Once Again on the Tabgač Language

ALEXANDER VOVIN MANOA

It appears that scholars still cannot agree whether Tabγač, the language of the Toba-Wei Empire in Northern China (386–525 AD) was Mongolic or Turkic. Several factors make the identification difficult: we do not have any Tabγač texts that survived, and our only materials are glosses scattered in the Chinese sources. In addition, these glosses are in the Chinese transcription that presents an insurmountable obstacle for anyone not familiar with the Chinese historical phonology. As a matter of fact, there were numerous attempts in the past to interpret the reading of these glosses through the prism of either outdated Chinese reconstructions or even modern Mandarin characters readings. This, of course, constitutes an anachronism that can hardly lead to any positive results.

Like all Inner Asian empires the population of the Toba-Wei was certainly multiethnic and multilingual, so the problem of the linguistic affiliation of the Tabγač language essentially boils down to the identification of the language of the dominant group of the nomadic population that has founded the empire. It seems that Shiratori Kurakichi was the first scholar to claim that the Tabγač language was Mongolian (1912, 1913). However, other scholars of the first half of 20th century were generally in favor of viewing Tabγač as Turkic speaking (Pelliot 1925: 255), (Boodberg 1936), (Bazin 1950). That clearly was a dominant opinion until 1970s, and Clauson even tried to connect the Old Turkic ethnic name *Tabγač* with the modern ethnic name *Čăvăš* 'Chuvash' (Clauson 1962: 38). Although phonetically this is plausible, and modern Chuvash is without any doubt a Turkic language, unfortunately this identification does not reveal to us the actual linguistic affiliation of the Tabγač language: by the same token we would be able to claim that Russians speak the language of Vikings and French the language of Francs.

A new support for Shiratori's point of view was introduced by Lajos Ligeti, who argued that both Xian-bei and its Tabγač dialect were in fact Mongolic languages (1970). This view was adopted by other scholars, such as Doerfer (1992, 1993), and Janhunen (2003).

Ligeti's seminal article is an outstanding contribution that is still important today. It suffers, however, from one important limitation that largely was not the fault of the great Hungarian scholar, but was due to the general level of Chinese historical phonology at that time. The main problem is that Ligeti used the now (and even by then) outdated reconstruction of Middle Chinese by Karlgren. Should he opt to use more extensively the work of Pulleyblank (1962), the results might be much better, but in any case he could not possibly have access to all the recent advances in Chinese historical phonology that we have today.

The pro-Turkic point of view was recently resurrected by Chen, who argued again in favor of the position that Tabγač was a Turkic language (2005). This article is practically void of actual examples and their discussion, but the major argument of Chen, if I understood is it correctly, is based on his observation that Tabγač personal names were shortened to monosyllabic words from bi- and tri-syllabic, and that this process is due to the oxytonic nature of stress in Turkic, while it would not be possible in Mongolic that has the prototonic stress (Chen 2005: 165–70). Needless to say, it is impossible to take this as a serious argument for several reasons: a) languages can undergo changes in their stress patterns, b) although most if not all modern Turkic languages indeed show the strong preference for the oxytonic pattern, the existence of the oxytonic pattern in Old Turkic, let alone Proto-Turkic is more questionable, c) reductions taken per se cannot prove or disprove a language relationship which has to be demonstrated by conventional means based on the comparative method, d) such a demonstration has to be based on the philologically attested materials. Neither (c) nor (d) is done in Chen (2005); therefore I think that this attempt to revive the Turkic affiliation of Tabγač was not successful.

Thus, it appears that the main issue of the Tabγač controversy is the debate on its Turkic or Mongolic genetic affiliation. This is not to say that Tungusic connection was not proposed, but the arguments for it are either based on ethnic names, comparing Xian-bei (*Sar-pi) with Ewenki (Shavkunov 1968: 25-26), which is not only methodologically weak, as mentioned above, but also is completely unrealistic from the phonetic standpoint; or on the etymology of the Tuyuhun (not Tabγač! – A.V.) personal name 戊寅 EMC *muwH yin* (HHS 97: 7a), that is then compared to Ewenki *mayin* 'protector spirit,' Ewen *mayis* 'protector spirit of a shaman,' Neghidal *mayin* 'the spirit-master of the upper world, Nanai *mayin mafa* 'id.' (Cincius 1975: 521), plus some other similar sounding but unlikely related Tungusic words (Simonov

1996: 79-81). Even leaving aside the phonetic implausibility (Tungusic /a/ vs. Tuyuhun /u/), this proposal should be abandoned as completely *ad hoc*, because there is no philological basis for it in the text of the HHS that leaves this name without a Chinese gloss that would allow us to know its meaning.

The goal of this article is two-fold. First, I will review all the Tabγač and Xian-bei words discussed by Ligeti with the main purpose of revising their reading from the view point of the recent advances in the Chinese historical phonology with some additional comments in order to make the presentation that they are indeed Mongolic, or more exactly, para-Mongolic, more tight. Second, I will present some modest, but new additional evidence in favor of the fact that Tabγač was indeed a para-Mongolic language.

