A MODEST PROPOSAL ON THE DECIPHERMENT OF THE KHITAN-JURCHEN BILINGUAL TEXT OF 1134 (THE LANGJUN INSCRIPTION)[1]

Alexander Vovin

University of Hawai'i at Mānoa

With the publication of the *Qidan xiaoji yanjiu* (契丹小字研究), a funfdamental volume on Khitan script and language by the Khitan research group (Qingge'ertai et al. 1985), the tide of the decipherment og Khitan seemed to have been turned, as this publication brought a real revolutionary break through in the long history of failed attempts to break the Khitan code: a number of Khitan signs were successfully identified, which made possible the reading of at least some portions of the Khitan epigraphic material. Although it seemed for a while that the final decipherment was close at hand, the tide ebbed. Today, as twenty two years ago, we still cannot read completely any Khitan inscription, as many more signs remain unknown and we still have only vague idea about the Khitan language itself. And this is in spite of the fact that several important works on Khitan appeared in China and other countries during this period: (Kara 1987), (Qingge'ertai 1992), (Chinggeltei 1997a and 1997b), (Kane 2006), and (Shimunek 2007).

I believe there are several reasons for a relatively slow progress in the decipherment of Khitan since 1985. First, it is the nature of the language itself: although it is Mongolic, it appears to be more distantly related to any extant Mongolic language than any extant languages to each other. Judging by several words that has been identified, Khitan looks like quite a different (and in many aspects rather phonologically innovative) offshoot of Mongolic. Therefore, the definition of the Khitan language as a Para-Mongolic rather than a Mongolic proper (Janhunen 2003: 392ff) seems to have a good foundation. This state of affairs makes it difficult to use data from the extant Mongolic languages for both initial guesses and final identification of the Khitan lexical and grammatical elements.

Second, phonetic identification of Khitan signs that were used to transcribe Chinese titles etc. and which served as a key to the decipherment on the first stage of work by the Khitan research group (Qingge'ertai et al. 1978), (Qinge'ertai et al. is based either on the outdated versions of the Middle Chinese reconstruction, such as Karlgren, Li Fang-Kuei, Li Rong, and others, or, even more anachronistically, on modern Mandarin readings of characters transcribed. This certainly creates additional difficulties for recognizing even those Khitan words and morphemic elements that are potentially similar to the same elements in other Mongolic languages. The reading of Chinese characters changed rather drastically during the last one thousand years, and it would be impossible to identify a Khitan word if it is read incorrectly on the basis of the anachronistic Chinese readings.

Third, there seems to be a poor recognition of the polyphonic principle in the Khitan writing system: one and the same sign may have different phonetic values, while the same phonetic value can be represented by different signs. The polyphonic principle is typical for all syllabic East Asian scripts ultimately derived from Chinese. A good example how the polyphonic principle works may be given on the basis of Old Japanese script *man'yōgana*. Thus, a Chinese character 木 'tree' may be used as semantographic sign, representing the Old Japanese word *kï* 'tree.' Since the Early Middle Chinese reading of this character was [məwk], it was also used as a syllabic sign representing syllable /mo/, as in e.g. 木理 /mori/ 'forest.' In addition, the Old Japanese reading /kï/ could have also been used phonetically, as in, e.g. 木之 /kïsi/ 'shore, coast.' Finally, not only syllables /kï/ and /mo/, but eventually every Old Japanese syllable could have been written by a number of other characters, e.g. /kï/: 奇, 貴, 紀, etc., /mo/: 茂, 望, 聞, etc., /ri/: 里, 利, 隣, etc., /si/: 志, 思, 斯, etc. Without any doubt, Khitan writing system enjoyed the same level of complexity, but in the chart of Khitan signs with known readings compiled by the Khitan research group we find very few characters with polyphonic readings (Qinge'ertai et al. 1985: 152-53). As a matter of fact, the only part of the polyphonic principle that is recognized by the Khitan research group seems to be the phenomenon that one and the same syllable can be written by the different signs. However, there is no recognition that one and the same sign can also have different phonetic values.

