are attested in any native word.

The Khitan word for "blood" ⟨c.i.is⟩ (pp. 85, 114) can be related to MM *čisü.n* "blood". This pair of cognates presents a correspondence *či : či*, which could reflect the fact that Proto-Khitan-Mongolian *ti palatalized and became *či* in both languages. Unfortunately, this is not proof that palatalisation of dentals occurred in the two branches, as the initial could also have been *č in the proto-language.

Third, Khitan presents simplification of the cluster *Vrs(V) to Vs(V). The main example of this phonetic change is the word for "nine". It is *yisün* in MM, but must be reconstructed *yersUn in Proto-Mongolic, as the Bonan cognate *jirsəŋ* "nine" shows (Janhunen 2003a: 9). Khitan has ⟨is⟩ for "nine" (p. 109), showing that the same change took place. However, it seems not to have happened accross morpheme boundaries, since the plural of "country" ⟨g.úr⟩ is ⟨g.úr:se⟩ (p. 89):[7] final *-r* does not drop after addition of the plural suffix.

Fourth, MM intervocalic voiced stops often correspond to zero in Khitan, in examples such as ⟨heu-úr⟩ "spring" (p. 119, MM *qabur*) and ⟨u-ul⟩ "winter" (p. 126–127, MM *ebül*). It is probable that these intervocalic stops existed in Proto-Khitan-Mongolic and then underwent lenition in Khitan, as they did later and independently in many Mongolic languages (Janhunen 2003b: 398).

In some cases, it is not obvious which of Khitan or Mongolic (or both) has innovated. In particular, we find several words that have a diphthong in Khitan corresponding to a simple vowel in MM:[8]

**Table 2.** The correspondence Vir/l : Vr between Khitan and Mongolian.

| Khitan | Meaning | Reference | MM | meaning |
|---|---|---|---|---|
| *nair | sun | | *nara.n* | sun |
| *sair | moon | p. 107 | *sara.n* | moon |
| ⟨x.ui⟩ or ⟨x.uil⟩ | to arrive | p. 89 | *kür* | to reach |

7. Shimunek (2007: 61) reconstructs *gur-s.

8. The pronunciation of the two first words is not known from their spelling in the Khitan small script, but from Chinese transcriptions, which is why we use an asterisk before it instead of cone brackets, which we reserve for transcriptions of Khitan small characters; Shimunek (2007)'s reconstructions of these two words are used.

© 2010. John Benjamins Publishing Company
All rights reserved

The correspondence Khitan *Vi to a single vowel in MM has a clear phonetic conditioning: it only occurs when the vowel is followed by r (or l). Since the syllable /ar/ is attested in Khitan (p. 49), Khitan cannot have innovated a rule such as *Vr > *Vir. On the other hand, there was no constraint against a group Vir in MM, and we cannot argue that Mongolic languages underwent a change *Vir > Vr. This issue deserves further research.

## 3. Some Mongolic etymologies

Although K sometimes gives Mongolian or Manchu etymologies for Khitan words, in other cases he refrains from doing so, presumably in an excess of caution.[9] In this section, we shall indicate a few possible Mongolic etymologies not mentioned in the book under review.

⟨êm.ci⟩ "to drink" (p. 93) could be tentatively compared with Classical Mongolian *emkü-* "to put into the mouth"; this would imply a correspondence between Khitan ⟨ci⟩ and Mongolic /kü/, which remains to be demonstrated by other examples.

⟨ku⟩, ⟨ku.u⟩ "man" (p. 105) is probably relatable to MM *kü'ü.n*, Classical Mongolian *kümün* "man".

⟨xe.ci⟩ "border" (p. 86) is compared to Jurchen *hečen(i) by K. Another possibility would be MM *kiǰa'ar* "edge", or, as Shimunek (2007: 78) proposes, Daur *kətʃ* "edge".

⟨go.er⟩, plural ⟨go:t⟩ "tent, family" (pp. 102 and 192). An etymology that could be considered for this word is MM *ger* "tent, house" (plural *ger-üt*), if the unexpected correspondence *o : e* can be explained.

The Khitan small character 68 is reconstructed as ⟨us⟩ by K (p. 42). This character occurs in the word ⟨68.gi⟩ meaning "character, writing", and K proposes a comparison with Classical Mongolian *üsüg* and reads it as ⟨us⟩. However, it is equally probable to compare it with MM *bičig* (a loanword from Uighur) and read ⟨68⟩ as *bit or *bič.