In his 1970 article Ligeti presented two Xian-bei and nine Tabγač words. Below I will not repeat Ligeti's transcription of Middle Chinese, but will replace it according to EMC readings presented in Baxter's on-line dictionary (2000),[1] and LMC and EM reading according to Pulleyblank (1991). Let me start with two Xian-bei words:

(1) 鳥候秦 EMC *tewX huw dzin* (SKZ 30: 6a-b), 託紇臣 EMC *thak hot dzyin* (SS 84: 20b), 土護眞 EMC *thuX huH tsyin* (LMC *thuə́ xɦuə̀ tṣin*) (XTS 219: 5a) 'soil, dust.' Ligeti compared it with Kitan 陶猥思 EM *t^{h}aẃ uǰ sz̩*, MM *to'osun ~ to'usun,* WM *toγusun* 'dust' and other related Mongolic forms (1970: 288). At the first glance the EMC transcriptions of the first syllable look completely unsystematic. However, this bewildering variety has a simple explanation if we look into the gaps of distribution in the EMC syllabary. The syllable *to simply does not exist in EMC, therefore the variation *tew ~ thak ~ thuX* represents an attempt to transcribe the syllable /to/ with a set of syllables available in EMC that could represent it approximately. Consistent presence of initial /γ-/ in the second syllable (Baxter's /h-/) confirms the presence of -γ- in the Xian-bei form. The last syllable really does not match, with Xian-bei forms implying *-čin* versus MM and WM *-sun*. We should not, however, forget that MM and WM *-sun* is a well-established nominal suffix (Poppe 1964: 44), and that Xian-bei is not a Mongolic, but a para-Mongolic language.

(2) 紇眞EMC *hot tsyin* 'thirty' (YHJXTZ: 14: 13a). Ligeti read this word as *yučči̇n ~ yučin* 'thirty' and compared it with MM *qučin ~ yučin* 'id.' and other related Mongolic forms (1970: 290). While this word certainly resembles MM *qučin ~ yučin* 'thirty,' a couple of observations are in order, since at the first glance EMC 紇 /hot/ [γʌt]

looks like a poor equivalent for Mongolic /γu/. The answer again might lie in the distribution: there is no syllable *hwo [γo] in EMC, although there is a syllable *hu [γu]. The reason why the latter was not used is probably connected with the fact that we are dealing here with a para-Mongolic language that might have had a different vocalism. I tentatively read this word as *γočin 'thirty.' The absence of the syllable *hwo [γo] in EMC may also explain why a syllable with a final stop /-t/ was used in the transcription. In any case, there is no resemblance to OT *otuz* 'thirty.'

Now let me turn now to the Tabγač words:

(3) 乞萬眞 EMC *khjït mjonH tsyin* 'interpreter' (NQS 58: 3b). Ligeti read this word as *kelmürčin* comparing it with WM *kelemürči* 'interpreter' (1970: 292).[2] The transcriptional evidence in favor of EMC /-n/ reflecting syllable-final /-r/ in 'Altaic' languages is abundant, and while the vowel in EMC /mjonH/ can seem to be a far cry from the Mongolic /mür/, in spite of the presence of the syllable /mu/ in EMC, there was no syllable *mun, let alone *mün, thus the transcription of the second syllable should not present any problem. The rendering of Mongolic /-l/ by EMC /-t/ is much more unusual, as we would normally expect EMC /li/, /lje/ or /lej/ transcribing it. In addition, there is EMC syllable /ket/, but as far as I can tell there is no syllable *khet. Mongolic voiceless /k/ was probably aspirated as it is today and the absence of the syllable *khet might have prompted the usage of the syllable /khjït/, especially if the vowel /e/ was phonetically [ə]. This word as well as the words (4) – (8) below has one apparent Turkic feature: the *nomen agentis* in *-či*, which is well attested in Mongolic, but certainly represents a loan from Turkic. Nevertheless, Tabγač *kelmürčin* 'interpreter,' as pointed out by Ligeti (1970: 292) also includes a purely Mongolic deverbal derivational suffix *-mür- < *-mar- ~ -mer-* 'suitable to do X' (Poppe 1964: 48), so *kele-mür-* can be understood as 'suitable to do talking.' It is also necessary to note that this and the next following Tabγač words all end in *-n*, which is not found in Turkic after *-či*, and, therefore, must be a para-Mongolic suffix, most likely comparable with the Mongolic 'fleeting' *-n*. Thus, there is nothing Turkic in this word, except one derivational suffix *-či*.

(4) 拂竹眞 EMC *phjut trjuk tsyin* 'servant at a post station' (NQS 58: 3b). Ligeti read this word as *pürtükčin* and compares it with WM *örtegeči(n)* 'one who attends to the horses at the post station, coachman, one on the postal relay duty' and other related Mongolic forms (1970: 293-94). There are no EMC syllables *phjot or *phwot so the transcription of the first syllable seems to be impeccable within

the limitations of the EMC syllabary. We have certainly seen in (2) above that Xian-bei syllable-final /-r/ was transcribed as /-n/ rather than /-t/. Since these two transcriptions come from the same text, the discrepancy is hard to explain, but like the transcription of the syllable-final /-r/ with /-n/, the transcription of the syllable-final /-r/ with /-t/ is also well known. The transcription of the second syllable with the retroflex initial /tr-/ may be puzzling, since there are EMC syllables /tok/ and /thok/, but it is quite possible that a Chinese scribe made an attempt to indicate Tabγač /ü/ by choosing a syllable with medial /-j-/, which does not occur after dentals in EMC. I agree with Ligeti's transcription of this word as *pürtükčin*.

