Notes on "Khitan Transcriptions Revisited"

Marc Hideo Miyake
University of Hawai'i at Mānoa

This short piece is best read as a series of footnotes to "Khitan Transcriptions Revisited". I have decided not to revise the original abstract, for I believe that the numerous errors and dubious hypotheses therein are worth keeping, if only to serve as targets of criticism here. Besides, yesterday's bad idea might turn out to be a good idea tomorrow.

Notes on the introduction

Unfortunately, my exposure to works on Khitan has been principally limited to the Menges/Wittfogel/Fêng discussion in Wittfogel and Fêng (1949), the discussion in Tamura Jitsuzô's *Keiryô*, Daniel Kane's *The Sino-Jurchen vocabulary of the Bureau of Interpreters* (1989), Nishida's articles in *Gengo* (1980, 1981), and Zhen Naixiong's "Qidan zi yanjiu shulüe" in the *Festschrift for Nishida Tatsuo* (1988)--all the works known to me to be available in the University of Hawaii library. Hence I have not yet seen the vast bulk of work on Khitan, so my pronouncements such as "scholars seem to assume that the Chinese of the time [i.e. the Liao Dynasty period] was not unlike Old Mandarin..." may not turn out to be true after all.

However, assuming that Kane and Nishida are not distorting the claims of the Chinese research I have yet to see first-hand, I am not impressed by what I see. There is surprisingly no discussion at all about what the Chinese of the time was like. In Nishida's 1980 series of articles on Khitan, Mandarin-like forms (*zhi*) coexist side by side with what looks like Old Mandarin or even Middle Chinese (i.e. forms ending in *-m* or with *rusheng* stops such as *tshak*, oddly given as a gloss for the *pingsheng* graph 奢, Late Middle Chinese *śa, not *tshak!). I would think that Khitan decipherment would first begin with an intensive analysis of the sinographic transcriptions, such as those found in the *Guoyu jie* of the *Liao shi*, but this doesn't seem to be the case. From Kane and Nishida, I get the impression that not much thought has been put into developing a coherent reconstruction of Liao Dynasty period Chinese. "Qidan zi yanjiu shulüe" disappointingly makes almost no reference to Chinese glosses at all, though it does provide a long list of possible 'small' script grapheme readings. If my worst fears are true, then the reconstruction of Khitan is based on incoherent notions of Liao Dynasty Chinese.

Notes on the reconstruction of Late Middle Chinese (LMC)

Was LMC really "the language of the *Yunjing* rhyme table[sic; correct to "tables"]"? I'm not so sure, as the history of the *Yunjing*, to the best of my knowledge, is rather vague--no one seems to know exactly when or where it was written. Perhaps a better definition of LMC for my purposes here would be the "Liao Dynasty variety of Chinese underlying the sinographic glosses of Khitan words." I assume (perhaps wrongly) that this LMC is close to, if not the descendant of, the northeastern variety of Middle Chinese that underlies Sino-Korean.

I brought up Tibetan and Turkic transcriptions here because like Sino-Korean, they confirm a final *-r corresponding to Early Middle Chinese *-t in northern LMC (but not in southern LMC--literary strata in southern Chinese languages and Sino-Vietnamese indicate *-t there).

I did not mention Sino-Japanese at all, since I believe the Kan-on stratum (and of course the Go-on strata) is older than most strata of Sino-Korean, which in turn are probably older than the language underlying the sinographic transcriptions of Khitan. The so-called 'Tô-sô-on' of Sino-Japanese are too new (dating from Southern Song), too few, and from too far away (Hangzhou as opposed to northern China) to be of any relevance.

Of course, there is a possibility that the authors of the *Liao shi* etc. spelled Khitan words phonetically according to the Chinese of their time (i.e. Old Mandarin), but I think that it is more probable that the words in the *Liao shi* etc. are traditional spellings dating back to Liao times, if not earlier in a few cases. If there has been any philological work done on the first attested appearances of Khitan words in Chinese texts (outside the dynastic histories penned long after Liao times), I would like to know about it.

The devoicing of the 'muddy' initials is clear from Kan-on, based on a variety of LMC predating that of Sino-Korean and northeastern Liao Dynasty period Chinese. Furthermore, there are apparent cases of 'muddy' and 'second clear' initials being used to write the same Khitan word--步 LMC *pɦo and 普 LMC *pho both appear as the second syllable of the tribal title (?) Khitan *mapo?.

My interpretation of the four *deng*s, particularly the second and third *deng*s, is admittedly rather unusual, and I cannot give anything approaching a complete defense of it here. The notion of a second *deng* with rhotacized initials was inspired by the fact that the retroflex initials are exclusive to the 2nd *deng* in Middle Chinese. I do not put much faith in this view of the second *deng*. (See below for an alternate view.) The notion of a third *deng* with velarized initials was inspired by the fact that Sino-Korean has *-u-* and *-e-* in third *deng* rhymes contrasting with *-i/y-* and *-ye-* in fourth *deng* rhymes--Sino-Korean represented the contrast between LMC *$k^{ɯ}$i (*$k^{ɯ}$ɨ?; *deng* III) and *k^{y}i (*deng* IV) as *kuy* (*[kɯi] or *[kɨy]?) and *ki* (*[ki]).