Fourth, until recently there was no recognition that the one of the most powerful keys in deciphering Khitan lies in the recognition of the fact that both Jurchen and Manchu are inundated with Khitan loanwords. The recent article by Daniel Kane demonstrates excellently how these loanwords can be used for successful

[1] This article is based on the presentation at the 43rd PIAC meeting, Chateau Pieterscheim, Belgium, September 3-8, 2000. It seems that the proceedings volume of that conference where the first variant of this article was supposed to be published never appeared.

identification of phonetic shapes of Khitan words and consequent decipherment of several new Khitan signs (Kane 2006).

Fifth, it appears that the major attention of the Khitan research group was turned towards the analysis of the Khitan transcriptions of Chinese titles etc. in the inscriptions, while the *Langjun* bilingual inscription[2] itself has not been used to a full extent. It is very well known from the history of decipherment how important bilingual texts are.

Sixth, it seems that the Khitan research group treated Khitan writing system as a mixed semantographic-alphabetic-syllabic script: in addition to a majority of purely syllabic signs with V, CV, VC, and CVC structures one finds a number of signs representing just consonants alone, such as #28[3] 戈 [š], #149-50 子-刘 [č], #214-215 𘖵-久 [t], #261 𘓱 [l], #311 𘐷 [p], etc. This seems hardly possible on the basis of the history of writing in general and in East Asia in particular. Mixed alphabetic-syllabic scripts are not unknown, e.g. modern Cyrillic one, with syllables /ya/, /yo/, /yu/ spelled by one letter each: я, ё, ю, but this is a later development from original purely alphabetic Old Church Slavonic script. In the East Asian historical context such a unique development in Khitan seems unprecedented: it is quite apparent that in the Jurchen script that was influenced by Khitan to a certain degree, there are no alphabetic signs -- only syllabic ones. The real alphabetic writing in East Asia was virtually unknown until mid-15th century, when the Korean alphabet Hankul was invented, although few isolated syllable-final signs in the Old Korean writing *hyangchal* could be potentially interpreted as alphabetic. However, this is not quite clear, and since this problem lies outside the scope of this paper I will not discuss it here. I suppose that Khitan signs reconstructed by Khitan research group are better off for a moment as syllabic CV signs with an unspecified vowel that I will denote below as (V). The usage of a consonant with an unspecified vowel has at least a historical precedent: Egyptian syllabic writing (Gelb 1963).

While it is not possible to deal with the first problem at the moment, I will try to demonstrate how proper accounting for other points outlined above, can lead to a further progress in the decipherment of the Khitan texts and improve our understanding of the Khitan language. I will concentrate in this article on the famous bilingual Langjun inscription of 1134 A.D., the bacl lines of which were the basis for the revolutionary breakthrough done by the Khitan research group in late seventies and early eighties. Rather than dealing with the Langjun inscription as a whole, which is a major undertaking, I will deal in this short article with the middle portion of the line four of the main text in the this inscription, namely comprising written block 4-8 to 4-13.[4] The Langjun inscription is a peculiar one among other Khitan inscriptions, because it was inscribed not by the Khitan themselves, but by the Jurchen after their conquest of Liao empire in 1125. It commemorates a visit by a younger brother of the Jurchen Jin (金) dynasty emperor, who had a title of Langjun (郎君) to a pavilion Tangqianling (唐乾陵) in Liyang (醴陽) prefecture.

In the following discussion I will use Late Middle Chinese (LMC) reconstruction by Pulleyblank (1991), which is much more up-to-date than reconstructions by Bernhard Karlgren, Li Fang-kuei, Li Rong, and others. Nevertheless, one sufficient correction must be born in mind when utilizing LMC readings: as recently convincingly demonstrated by Shimunek (2006), the Northern LMC dialect that was used to transcribe Khitan probably already had no final stops *-p, -t,* and *-k*.

Since we are dealing with a bilingual text, we should first of all pay attention to the Chinese equivalent of the text we are dealing with. The Chinese text that roughly corresponds to the Khitan written blocks 4-8 to 4-13 is: 與醴陽太守酣飲而歸 '[Langjun] drank heavily with the governor of Liyang and [then he] returned.' Consequently, our task is essentially to identify the bits and pieces of the Khitan text that correspond to it. The rest of this article is dedicated to the attempt of this identification.