(5) 咸眞 EMC *xamX tsyin* or *hɛm tsyin* 'official at a post station' (NQS 58: 3b). The issue is complicated here by the fact that the character 咸 has two different EMC readings: *xam* or *hɛm*. Ligeti read this word as *γiamčin* and compared it with MM *ǯamučin* 'post office people,' *ǯam* 'post office' and other Mongolic related forms (Ligeti 1970: 295). He further noted Ma. *giyamun* [g'amun] 'relay post station' that represents, in his opinion, an ancient loan from a Xian-bei dialect. Indeed, Ma. *giyamun* can be only a loan from a para-Mongolic language, since Manchu has initial /ǯ-/ that would likely surface if the word was borrowed from a Mongolic language with initial /ǯ-/. Another loanword from para-Mongolic into Manchu that Ligeti does not mention but which demonstrates the same development of pre-proto-Mongolic *γi- to Mongolic *ǯi-* is Ma. *gida* 'lance, spear' that is also found in virtually any other Tungusic language (Cincius 1975: 148–49), cf. MM *ǯida* (MNT 232, HYYY 9b.5, Mu 209, IM 436), WM *ǯida* (Bur. 20b, 55a, 57a, 59a),[3] and the related forms in other Mongolic languages such as Khalkha, Buriat *жад*; Dagur *giad*, Dongbu Yugu *ʒəda*, Manguer *ʒidaa*, Santa *ǯɯda*, etc. 'spear.'[4] Thus, it is almost certain that the Chinese character 咸 is used in this case with its EMC reading /hɛm/ [γɛm]. Since there is no EMC syllable *hjam (the only other options are EMC /hæm/ and /xjem/), it seems that EMC 咸 /hɛm/ is a good approximation for the hypothetical Tabγač *γiam. I agree with Ligeti's reading of this word as *γiamčin*.

(6) 可薄眞 EMC *khaX bak tsyin* 'door-keeper' (NQS 58: 3b). Ligeti read this word as *qabaqčin* (and I concur with his reading), and although he tried to argue against the Turkic etymology, while citing the relevant Turkic forms such as OUIG *qapaγči* ~ *xapaγči* 'door-

keeper' on the basis of MM *qa'a-* 'to close' ~ *qa'alqa* 'door,' and other related Mongolic forms (1970: 296–98). The latter are certainly to be recognized as Turkic loanwords in Mongolic. I see no potential harm for the Mongolic hypothesis of the Tabγač linguistic affiliation in admitting this word as a Turkic loanword, because cultural loanwords from Turkic are certainly to be expected in Tabγač. There is, however, one feature that Ligeti did not mention, and that ultimately betrays this word as being para-Mongolic, albeit based on the Turkic loanword: namely, the suffix *-n*, which is clearly non-Turkic and that was already mentioned in (3) above.

(7) 比德眞 EMC *pjijX tok tsyin* 'secretary' (NQS 58: 3b). Ligeti read this word as *bitekčin*, with a note that while the oldest Turkic forms are *bitigüči* ~ *bitkäči* 'id.', the form *bitikči* must be more recent (1970: 298–99). This, with all respect due to the great Hungarian scholar, makes no sense at all. The oldest Turkic forms are all with /-t-/ unlike all the Mongolic forms cited by Ligeti that invariably have /-č-/ or /-c-/ (1970: 300–02). The palatalization /t/ > /č/ ~ /c/ /__i clearly indicates that all Mongolic forms are subject to the common Mongolic process of palatalization *ti > či (cf. also (9) below) that did not affect para-Mongolic Tabγač. EMC speakers certainly were able to differentiate at least between dental stops and several types of affricates, so the substitution of /t/ for /č/ makes no sense here. In addition, Tabγač is at least two centuries older than the earliest attested Turkic forms. In all probability, the Tabγač word in question was borrowed from a non-Bulghar Turkic language, but since it has the Mongolic suffix *-n*, the same line of reasoning as in (3) and (6) above should be applicable. Finally, a minor correction to Ligeti's reading *bitekčin* should be proposed. Syllables /tek/, /thek/, and /dek/ are perfectly attested in EMC, but there are no syllables *tik, *thik, or *dik. Therefore, the usage of EMC 德 /tok/ probably implies that a high front vowel /i/ or a high central vowel /ï/ was involved here. Given that the vowel in the first syllable clearly renders Tabγač /i/, I opt for the reading *bitikčin*.

(8) 契害眞 EMC *khjit hajH tsyin* 'assassin' (NQS 58: 3b). Ligeti read this word as *qityaičin* and compared it with MM *kidu'ači* 'killer,' *kidu-* 'to kill', WM *kidu-* 'id.' and other related Mongolic forms, with a proposal that the word involves Mongolic *nomen agentis* in *-γči* (1970: 303). The reduction of the final vowel in the second syllable probably should present no problem, because it is a phenomenon frequently observed in para-Mongolic. It is more of a problem to connect Tabγač *-γaiči* with the Mongolic *-γči* as we have to explain the diphthong /ai/ in Tabγač. The choice on the part of Chinese scribe

to render a zero vowel with /ai/ seems to be unrealistic. It is not inconceivable that in Tabγač *kid-* was actually an intransitive verb 'to die' rather than 'to kill': it is impossible, of course, to prove this assumption positively, given our limited sources on Tabγač, but the possibility should not be discarded. Then Tabγač *-γai-* transcribed with 害 can nicely correspond to Mongolic *-γa-* ~ *-γe-*, a causative suffix. There are only two competing syllables in EMC that could be used to transcribe /-γe-/, a variant that we would expect here by the vowel harmony (if it was present in Tabγač like in Mongolian): /hæ/ and /hɛj/. The validity of this analysis is further confirmed by the observation that we have the same 'tail' as in the preceding words (3)–(7): *-či-n*, consisting of the Turkic derivational *-či*, apparently borrowed by Tabγač, and Mongolic *-n*. Therefore, I propose to change the reading of this word into *kityečin* (*kit-γe-či-n* with a morphological analysis).