K does not give a reconstruction for the character ⟨191⟩ meaning "six", but suggests either *nil or *jir based on the comparison with Manchu *niolhon* "the sixteenth day of the month" (a loanword from Khitan) or with MM *ǰirqo'an*, Classical Mongolian *ǰirɣuɣan*. In fact, CM *ǰirɣuɣan* cannot be used to reconstruct the Khitan form, since it is a compound from /ǰir/ meaning "two" and /ɣu/, a reduced form of *ɣurban* "three" (Janhunen 2003a: 16–17). Therefore, the Khitan root for

9. Also, regrettably, the book does not contain an index of cited Mongol, Jurchen and Manchu forms, which would have been very helpful.

© 2010. John Benjamins Publishing Company
All rights reserved

"six" is most probably preserved in the loanwords found in Manchu and Jurchen, though its exact shape is difficult to determine.

## 4. Language contact with Tibetan and Tangut

An almost untapped source of data on Khitan reconstruction can come from Tangut and Tibetan. Although K makes very welcome use of Tangut transcriptions of Chinese in his discussion of the phonology of the Chinese dialect Khitan was in contact with, Tangut and Tibetan data are not directly mentioned in his book.

However, at least one borrowing from Tibetan to Khitan, the word for "marmot", is known to exist, as Róna-Tas (2004) has shown. Another loanword is most probably the word for "country" ⟨g.úr⟩ (p. 89) from Tibetan *gur* meaning "tent". Manchu *gurun* comes from the same source, perhaps via Khitan. Given the dominance of the Tibetan Empire, it is unlikely that the loan could have gone in the opposite direction.

Khitan borrowings into Tibetan have not been detected up to now except for the transcription of the name "Khitan" Ge-tan / Ge-tang in the Pelliot Tibétain 1283, line 559–560 (see for instance Ligeti 1971: 169), a text from the VIIIth or IXth century. The Khitan autonym is reconstructed ⟨qid.ún⟩ in K (pp. 162–163). It is probable that this name was borrowed by way of another language (possibly Uighur) and that it is not a genuine loan from the Khitan language into Tibetan.

In Tangut, the only transcription of a Khitan word we have found up to now is also the name of the Khitan people *tɕhjɨ¹tã¹*. It is attested for instance in the following passage:

(1) *mjɨ² njaa² kjir¹ ljɪ¹ dʑjij¹ tɕhjɨ¹ tã¹ bji² dʑji lwẹ²*
Tangut brave go Khitan step walk slow
The Tangut go forth, brave and strong, the Khitan slowly walk step by step.
(*Fragments of Gold in the palm*, 13.5.6)[10]

The origin of this name is puzzling. The corresponding Chinese word for Khitan is 契丹 Qìdān, a word that had a velar initial in all known forms of Chinese before the thirteenth century (the first syllable was transcribed *khÿi* [k'ji] in 'Phags-pa, #199 p. 125); In loanwords, Chinese velars always correspond to Tangut velars, even before front vowels. For instance, in a personal name, Chinese 起 *qǐ* ('Phags-pa *khi* [k'i] #158 p. 120) is transcribed as Tangut *khji¹*. Therefore, *tɕhjɨ¹tã¹* cannot have been loaned from Chinese. The presence of nasality in the second syllable indicates that it is a recent loan, as nasal vowels normally only appear in Chinese

10. Kychanov (2008: 423–433) provides a Russian translation of this text.

© 2010. John Benjamins Publishing Company
All rights reserved

loanwords. Given the intimate contact between Tanguts and Khitans (the first Tangut emperor was married to a Khitan princess), it is possible that *tɕhjɨ¹tã¹* originates from a dialect of Khitan where dorsal initials had palatalized before front vowels. Incidentally, it is interesting to note that the Khitan demonstrative ⟨qi⟩ attested in the expression ⟨qi.po⟩ "that time" (p. 121) could correspond to the Tangut demonstrative *tɕhjɨ¹*, which almost only appears in the form *tɕhjɨ¹zjo̱²* "that time" in texts[11] and has no obvious cognates in Qiangic languages. This would present the same correspondence *tɕhjɨ*: *qi found in the name of the Khitan. However, it is not likely that the Tanguts would have only borrowed a demonstrative without borrowing much other vocabulary.

