Tocharian-Turkic lexical correspondences, I

Hakan Aydemir[*]

Abstract: There are many unsolved problems and unknown aspects with regard to Tocharian-Turkic relations. The most fascinating of them are especially the Tocharian-Turkic lexical correspondences. One of these lexical correspondences is the Tocharian B *kärk-* 'to hack up' and the Karachay-Balkar Turkic *kärk-* 'to hack (up), to notch'. Since the structure of the Tocharian word is unclear and there is no acceptable etymology of it, the question arises as to whether it is a borrowing from Turkic. The present study investigates whether the Turkic word is the source of the Tocharian B word. The study states that mainly on the basis of the Turkic and Tocharian morphological criteria, the Tocharian B *kärk-* seems to be a borrowing from Turkic.

Key words: Tocharian, Proto-Turkic, language contacts, lexical correspondence, verb borrowing, Turkic etymology.

Toharca-Türkçe Sözcük Denklikleri, I

Özet: Toharca-Türkçe ilişkilerinin çözülmemiş pek çok problemi ve bilinmeyen pek çok yönü vardır. Bunlardan en heyecan verici olanı özellikle Toharca-Türkçe sözcük denklikleridir. Bu sözcük denkliklerinden bir tanesi, Toharcanın B diyalektindeki *kärk-* 'doğramak, yarmak, kesmek' ve Karaçay-Balkar Türkçesindeki *kärk-* 'yontmak, çentik atmak' denkliğidir. Toharca sözcüğün yapısı anlaşılamadığından ve kabul edilebilir bir etimolojisi bulunmadığından, bu sözcüğün Türkçeden bir alıntı olup olmadığı sorusu akla geliyor. Çalışma, Toharcanın B diyalektindeki bu sözcüğün Türkçeden bir alıntı olup olmadığını inceleyerek, başlıca Türkçe ve Toharcadaki morfolojik kriterlere dayanarak Toharcanın B diyalektindeki *kärk-* sözcüğünün Türkçeden bir alıntı olabileceği sonucuna varıyor.

Anahtar sözcükler: Toharca, Ana-Türkçe, dil ilişkileri, sözcük denklikleri, fiil alıntılama, Türkçe etimoloji.

---

[*] PhD. Istanbul Medeniyet University, Department of Linguistics, aydemirhaakan@gmail.com.

I

As has already been indicated in my previous studies, there are many unsolved problems and unknown aspects regarding Tocharian-Turkic linguistic and ethnic relations. One of the fundamental problems concerns "when" and "where" these relations started and how long they took.[1] The first two questions are – for the time being – very difficult to answer. I think, however, I answered the last one (i.e. 12th-13th c.) in my previous study on Tocharian-Turkic relations.[2] The other unsolved and fundamental problem was and is whether or not the Tocharian ethnonyms *Twqry* and *Twq'ry* occurring in Old Turkic sources can be shown in the toponymy of today's Xinjiang. This issue has not actually been raised in literature so far. Although I have not yet been able to clarify this question, I think I have been able to map the former and factual geographical distribution of Tocharians in Xinjiang revealing many hitherto unknown Tocharian ethnotoponyms and ethnohydronyms. My forthcoming study on this topic has shown that in today's Afghanistan (i.e. in the former Tocharistan) and in the neighboring countries there are also many hitherto unknown Tocharian ethnotoponyms and ethnohydronyms which are very important in view of the history and historical geography of the Tocharians in the former Tocharistan.

II

Following the studies clarifying the historical, chronological and geographical background of the Tocharian-Turkic relations, I am starting to examine the most fascinating problem now, i.e. the problem of Tocharian-(Proto-)Turkic lexical correspondences; firstly, (1) I will explore those that have not yet been discussed by any scholars (e.g. Toch. *kärk-* = Turk. *kärk-*; Toch. *yāṇk-* = Turk. *yaŋ-*; Toch. *ku-* = Turk. **ku-* in: *ku-d-*; Toch. *ār-* = Turk. *ār-*; Toch. *or* = Turk. **or* in: *or+man*; Toch. *ārśe* = Turk. *ars ~ ās ~ as*; Toch. *päl* = Turk. **bal* in: *balbašı* and *bālıg < *bal+lıg*; Toch. *keme, kam* = Turk. *kämä ~ gämä*, *kämir-*, Mo. **kämä* in: *kämä+lä-*; etc.). In this first part, however, I discuss only the Tocharian B *kärk-* 'to hack up' and the Turkic (Karachay-Balkar) *kärk-* 'to hack (up), notch'. I will discuss the rest of the words mentioned above in the next parts of this series. In the later parts of this series (2) the problematic ones will also be addressed which have already been discussed by excellent scholars, but for which the Turkic background has not been clarified or where there is an uncertainty in view of their etymology on the Turkic and/or Mongolic side; and finally, (3) I will deal with the false lexical correspondences that are based on misconceptions and speculations.