(9) 莫堤比 EMC *muwX tej pjijX* 'military governor of a province'[5]. Ligeti read this word as *modi bi* and compared it to WM *moǯi ~ muǯi* 'province' and other related Mongolic forms (1970: 304). He further indicated (with a reference to Bazin (1950: 306, 309)) that the last element *bi* occurs in the numerous designations of official positions. I must also add that MM *moǯi* in the meaning 'position' is also attested: bura'ul-ba noyan-i *moǯi*-asa '[he] relieved the noble from [his] *position*' (Mu 127). The Mongolic etymology seems quite impeccable, and it is also necessary to note that this word starts with /m-/, which does not occur in Turkic except in loans and in first persons pronouns, where it is a result of regressive assimilation of the primary /b-/ to the last stem consonant /-n/. The Chinese transcription of the first syllable is unproblematic, since there is no syllable *mwo in EMC. The same can be said about the transcription of the second syllable: there is only one character 地 'earth' that has the reading /dijH/ in EMC, but it is an aberrant case, because dental initials do not combine with the following /i/. Thus, it is quite possible that EMC reconstruction here is erroneous. Therefore, the foreign /ti/ or /di/ was normally transcribed with EMC /tej/, /thej/, or /dej/. There is no syllable *di in Mongolic, since it was palatalized to /ǯi/, however, we would not expect that the same process took place in para-Mongolic Tabγač. Consequently, when Chinese scribes were faced with a dilemma whether to sacrifice accuracy of the consonant or a vowel, they opted to preserve the consonant followed with only an approximate vowel that was available.

(10) 宥連 EMC *hjuwH ljen* ‘cloud’ (WS 113: 44a). Ligeti read this word as *üglen* and compared it with MM *e’ülen ~ e’üle ~ eülen*, WM *egülen* ‘id.’ and other related Mongolic forms (1970: 304). I believe that the identification itself is correct, but I think that the Ligeti’s reading is erroneous. Ligeti correctly did not include the initial vowel /e-/ into his reading, as there is no indication in Chinese transcription of it. However, there is no indication of the metathesis /gü/ > /üg/, either. We have already seen in (4) above that EMC medial /-j-/ may represent an attempt to indicate a front rounded vowel /ü/ that was not present in EMC. But there is clearly an EMC /h-/ (= [γ]) present in the beginning of the word. Thus, I read *gülen* or *γülen*,[6] and the word must be a result of the aphaeresis that pre-Proto-Mongolic *egülen (or *exülen) underwent in Tabγač. This word is very important for determining the linguistic affiliation of the Tabγač language, since it belongs to the basic vocabulary like the Xian-bei words (1)–(2) above, and unlike all other previous Tabγač words (3)–(9) that represent cultural vocabulary. The Turkic word for ‘cloud’ is *bulut*, and it is certainly not applicable here.

(11) 叱如 EMC *tsyhit nyo* ‘wolf’ (WS 113: 45b). Ligeti read this word as *činno* ~ čino and compared it with MM *čino ~ činō ~ činā*, WM *činu-a* ‘id.’ and other related Mongolic forms (1970: 305–06). Ligeti’s identification is beyond any doubt, but one of his readings, namely *činno* and the Chinese transcription merit somewhat more extended discussion. The syllable /tsyhi/ without a final consonant is perfectly attested in EMC, and the same is true about the syllable /tsyhin/. Thus, if the form *činno really existed it would make much more sense to transcribe it as /tsyhin nyo/. It makes even less sense to transcribe *čino* as /tsyhit nyo/ and not /tsyhi nyo/. Therefore, it seems to be a puzzle why Chinese scribes used here the syllable /tsyhit/ that has final consonant /-t/. In the Tabγač words recorded in the NQS that were cited above the Chinese final voiceless stops were clearly functional in rendering certain consonants of Tabγač, see (3)–(4) and (6)–(8). The NQS is just about twenty years earlier than WS, so this discrepancy might seem strange. However, there is a possible explanation. We know that Xiao Zi-xian (蕭子顯, 489–537 AD), the author of the NQS, was a descendant of the founding emperor Gao of the Southern Qi (南齊), and that he wrote the NQS during the next Liang dynasty. Thus, Xiao Zi-xian was himself a Southerner, and Southern dialects of Chinese are notorious for preservation of EMC final voiceless stops even today. On the other hand, the author of WS was Wei Shou (魏收, 506–572 AD), born in Hebei, and thus certainly a Northerner, or more exactly Northeasterner. Although we know that

in the Northwestern Chinese dialect of Chang'an the final stop *-t* was lenited to *-r*, and that consequently all final voiceless stops disappeared in the North by the Yuan period, the evidence for the history of final voiceless stops in Northeastern Chinese is very meager. We certainly know that in the Beijing dialect of the eleventh century final consonants *-t* and *-k* have likely shifted to the final glottal stop *-ʔ* (Iakhontov 1980: 41). Unfortunately, we have neither the evidence of the similar process nor to the contrary for the Northeastern dialect of the mid-sixth century when the WS was compiled. However, another Tabγač word from WS that I am going to discuss below also seems to support the fact that final *-t* and *-k* have already disappeared from Northeastern Chinese by the mid-sixth century. Another problem may seem the transcription of the second syllable as 如 /nyo/ with palatal /ny-/ (= /ń-/) rather than with plain dental /n-/. The gaps in distribution in the EMC syllabary again are helpful for explaining this oddity. While there are EMC syllables /nu/ and /nuw/, there are no syllables *no or *now. Consequently, in this case the choice was made in favor of the vowel and accuracy of the consonant was sacrificed. Thus, I read the Tabγač word as *čino* 'wolf.'