I formerly advocated a view of the four *deng*s which I called the 'А-Я-Ы-И hypothesis' (via correlating reconstructed LMC vowels [and palatalizations of preceding consonants] with equivalent Russian letters):

	initial	vowel		
I:	non-palatalized	back	(*ɑ, *o, *u)	"А, О, У"
II:	palatalized	low front	(*a)	"Я"
III:	non-palatalized or velarized	high~mid central~front	(*ɨ, *e)	"Ы, Э"
IV:	palatalized	high~mid	(*i, *e, *o, *u)	"И, Е, Ё, Ю"

Although now obsolete, this system (and its successor) both have velarized initials (or should it be velar semivowels?) for the third *deng* contrasting with palatalized initials (or a palatal seminvowel) for the fourth *deng*. This hypothesis feeds my hypothesis about how the sinographic transcriptions correlated with Khitan vowels, as I will explain below.

The proposed eight-vowel system frequently correlates directly with Sino-Korean and Sino-Vietnamese (given here in phonetic transcription, not *quôc ngu'* orthography):

LMC	SK	SV	LMC	SK	SV	LMC	SK	SV	LMC	SK	SV
*a	*a*	a	*i	*i*	i	*ɨ	*u, o*	ɨ	*u	*wu*	u
*ɑ	*a*	a	*e	*e*	e	*ə	*o, a*	ə, a	*o	*wo*	o

For reasons I can only briefly go into here, I believe that the Korean vowel shift (if it ever existed at all) must have taken place before the importation of most (or all) of the Sino-Korean strata. (If a Korean vowel shift took place *after* the borrowing of Sino-Korean, then why do the Sino-Korean vowels in their post-shift pronunciation match up so well with Sino-Vietnamese, Sino-Japanese, and the literary pronunciations in Amoy Min, unless *three independent vowel shifts in the same direction* took place in Vietnamese, Japanese, and Amoy Min, three genetically unrelated and geographically distant languages, after borrowing from LMC, a highly unlikely possibility?) Hence I believe that the phonetic values of the Korean vowels at the time of importation were close to those of modern Korean, except that *alay a* was [ə] (appropriate for a 'reduced vowel') and *e* was [e].

Currently, I think that the velar and labial *rusheng* endings in 10th century northeastern Chinese were voiced fricatives ('gamma' and 'beta', not 'chi' and 'phi'). However, since I think LMC *rusheng* codas could represent either voiced or voiceless Khitan codas (table 5, page 4), the voicing of the LMC *rusheng* codas doesn't matter.

Notes on table 1

Any similarities between the reconstructed finals on page 2 and those of Sino-Korean are intentional. Note, however, that the SK *-u-* of *deng* III and SK *-y-* of *deng* IV correlate with the velarization and palatalization of the initials listed on table 2, page 3, not with anything on table 1, page 2.

Notes on table 2

*fi in LMC forms is Pulleyblank's notation (in turn taken from Chao Yuan Ren) for murmur or voiced aspiration--his interpretation of the 'muddiness' of the LMC descendants of earlier voiced initials.

I have assumed that LMC voiceless aspirated : Khitan voiceless and LMC voiceless unaspirated : Khitan voiced, on the model of the pattern found in Chinese transcriptions of Jurchen. This could turn out to be utterly wrong, particularly if it turns out that Khitan had a three-way distinction between, say, *p : *ph : *b. That doesn't seem typologically likely, though. It may, however, be possible, that LMC aspirates sometimes represent Khitan fricatives (i.e. LMC *qh : Khitan *x). But Khitan didn't seem to have any labial fricatives.

I have assumed that Khitan backed */k/ to *[q] before back vowels, for two reasons. The first is typological (cf. other 'Altaic'-type languages); the second is the fact that the transcription sinographs used to write Sanskrit velar initial-back vowel combinations such as 伽 LMC *kfi^{ɯ}a were never used in Khitan transcription, suggesting that Khitan did not have *ka; we find 可 LMC *qhɑ instead, presumably for Khitan *qa. But since the Sanskrit transcription sinographs were probably 'off limits' for writing non-Sanskrit, particularly 'barbarian' languages, so the absence of sinographs like 伽 may actually signify nothing.

Did Khitan have an backed allophone of */g/ before back vowels? I really don't know, but the fact that I find sinographs like 古 LMC *qo (Khitan *go) but not 居 LMC *k^{ɯ}o seems to indicate allophonic backing of Khitan */g/. Perhaps I should have written (*[G] before back vowels) in the table entry for LMC *k.