As an 'Altaic' language, Khitan is likely to have an SOV word-order. If this is true, the last block 4-13 𘐷币𘓱𘖵 is likely to be a verb corresponding to Chinese 歸 'to return, to come/go back.' A similar block 𘐷币𘓱𘍿 is found in *Xu wang muzhi* inscription 14-14, also apparently at the end of a sentence. Only three signs in the block 4-13 of the Langjun inscription have been established by the Khitan research group: #295 𘐷 /ph/f(V)/, #261 𘓱 /l(V))/, and. However, sign #295 𘐷 transcribes initial in LMC 平 /pɦiaŋ/, 妃 /phuaj/, 奉 /fɦyawŋ ~ fɦəwŋ/. In two cases out of three it transcribes LMC initial *pɦ-/fɦ-* with voiceless onset but voiced offset. I surmise that it was voicing and not aspiration that was distinctive, and therefore I believe that the sign #295 𘐷 renders /b(V)/ rather than /p(V)/. If this solution is true, there is a likely Mongolic cognate for the block 4-13 𘐷币𘓱𘖵 is spite of the fact that the phonetic value of the sign #16 币 is unknown: Middle Mongolian (MM) and Written Mongolian (WM) *bučа-* 'to come/go back, return.' Therefore, it is possible to

[2] The full name of the inscription is: *Dà Jīn Huáng Dì Dūtóng Jīnglüè Lángjūn Xíngjì* (大金皇弟都統經略郎君行記).

[3] The enumeration of the Khitan signs follows (Qinge'ertai et al. 1985).

[4] The Khitan writing system used to combine written signs into blocks that comprised either semantographic sign alone or several syllabic signs, or a semantographic sign plus syllabic sign(s).

surmise that the phonetic value of the sign #16  is /ča/. Ji Shi also interprets the block 4-13 as Mongolic /bučala/ 'returned,' although he takes the last sign to be #144 , and not #143 (Ji 1996: 268). The last sign is written in a cursive style, so it can be either #143 or #144 . However, Qingge'ertai et al. read #144 as /da/ (1985: 152), and Ji Shi himself reads it as /in/ (1996: 442). Neither /da/ nor /in/ could possibly correspond to /a/. Therefore, it is more likely that #143 read by Ji Shi as /γ^a/ (1996: 442) is actually used here. Consequently, I believe that the remaining portion of the block 4-13 should be read as /l(u)-γa/. There are three reasons for that. First, Khitan sign # 261 /l(V)/ is interpreted by Khitan research group as alphabetic sign /l/, which transcribes MC initial /l-/. However, a combination of both syllabic and alphabetic signs in the same system of writing is typologically uncommon and historically unverified, as was already noted above. It is more likely that Khitan transcription was modeled after Chinese *fanqie* transcriptional method, with the sign # 261 /l(V)/, being universally used for Chinese syllables with the initial /l-/. Second, it is further possible that for the indication of the lack of the vowel after /l-/, a sign with a high vowel was used (cf. typologically the same case in Japanese and Korean writing systems, where high vowel is used to indicate an absence of the vowel in a foreign word). Third, read as /l(u)-γa/ nicely fits with the Mongolic perfective tense marker: MM *-lu'ai/-lü'ei, ~ -lu'a/-lü'e*, WM *-luγa/-lüge*, (preclassical *-luγai/-lügei*; *-laγa/-lege*), etc.

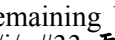
The block 4-8 (consisting of #332 with an unknown reading and #254 /t(V)/) has been identified as 'governor' by the Khitan Research Group, therefore blocks 4-9 to 4-13 must correspond to Chinese與醴陽太守酣飲而歸 '[prince] drank heavily [lit. intoxicated himself] with (governor [of Liyang] = 4-8) and then returned.' I have already identified above the block 4-13 as the final predicate 'returned.' Thus, the blocks 4-9 to 4-12 correspond to the rest of the phrase.