To recapitulate, I believe that Ligeti's proposal (1970) that Tabγač was a language related to Mongolic should be accepted and that the theory that Tabγač is related to Turkic should be discarded. Namely, within the above list we have two Xian-bei words ('dust,' 'thirty') and several Tabγač words ('interpreter,' 'assassin,' 'cloud,' wolf') either belonging to the basic vocabulary or being derived from Mongolic basic vocabulary. Two words that can be traced back to Turkic vocabulary ('door-keeper,' 'secretary') belong to cultural vocabulary. In spite of the presence of the Turkic derivational suffix *-či*, we also have two Mongolic derivational suffixes *-mür-* and *-ge-*, as well as the inflectional suffix *-n*. The revision of Ligeti's readings of Tabγač words in Chinese transcription presented above essentially eliminates all doubtful cases and demonstrates that the Chinese transcription is absolutely systematic within the limits of the EMC syllabary.

In conclusion, I am going to provide an extra evidence for the Mongolic genetic affiliation of the Tabγač language by discussing two words that were not used by Ligeti. Similar to (10) and (11) above they are also found in the WS.

(12) 若干 EMC *nyak kan* 'dog' (WS XX). The word is clearly para-Mongolic, cf. Kitan 狀ㄌ *ńaxa (Dao 4.12, 4.17; Xing 1.12), 捏褐 *ńaXa (NLMC *niaʔ-xɦaʔ) (LS 53:8a)[7] 'dog,' MM *noqai* (MNT 21, 153; HYYY 6a.3; Mu 136, IST 58), *noγai* (IM 442) 'dog,' WM *noqai*

(Qqor. I: 11, Sub. II: 7b, Qqot. IV: 13) 'id.,' Ma. *niyahan* 'puppy' (MYK 19a, MDS 17a), *niyaha-ša-* 'to hunt with dogs,' apparently borrowed from a proto-Mongolic language, possibly even from Tabγač.[8] On the other hand, Jurchen 捏哈 *nieha* 'small dog' (Kane 1989: 234), looks more like a direct loan from Kitan. It is interesting to note that the first syllable of the Tabγač word *nyak kan* is transcribed by an EMC syllable *nyak* ending in *-k*. The syllable *nya was not available in the EMC syllabary, although there is a syllable /nyæ/. Thus, the usage of the final *-k* in the first syllable may be explained by the fact that Wei Shou tried to imitate the vowel as close as possible. However, two other explanations also seem feasible. First, it is the explanation brought up in (11) above: there a possibility that final *-t* and *-k* were gone from Northeastern Chinese as early as in the sixth century, although it is almost impossible to prove this claim. Second, it is possible that a combination of the final *-k* in the first syllable *nyak* and the initial *k-* in the second syllable *kan* was indicating the 'heavy' uvular consonant /-q-/ in the Tabγač form, since there is no uvular /q/ in EMC. I will return to this problem in the discussion of the next word that is also attested in WS, and presents exactly the same combination *-k k-* corresponding to /-q-/ in another Mongolic word.

(13) Much more enigmatic 木骨閭 EMC *muwk kwot ljo* 'bald head, baldy' is found in a passage from WS that I reproduce below:

> 始神元之末掠騎有得一奴髮始齊眉忘本姓名其主字之曰**木骨閭木骨閭**者首禿也
>
> In the beginning of the end of the Shen-yuan era, a marauding [Toba-Wei] mounted warrior obtained a slave. [The slave's] hair started only from the level of [his] eyebrows, and [he] did not remember his name and surname. The master nicknamed him *Muwk kwot ljo*. *Mwuk kwot ljo* means 'head-bald.' (WS 103: 1a).

The Tabγač word *muwk kwot ljo* was discussed before by Boodberg (1979[1933]: 94) and Taskin (1984: 399). Boodberg compared it to the Ruan-ruan clan name 郁久閭 *ʔyuwk kjuwX ljo*, hypothesizing that these two words represent variations of the same word due to *Ø-* ~ *m-* initial variation that, of course, cannot be proven for 'Altaic' languages, let alone for Ruan-ruan. In addition, while 郁久閭 *ʔyuwk kjuwX ljo* is the Ruan-ruan clan name, 木骨閭 *muwk kwot ljo* is clearly a Tabγač word, since it was used by a Tabγač mounted warrior. Taskin's etymological attempt fares even worse, because he

compared Tabγač *muwk kwot ljo* with WM *mayu* 'bad' plus Monguor *rāwa* 'hair' using modern Mandarin readings /mu gu lü/ of 木骨閭.