The retroflex stops probably already affricated by the tenth century, although they are still reflected as stops in Sino-Korean (currently palatals in modern Korean pronunciation) and in Sino-Vietnamese (*tr* in *quôc ngu'*). Late Tang and Five Dynasties Tibetan transcriptions have palatals, not *-r-* clusters, for these initials. Hence LMC retroflex stops *ʈ, *ʈh, *ʈfi should probably be rewritten as LMC *ʈʂ, *ʈʂh, *ʈʂfi, corresponding to Khitan *j, *c, *c, not Khitan *d, *t, *t. (But then how were Khitan syllables like *ti written? Since there were no syllables like *thi or *ti in LMC, the choices would be affricate-initial LMC *chi, *ʈʂhi, *tsh^{y}i or stop-glide cluster initial LMC *thwi.)

I seriously doubt that Khitan had a three-way ts-ʈʂ-c distinction, at least in native words. (Who knows how many marginal phonemes were in 'Sino-Khitan'?) Hence, I conflated the LMC affricates, considering them all to either represent Khitan *c or *j. When analyzing the Khitan writing system, however, one should still be on the lookout for any graphemes making a partial or total distinction between dental, retroflex, and palatal affricates, especially in small script compound graphs representing Chinese loans.

LMC *h and *hfi could be rewritten as *x and *xfi; there is no hard evidence (other than the current velar pronunciation of Mandarin) to reconstruct these as glottals or velars. I currently prefer reconstructing them as velars.

Khitan might have had an *h as well as an *x (or just an *h and no *x) corresponding to LMC *h and *hfi or even *kh and *kfi, but it's difficult to tell from the sinographic transcriptions alone.

Notes on table 4

This table is highly conjectural. The hypotheses of varying quality underlying it are:

--*deng* III LMC *i with velarized initials corresponds to Khitan *ï (central or back unrounded).

--LMC *ə and *deng* III LMC *e and o with velarized initials correspond to Khitan *ö (front mid rounded) due to the acoustic similarity of unrounded back or central vowels (here, not a vowel, but the LMC *-^{ɯ}- of *deng* III) with front rounded vowels.

--*deng* III LMC *u with velarized initials corresponds to Khitan *ü (front high rounded) for the acoustic reasons given above.

--since velarization in *deng* III applies only to [+grave] initials and since a preceding velarized initial is the key to determining whether an LMC *i corresponds to Khitan *i or *ï, LMC *i after [-grave] initials lacking velarization corresponds to indeterminate Khitan *I (*i or *ï?). Similar logic motivates reconstructing *E (*e or *ö?), *U (*u or *ü?), *O (*o or *ö?) after [-grave] initials.

--*deng* IV LMC *o, *u with palatalized initials correspond to Khitan *ö, *ü.

I have totally ignored the question of how to interpret LMC vowel-glide sequences like *-iw which might correspond to Khitan *-iu, *-iw, or even *-ü. Tentatively, by analogy with Sino-Japanese, I would interpret LMC vowel-glide sequences as representing Khitan vowel-vowel sequences; see 'hare', table 8, page 6, for an example.

Notes on table 5

LMC *-r should correspond to Khitan *-R, which may be *-t, *-d as well as *-l, *-r.

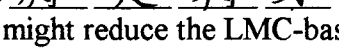
Notes on tables 6 and 7

These tables are merely tabulations of the guesses in previous tables and should not be taken too seriously.

Notes on table 8

The shared small script element  *-a? in 'prince' and 'hare' seems to suggest *śaali and *taulla, but my reconstructions are based purely on the LMC data alone.

The reconstruction for 'dog' should be corrected to *NyeRXaR with a -y-.

The gloss *chi shi de ben* for 'filial piety' is not in the *Liao shi guoyu jie* itself; I took it from the correction given in the *Liao shi shiyi, juan* 23. The *Liao shi guoyu jie* gives  *de shi de ben* LMC *təx śir təx pon (Khitan *döXśIRdöXbon?). Analysis of the script might reduce the LMC-based construction down to something with fewer syllables.

Notes on the conclusion

Parts of the reconstruction should be segmented: l- -aŋ, d- -Uŋ.

The 'Sino-Khitan' proposed here is intended to be another branch of Sinoxenic: a large-scale stratum or strata of Chinese borrowings with highly regular correspondences to Chinese forms. Although here I hypothesize that Sino-Khitan may be based on LMC pronunciation (with small script spellings of Chinese words perhaps fixed after the invention of the small script in 925), it might be possible that Sino-Khitan could be based on an even earlier stage of Chinese. If I accept Janhunen's (1994) hypothesis concerning the Bohai origin of Manchurian scripts (i.e. Khitan and Jurchen large scripts), then this "even earlier stage of Chinese" might be a Khitan borrowing of Sino-Bohai, which in turn would be derived from Sino-Kokwulye.