The block 4-9 immediately following the block 4-8 'governor' can functionally correspond only either to a comitative case marker or to a postposition with the comitative meaning. Since the readings of the signs constituting this block are now known: #101 /təu/ (Ji 1996: 440), #97 /ur/ (Ji 1996: 657), (Kane 2006: 123-24), it is not possible to compare the block 4-9 with MM comitative case marker *-lu'a/-lü'e*. The resulting /təwur/ is likely to be a Khitan postposition with a comitative function. Given a tendency of *-b- > *-w-* in Khitan (Kane 2006: 124), it is hard not to notice the striking similarity of Khitan *təwur* < *təbur 'with' to MK *təpïrə* [təbïrə] 'with.' Unless this is a chance resemblance, the postposition must have been borrowed by Khitan or its immediate ancestor from either Koguryo or Parhae dialects of Old Korean, since the MK postposition *təpïlə* 'with' is morphologically transparent: it represents a grammaticalization of the converb in *-a/-ə* from the verb *təpïr-* 'to share, to accompany.'

The block 4-11 contains two signs with known readings: sign #270 /em/ and sign #162 /č(V)/. I tentatively interpret it as *em-či 'drink-GER.' Although there is no verb *em-* in extant Mongolic languages, there is a WM *umdan* 'drink' (n.), and the root is abundantly attested in Tungusic with interesting reflexes: Ewenki *um-*, Solon *emu- ~ imu-*, Neghidal *om-*, Oroch *imi-*, Udihe *umi-*, Ulcha *umi-*, Orok *umi-*, Nanai *omi-*, Manchu *omi-* (Cincius 1977: 266b). While the vowels of the first syllable show regular correspondences in Ewenki, Neghidal, Udihe, Ulcha, Orok, and Nanai, the reflexes in Solon, Oroch, and Manchu are irregular. Thus, the Solon, Oroch, and Manchu forms cannot be traced to the common Tungusic archetype. While Manchu *omi-* can be explained as a inter-Tungusic loan from a Nanai-like form *omi-*, Solon and Oroch forms cannot be explained along the same lines. Consequently, these forms are likely to be loans from Para-Mongolic, suppoting therefore the tentative Khitan *em- 'to drink.' The sign #162 probably could be safely identified with MM imperfect gerund marker *-ju/-ču/-ji/-či*. It is necessary to add that this sign frequently occurs in the last position in the block, e.g. another example in the same inscription is found in the block 4-2 , consisting of #319 /ko/ and #162 /č(V)/.

The block 4-10 also appears as the block 59-21 of the *Xu wang muzhi* inscription. The first sign #342 occurs in Khitan texts exclusively in initial position. Therefore, it likely represents a semantogram, and not a phonetic sign. Sign #312 is interpreted by Khitan reserach group as /p/. I believe it rather represents /b(V)/, as it was used to transcribe only two types of LMC initials: voiceless aspirated with voiced offset /pɦ-/ (twice) and voiceless unaspirated /p-/ (three times) (Qinge'ertai et al. 1985: 103-04), but never voiceless aspirated with voiceless offset /ph-/. The position of this block between the comitative case marker and the verb suggests that the block 4-10 represents the object of the verb 'to drink,' and thus it probably means 'wine,' 'liquor.' I suppose that the sign #342 is a logogram meaning 'to distill wine,'[5] and sign #312 represents a nominalizer *-b(V)*. This *-b-* must be from an earlier *-p-, because the original *-b- lenites to *-w-* in Khitan. Since Proto-Mongolic *-p- is reflected as *-'-* in MM and as *-γ-* in WM, this Khitan nominalizer *-b(V)* < *-p(V) may be comparable with MM nominalizer *-'a* (WM *-γa*).

[5] A possible candidate is WM *nere-* 'to distill wine,' although there is a potential problem, because this word is not attested in MM and in early (pre-17th century) WM texts.

The remaining block 4-12  consists of signs that all have known readings: #340  /x(V)/, #339  /i/, #33  /is/, #372  /u/, and #262  /uei/ ~ /ui/ or /ul/.[6] This block can only correspond to Chinese 'to get drunk, to get intoxicated,' but unfortunately there is no Mongolic word with the same phonetic shape and meaning. There is, however, a Manchu verb *hisala-* 'to make a libation,' that has no cognates in other Tungusic languages, and given also the absence of the similar word in Mongolic and Korean, it is quite likely that the Manchu word represents an old loan from Khitan. The internal structure of this verb seems to be quite clear: it must be a denominal verb, derived with a Manchu verbalizer *-la-* from a noun *hisa- 'libation,' which unfortunately does not present itself. While the difference in vocalism of the second syllable is difficult to explain at this point,[7] the block probably should be read as /xisu-wei/ < *xisu-bei, a preterite form of the verb *xisu- 'to make a libation, to imbibe,' which shows the *-b- > *-w-* lenition typical for Khitan that was already mentioned below.