I will apply to this word the same line of reasoning as was used above. Whether the final *-k* and *-t* might have already disappeared in the Northeastern Chinese by the mid-sixth century or not (see (11) and (12) above), it is also important to keep in mind that EMC has syllables /mu/, /muw/, /mok/, /mwot/, /ku/, /kuw/, /kwot/, /khu/, /khuw/, /khwot/, and /kop/, but no syllables *mo, *mwo, *ko, *kot, *kwo, *kho, and *khwo. Thus, even if the final *-t* was not lost in the Northwestern Chinese by the mid-sixth century, the syllable /kwot/ might have been the best approximation for the foreign /ko/. It is more difficult to explain the choice for /muwk/ with a final *-k*, but we should keep in mind that the loss of the final *-p, -t, -k* in Chinese dialects always started with a velar *-k*, with labial *-p* being lost the last. There is also another possibility: the double *-kk-* in the Chinese transcription might indicate a uvular /q/ that was not present in EMC, cf. also the same point made about (12) above. I believe that EMC transcription should be interpreted as *muqolo ~ *moqolo 'bald.'

Let me note that *muqolo ~ *moqolo 'bald' cannot be a Turkic word, not only because Turkic has a different word for this notion: OT *taz* (MK I: 26), but also because it starts with *m-*, a initial consonant that does not occur in the Turkic native vocabulary. At the first glance it may seem that there are no Mongolic etymologies either, but Manchu offers us some help here as it has *mohoto* 'hairless (of a horse tail),' (Norman 197X: XX), *mokto* 'bald tail' (spelled *mogto*) (SBB III: 3575), and Ma. *moholo* 'hornless.' Contrary to Cincius, these words cannot be connected genetically to their Tungusic pseudo-cognates,[9] such as Ewenki *muku-* 'to become blunt,' 'to break off (of a knife)' Solon *moxō* 'blunt (of a knife),' *moxor* 'hornless,' Ewen *muk-* 'to cut off,' *muktar-* 'to fall off (of a reindeer's horn),' *mukter-* 'to become blunt' (1975: 553-54), [10] because both words are morphologically indivisible in Manchu: there is no Ma. verb or adjective *moko- 'to become blunt, to be hornless,' there is no Ma. nominalizer *-lo (only verbalizer *-lo-* is attested), and there is no Ma. nominalizer *-to. I believe that *-lo* represents a para-Mongolic comitative suffix *-lo cognate to MM *-lu'a* and WM *-luγa*. We have already seen above that MM *činō* and WM *činu-a* 'wolf' (Turf. 53: 24) corresponds to Tabγač *čino* 'id.'. In other words, the pre-proto-Mongolic sequence *u[x]a contracts to /o/ in Tabγač. Ma. *mohoto ~ mokto* 'hairless tail' is likely to be an early loan[11] from WM *moγutur* which is used as its translation equivalent in the SBB (III: 3575). WM *moγutur ~ moqutur* usually means 'blunt, hornless,' but the meaning

of the Mongolian SBB gloss is further confirmed by Tibetan *mjug rdum* 'tailless,' and Chinese 禿尾 'bald tail.' We then can see the semantic variation 'blunt' ~ 'hornless' ~ 'hairless' ~ 'bald.' The first two meanings are well supported by other Mongolic comparative data, cf. MM *muqular* 'hornless' (MNT 121, 214),[12] MM *moqo-dag* 'hornless,' 'blunt' (HYYY 27b: 5, Mu 238), WM *moqu-γa-* 'to make blunt' (Bur. 40b, Bod. 13a: 7). The evidence for 'hairless' ~ 'bald' is more limited, as it is supported only by Ma. *mohoto ~ mokto* 'hairless tail' and the WM *moγutur* 'id.' in the SBB. However, a person can become bald either by losing his hair naturally or by cutting it off like a Buddhist monk. It is interesting to look from this point of view at Ewenki *muku-* 'to become blunt,' 'to break off (of a knife),' Solon *moxō* 'blunt (of a knife),' *moxor* 'hornless,' and Ewen *muk-* 'to cut off,' *mukter-* 'to become blunt,' *muktar-* 'to fall off (of a reindeer's horns)', and Ma. *moholo* 'hornless,' *mohoto ~ mokto* 'hairless tail' cited above. It is quite obvious that all these forms must go back (via some intermediate source) to the Mongolic archetype *muqu-, with absolutely expected and regular syncope in Ewen. The semantic range in Tungusic is much greater than in Mongolic, and I think that one can reconstruct on its basis a semantic archetype *'to be cut off' ~ 'to break off' ~ 'to fall off.' Thus, 'blunt,' 'hornless,' and 'hairless' are obvious semantic derivatives from this archetype, and apparently only these derivatives survived in extant Mongolic languages. Therefore, Tabγač *muqo-lo ~ moqo-lo* 'bald head' is probably to be analyzed as 'one which has cut off/fallen off [hair].' Let me note that it has exactly the same para-Mongolic suffix *-lo* as Ma. *moholo* 'hornless,' and its vocalism is closer to Manchu forms *moholo* 'hornless,' *mohoto ~ mokto* 'hairless [tail]' than it is to Mongolic or other Tungusic forms. This is probably is not a sheer coincidence, as both Manchu and Jurchen are inundated with loans from para-Mongolic, presumably from Kitan (Janhunen 2003: 400), although the comparable Kitan word has not yet been found either in Chinese transcriptions of Kitan, or in the Kitan inscriptions. Nevertheless, I hope that the above demonstration added two more Tabγač words with Mongolic etymologies to Ligeti's list, and, therefore, provides an additional evidence for the Mongolic affiliation of the Tabγač language.