It is possible, though, that Khitan loanwords in Tangut are waiting to be discovered. We should look for Tangut words without a Qiangic etymology and resembling Mongolic forms. For instance, Tangut *njijr²* "face" reminds of MM *ni'ur* "face" and might have been borrowed from Khitan.[12]

For future work, an area where Khitan loanwords could be found are the lists of personal names, several hundreds of which are attested in the book *Mixed Characters* (Terent'ev-Katanskij & Sofronov 2002). The Tangut Imperial family originally had the family name 拓跋 Tuòbá (Middle Chinese *takbɛt*), a Xiānbēi (Para-Mongolic) name. Therefore, other Tangut names might also have a Para-Mongolic or even Khitan origin. For instance, one wonders whether the Tangut name *.jɨ²rjir²* (in Chinese known as 野利 *yělì*, among other things the clan name of the inventor of the Tangut script) is not a transcription of the Khitan word ⟨i.ri⟩ "name, title" (K. p. 108; this word has been loaned into Middle Korean, see Shimunek 2007: 75).

In conclusion, K's book is a welcome addition to Khitan studies, the first of its kind in a Western language. It succeeds in making Khitan data available to a broader public, and presents an important number of original contributions. In recent years, new Khitan texts have been unearthed and, among all the undeciphered languages, Khitan is probably the one that has the greater chance of being one day fully understood. K's work is a significant step forward towards this goal.

**11.** This character is also used, however, to transcribe an unrelated verbal prefix, see Jacques (To appear).

**12.** The similarity between Tangut *njijr²* and Mongolian *ni'ur* was independently noticed by Saiyinjiya Caidengduoerji, a student attending my Tangut class, and me.

© 2010. John Benjamins Publishing Company
All rights reserved

## References

Coblin, Weldon South. 2007. *A Handbook of 'Phags-pa Chinese*. Honolulu: University of Hawai'i Press.

Jacques, Guillaume. 2004. *Phonologie et morphologie du japhug (rGyalrong)*. PhD thesis, Université Paris VII — Denis Diderot.

Jacques, Guillaume. To appear. "The structure of the Tangut verb". *Journal of Chinese Linguistics*.

Janhunen, Juha. 2003a. "Proto-Mongolic". *The Mongolic Languages* ed. by Juha Janhunen, 1–29. London: Routledge.

Janhunen, Juha. 2003b. "Para-Mongolic". *The Mongolic Languages* ed. by Juha Janhunen, 391–402. London: Routledge.

Kara, György. 1987. "On the Khitan writing systems". *Mongolian Studies* X.19–24.

Kara, György. 2009. *Dictionary of Sonom Gara's Erdeni-yin Sang*. Leiden: Brill.

Kychanov, Jevgenij I. 2008. *Istorija Tangutskogo gosudarstva* (*A History of the Tangut State*). St. Petersburg: Fakul'tet filologii i iskusstv, Sankt-Peterburgskogo gosudarstvennogo universiteta.

Ligeti, Louis. 1971. "À propos du 'Rapport sur les rois demeurant dans le nord'". *Études Tibétaines dédiées à la mémoire de Marcelle Lalou* ed. by Adrien Macdonald, 166–189. Paris: Librairie d'Amérique et d'Orient, A. Maisonneuve.

Qīnggéěrtài, Liú Fèngzhù, Chén Nǎixióng, Yú Bǎolín, & Xíng Fùlǐ. 1985. *Qìdān xiǎozì yánjiū* (*A Study of Khitan Small Characters*). Běijīng: Zhōngguó Shèhuì kēxué chūbǎnshè.

Róna-Tas, András. 2004. "A Khitan Word for marmot". *Acta Orientalia Academiae Scientiarum Hungaricae* 57:3.27–29.

Shimunek, Andrew. 2007. *Towards a Reconstruction of the Kitan Language, with Notes on Northern Late Middle Chinese Phonology*. MA Thesis, Bloomington.

Terent'ev-Katanskij, Anatolij Pavlovich & Mixail Viktorovič Sofronov. 2002. *Smeshannie znaki* (*Mixed Characters*). Moskva: Vostochnaja Literatura.

Takeuchi, Yasunori. 2008. *Kittan shouji de hyouki sareta kanji on kara mita Kittango on'in taikei no kenkyuu* (*A Study of the Phonological System of Khitan from the Pronunciation of the Chinese Forms Transcribed in Khitan Small Characters*). MA thesis, University of Kyoto, Department of Literature.

Vovin, Alexander. 2007. "Once again on the Tabgac Language". *Mongolian Studies* XXIX.191–206.

### *Reviewer's address*

Guillaume Jacques
CNRS (CRLAO) — INALCO
49bis Avenue de la Belle Gabrielle
75012 Paris, France

rgyalrongskad@gmail.com

© 2010. John Benjamins Publishing Company
All rights reserved