References:

Chinggeltei 1997b. 'Кидан хэл хийгед монгол хэлний арван хоёр жил [Twelve year cycle in the Khitan and Mongol languages].' *Althai hakpo* 7: 156-60.

Cincius Vera I., ed. 1975-77. *Сравнительный словарь тунгусо-маньчжурских языков [A comparative dictionary of the Manchu-tungusic languages]*. vol. 1 & 2.Leningrad: Nauka.

Gelb, Ignace J. 1963. *History of writing*. Chicago: University of Chicago Press.

Janhunen, Juha 2003. 'Para-Mongolic.' In: *The Mongolic Languages*. Ed. by Juha Janhunen, pp. 391-402. London: Routledge.

Ji, Shi 1996. *Mi lin wen jing (谜林问径) [Asking for a path in an enigmatic forest]*. Dalian: Liaoning minzu chubanshe.

Kane, Daniel 1989. *The Sino-Jurchen Vocabulary of the Bureau of Interpreters*. Uralic and Altaic Studies, vol. 153. Bloomington: Research Institute for Inner Asian Studies.

Kane, Daniel 2006. 'Khitan and Jurchen.' In: *Tumen jalafun jecen akū: Manchu Studies in Honor of Giovanni Stary*. Ed. by Alessandra Pozzi, Juha A. Janhunen and Michael Weiers. Wiesbaden: Harrassowitz.

Kara, György. 1987. 'On the Khitan Writing Systems.' *Mongolian Studies* X: 19-24.

Nishida, Tatsuo 1982. 'Kittan moji kaidoku no sin tenkai (契丹文字解読の新展開) [A new development in the decipherment of the Khitan script].' In: *Ajia no mikaidoku moji (アジアの未解読文字) [Undeciphered scripts of Asia]*. Tokyo: Taishūkan.

Nishida, Tatsuo 1992. 'Guanyu qidan xiaozi yanjiu zhong jibenxing wenti (关于契丹小字研究中的基本性问题) [On basic problems in the study of the Khitan Small Script]. *Minzu yuwen* 1992.2: 10-13.

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

Qingge'ertai [Chinggeltei], Yu Baolin, Chen Naixiong, Liu Fengzhu, Xin Fuli 1978. 'Qidan xiaozi jiedu xin tan (契丹小字解读新探) [New research on the decipherment of the Khitan Small Script].' *Kaogu xuebao* 3: 353-87.

Qingge'ertai [Chinggeltei], Yu Baolin, Chen Naixiong, Liu Fengzhu, Xin Fuli 1985. *Qidan xiaoji yanjiu (契丹小字研究) [A Study of the Khitan Small Script]*. Beijing: Zhonguo shehui kexue chubanshe.

Qingge'ertai [Chinggeltei] 1992. 'Qidan xiaozi zhong dongci fujia chengfen (契丹小字中的动词附加成分) [Verbal affixes in the Khitan Small script].' *Minzu yuwen* 1992.2: 1-9.

Qingge'ertai [Chinggeltei] 1997a. 'Qidan yu shuci ji qidan xiaozi pindu fa (契丹语数词及契丹小字拼读法) [Khitan numerals and the method of the Khitan language decipherment]. *Althai hakpo* 7: 143-52.

Shimunek, Andrew 2007. *Towards a Reconstruction of the Kitan Language, with Notes on Northern Late Middle Chinese Phonology*. University of Indiana unpublished M.A. thesis.

[6] See Kane (2006: 131) on the polyphonic nature of this character.

[7] Cf., however, Jurchen *sa'u-* in *nure sa'u* 'to pour a libation of wine' (Kane 1989: 347) that if related to Manchu *hisa-la-* can potentially explain the discrepancy of vocalism.

APPENDIX: *Dà Jīn Huáng Dì Dūtóng Jīnglüè Lángjūn Xíngjì* 大金皇弟都統經略郎君行記 inscription (reproduced from Qingge'ertai et al (1985: plate 18))