Abbreviations of languages:

Chuv.	Chuvash
EM	Early Mandarin (Pulleyblank)
EMC	Early Middle Chinese (Baxter)
LMC	Late Middle Chinese (Pulleyblank)
Ma.	Manchu

MM	Middle Mongolian
NLMC	Northern Late Middle Chinese (Shimunek)
OT	Old Turkic
OUIG	Old Uigur
WM	Written Mongolian

NOTES

[1] Several notes on Baxter's transcription: /o/ represents a mid-low unrounded vowel, while /wo/ is [o]. X after the syllable indicates the entering tone (Chin. *ru-sheng*) and H the departing tone (chin. *qu-sheng*). The initial /h-/ represents a voiced fricative /ɣ-/ or /ɦ-/, and initial /x-/ a voiceless /x-/ or /h-/. Combinations with /-y-/ are palatals: /tsy-/ is phonetically [č-], /dzy-/ is /ǯ-/ and /ny-/ is [ń-]. Combinations /tr-/, /dr-/, /sr-/, and /nr-/ are not the clusters, but the retroflex consonants.

[2] Starting with Ramstedt, Altaicists cited OUIG *kälmäči* (sic!) 'interpreter' alongside with Chuv. *kala-* 'to speak' as the 'proof' that 'Proto-Altaic' *kälä- 'to speak' existed also in Proto-Turkic (Ramstedt 1957: 47), (Starostin 1991: 57-58, 131, 292), (Miller 1998: 27). However, it seems that the OUIG *kälmäči* 'interpreter' is a *hapax legomenon*. The ultimate source for this word seems to be Radloff's monumental dictionary, where the word is given in Old Uigur script as kEla me[h *kälämäči* 'interpreter' (1899.2: 1116). Radloff also notes that this is likely to be a mistake for *tïlmači* 'interpreter,' and gives as his source 'Chinesisch-Uigurisches Wörterbuch' in the Asiatic museum (nowadays Institute of Oriental Studies in St. Petersburg), with no further specifications. Ligeti indicates that this is without any doubt a copy of the Chinese-Uigur *Hua-yi yi-yu* (Ligeti 1966: 118). In the copy published by Ligeti the word also appears as *kälämäči* 'interpreter' (Ligeti 1966: 172, 299). But the word is not attested in any other sources. The verb *kala-* (< *kälä-) 'to speak' is attested independently only in Chuvash, but not in any other Turkic language, including Old Uigur. It is a wishful thinking to reconstruct Proto-Turkic *kälä- to speak on the basis of a *hapax legomenon* in Old Uigur, another Karakhanid *hapax legomenon kelēčü* 'talk, conversation' (MK I: 445), and on Chuv. *kala-*. The Old Uigur word is in all probability a loan from the para-Mongolic word that we are discussing here, unless it is indeed a mistake as Radloff cautiously suggested. Admittedly, since Chuvash is at the highest level of branching of Turkic, Chuv. *kala-* 'to speak' can reflect a Proto-Turkic etymon, but it is equally possible that it is a loan either from Mongolic *kele-* 'to speak,' or from Finno-Ugric *keeli ~ *keele- 'tongue.'

[3] An extremely helpful tool for tracking down earliest WM attestations is (Tumurtogoo 2006).

[4] Only Dagur among all Mongolic languges preserves initial *g-* in this word: *giad* 'spear.'

[5] Ligeti does not indicate the source for this word.

[6] The symbol /ɣ/ here does not refer to the back velar customarily denoted as /ɣ/ in Mongolic transcription, but serves the only purpose to provide a closer approximation to the Chinese transcription.

[7] The Kitan and NLMC reconstructions are by Shimunek (2007: 13). His Kitan reconstruction follows Kara (1975).

[8] Shimunek believes that Ma. *niyahan* 'puppy' is a loan from Kitan *ńaXa* 'dog' (2007: 43), however, the presence of the final *-n* in both Manchu and Tabɣač and absence of it in Kitan rather indicate that it is more likely to be a loan from Tabɣač.

[9] Ma. /o/ cannot correspond to Ewenki or Ewen /u/ in real Tungusic cognates. Thus, all these forms must represent different layers of Mongolic loans.

[10] I omitted several language entries, such as Oroch onomatopoetic *mōkto* 'to gnaw through the neck (of a carnivore animal)' or Nanai *moqtoa-* 'to break' that clearly do not belong here.

[11] Final *-r* in Manchu is normally found only in onomatopoeia or in late loanwords.
[12] The morphemic division here seems quite apparent *muqu-* + verbalizer *-la-* + nominalizer *-r*.

BIBLIOGRAPHY

Primary sources:

Chinese

HHS	Hou Han shu (後漢書), ca. 440 AD
LS	Liao shi (遼史), ca. 1350 AD
NQS	Nan Qi shu (南齊書), ca. 530 AD
QGZ	Qidan guo zhi (契丹國志), 1247 AD
SKZ	San guo zhi (三國志), ca. 290 AD
SS	Sui shu (隋書), ca. 640 AD
WS	Wei shu (魏書), 551-54 AD
XTS	Xin Tang shu (新唐書), ca. 1060 AD
YHJXTZ	Yuan he jun xian tu zhi

Kitan

Dao	Daozong Huangdi Aice (道宗皇帝哀冊) inscription
Xing	Xingzong Aice (興宗皇帝哀冊) inscription

Manchu

MDS	Enetkek tubet monggo han se-i da sekiyen-i suduri, 18th c. AD
MYK	Manju yargiyan kooli bithe, 1781 AD
SBB	Han-i araha sunja hacin-i hergen kamciha manju gisun-i buleku bithe, 1790 AD

Mongolian

Bod.	Bodhicaryāvatāra from Olon Süme
Bur.	Arban qoyar ǯokiyangγui orosiba (Twelve deeds of the Buddha)
HYYY	Hua-yi yi-yu, 1389 AD
IST	Istanbul glossary
IM	Ibn-Muqannā glossary
MNT	Mongγol niuča tobča'an, 1240 AD
Mu	Muqaddimat al-Adab, 15th century
Qqor. I	Mongolian inscription from Qara Qorum, 1346 AD
Qqot. IV	A fragment from Qara Qoto IV
Sub.	Mongolian version of Subhāsitaratnanidhi, 16th c. (?)

Turf. 53	An unknown manuscript fragment from Turfan
Turkic	
MK	Dīwān Luγāt at-Turk by Maḥmūd al-Qāšγarī, 1072-74
AD	

Secondary sources:

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

Baxter, William H. 2000. *An Etymological Dictionary of Common Chinese Characters*. Preliminary Draft of 28 October 2000.

Boodberg, Peter A. 1933. Sino-Altaica I. *Hu Tien Han yüeh Fang Chu 5* (January 1933). Reprinted 1979 in *Selected Works of Peter A. Boodberg*, ed. Alvin P. Cohen, pp. 94-114. Berkeley: University of California Press.

——— 1936. The language of the T'opa Wei. *Harvard Journal of Asiatic Studies* 1: 167-85.

Chen, Sanping 2005. Turkic or Proto-Mongolian? A Note on the Tuoba Language. *Central Asiatic Journal* 49.2: 161–73.

Cincius, Vera I. (ed.) 1975–77. *Sravnitel'nyi slovar' tunguso-man'chzhurskikh iazykov, vol. 1-2*. Lenigrad: Nauka.

Clauson, Sir Gerard 1962. *Turkish and Mongolian Studies*. London: The Royal Asiatic Society.

Doerfer, Gerhard 1992. Mongolica in Alttürkischen. In *Bruno Lewin zu Ehren 3*, pp. 30–56. Bochum.

——— 1993. The Older Mongolian Layer in Ancient Turkic. *Türk Dilleri Araştırmaları* 3: 79–86.

Iakhontov, Sergei E. 1980. Pekinskoe proiznoshenie XI v. In *Razyskaniia po obshchemu i kitaiskomu iazykoznaniiu*, ed. S.E. Iakhontov, I.T. Zograf, N.V. Solntseva, A.G. Shprintsin, pp. 35-42. Moscow: Nauka.

Janhunen, Juha 2003. Para-Mongolic. In *The Mongolic Languages*, ed. J. Janhunen, pp. 391-402. London: Routledge/Curzon.

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

Kara, György 1975. A propos de l'inscription de 1150 en écriture khitane. *Annales Universitatis de Rolando Eötvös nominatae, Sectio Linguistica*: 163-67.

Ligeti, Lajos 1966. Un vocabulaire Sino-Ouigour des Ming. *Acta Orientalia Academiae Scientiarum Hungaricae* 19: 117-316.

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

Miller, Roy A. 1998. Altaic *kele(-) 'Tongue, To Speak' in Korean. *Studia Etymologica Cracoviensia* 3: 27-43.

Pelliot, Paul 1925. Les mots à *h-* initial, aujoud'hui anuie dans le mongol des XIII^{e} et XIV^{e} siècles. *Journal Asiatique* 206: 193-263.

Pulleyblank, Edwin G. 1962. The Consonantal System of Old Chinese. *Asia Major* 9.1: 58-144, 9.2: 206-65.

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

Radloff, Wilhelm 1893-1911. *Versuch eines Wörterbuches der Türk-Dialekte*, v. I-IV. St. Petersburg.

Ramstedt, Gustaf, J. 1957. *Einführung in die Altaische Sprachwissenschaft. Lautlehre*. Helsinki: Société Finnou-Ougrienne.

Shavkunov, E.V. 1968. *Gosudarstvo Bokhai i pamiatniki ego kul'tury v Primor'e*. Leningrad: Nauka.

Shimunek, Andrew 2007. *Towards A Reconstruction Of The Kitan Language, With Notes On Northern Late Middle Chinese Phonology*. Unpublished Indiana University MA thesis.

Shiratori, Kurakichi 1912. Tōko minzoku kō. *Shigaku zasshi* 23: 1243-1269.

——— 1913 Tōko minzoku kō, *Shigaku zasshi* 24: 17-45, 854-884.

Simonov, Mikhail D. 1996. Towards an etymology of the Xianbi anthroponym Wuyin Kehan. *Studia Etymologica Cracoviensia* 1: 77-84.

Starostin, Sergei A. 1991. *Altaiskaia problema i proiskhozhdenie iaponskogo iazyka*. Moscow: Nauka.

Taskin, V.S. 1984. *Materialy po istorii kochevykh narodov gruppy dunkhu*. Moscow: Nauka.

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