---

1   See Aydemir 2009:159, note 2.
2   See Aydemir 2009:159-180.

Turkic (Karachay-Balkar) *kärk-* 'to hack (up), notch'

On the Tocharian side the etymology of Tocharian B *kärk-* is – according to D. Adams – "unclear".[3] For the time being the Tocharian word seems to be attested only in Tocharian B. Van Windekens suggested that it has a connection with PIE *(s)ker-* 'cut'.[4] Adams, however, finds this kind of solution "less convincing".[5] Hilmarsson suggested a relationship with *kärst-* 'cut off'.[6] According to Adams, "such a suggestion provides an excellent etymology semantically but at the cost of complicating the phonological development unduly".[7]

As seen above, the structure of the Tocharian word is unclear and it has no convincing Tocharian or PIE etymology. On the other hand, however, there is no doubt that the Tocharian B *kärk-* and Turkic *kärk-* are one and the same word. This fact, nonetheless raises some important questions. The basic question is firstly (1) whether or not the word *kärk-* is a borrowing in Turkic or in Tocharian. If it is, then (2) "how", "when" and "where" was it borrowed from one language to another? Considering that the traditional verb forms are more difficult to borrow from one language to another than the nouns, and a verb borrowing presupposes a certain degree of bilingualism (mostly forced bilingualism), then the questions arise as to (1) "who" were bilingual, the Tocharians or the Turks, and (2) "when" and "where" were they bilingual? I will attempt to answer all these questions in the following parts of this series, but definitive answers to these questions can, of course, only be given if we examine all the Tocharian-Turkic lexical correspondences, especially the verb borrowings. However, even then there is no certainty that we will be able to answer all these questions.

In Turkological literature there are no remarks on the form *kärk-* 'hack up'. Among the Turkic languages, the word is attested only in the Karachay-Balkar language. It seems, however, to have so far escaped the attention of researchers. Therefore its connection with the other derivatives (*kärt-*, *kärki*, *kärt*) briefly discussed below has not been touched on to date.

The Karachay-Balkar *kärk-* 'hack up' and the Old Turkic words *kärt-* 'to gash, cut a notch' and *kärki* 'carpenter's axe' are the regular derivatives of a common etymon, i.e. of the verbal base **kär-* 'to hack up, cut (a notch)'. All three forms can quite regularly be analyzed as **kär-k-* (> *kärk-*), **kär-t-* (> *kärt-*) and **kär-ki* (> *kärki*). The form *kärt* (< **kär-(U)t*) 'a notch, notched' in Azeri Turkish, and the Chuvash *kart* 'a notch, mark' (*kart* < *kärt* < **kär-(U)t*) also clearly demonstrate that the etymologies of the words in question must be

---

3 Adams 1999:162-163; Adams 2013:172.
4 Windekens 1979:208.
5 Adams 1999:162; Adams 2013:172.
6 Hilmarsson 1996:94.
7 Adams 1999:162; Adams 2013:172.

based on the Proto-Turkic verb *kär-. As M. Stachowski in his book has already shown, the postulated form *kär- of the verb kärt- can actually be shown in the form kir- (< *kär-) 'to bite, gnaw (off)' in Yakut (Sakha) a Turkic language.[8] This very important fact, however, interestingly seems so far to have escaped the attention of all the Turkologists and Altaists, who have discussed the etymology of the words kärt- and kärki. The meaning 'to bite, gnaw (off)' of the Yakut verb kir- is obviously a metaphorical meaning, and can also be shown in kärt- 'to cut, chop; (rodent) to gnaw (off); to notch' in the Altai Turkic language.[9] This metaphorical meaning obviously developed from the meanings 'to cut, chop; to notch' of the verb *kär- in Proto-Turkic.

Regarding the morphology of the Turkic verb kärk- (< *kär-k-), it should firstly be noted that the Turkic kärk- is a transitive verb. So, first of all we have to clarify the status of the formative -k- in Old Turkic, if we start from a base *kär- (> Yakut kir-), because according to current knowledge of Turkology the formative -(X)k- does not form transitive verbs. For practical reasons, I use -(X)k^1- to denote this formative, because – as seen below – we have another formative -(X)k^2- in Old Turkic, which has not been discussed as a separate formative to date.

In Old Turkic there is a -(X)k^1- formative which – as Erdal stated – forms (from tr. or intr. bases) intransitive and inchoative verbs (e.g. böl-ük- 'to get into separate flocks and assemble in them' from böl- 'to separate'; kork- 'to be afraid' from korı- 'to protect', etc.).[10] Erdal also remarks that the "-(X)k- verbs [i.e. -(X)k^1- verbs, H. A.] are neither tr. nor passive, then, nor medial or reflexive. They are all anti-transitive" and "none is accompanied by a direct object or qualifies a direct object, and none has an agentive dative […] with it."[11] As the -(X)k^1- forms anti-transitive verbs in Old Turkic, it cannot be taken into consideration in the case of the verb kärk-. Thus, we have to find another formative, which is not anti-transitive and forms transitive verbs. Thus, we are faced with a verb category in Old Turkic, that has not been investigated so far.

In Old Turkic there is, however, a certain group of words, which Erdal did not discuss in his book Old Turkic Word Formation. They are as follows: (1) yak- 'to ignite, burn (something Acc.)', (2) kırk- 'to shear (sheep and the like Acc.)', (3) irk- 'to collect or assemble (things Acc.)', (4) alk- 'to use up, finish, come to the end of (something Acc.)', (5) sik- '(of the male only) to copulate (with a female Acc.)', (6) sık- 'to squeeze, press, compress', etc.[12] Although this group of words is not so large in Old Turkic, I think that there are still many other examples of this category in modern Turkic languages that have not been

---

8   Stachowski 1993:61.
9   See 'грызть (о грызуне) // (kemirgen) kemirmek' (Alt., AltGND.).
10   Erdal 1993:649.
11   Erdal 1993:650-651.
12   See Clauson 1972.

collected yet. The Karachay-Balkar *kärk-* discussed here is just one of them.

The number of examples from modern Turkic languages can, of course, be enlarged, but I think these examples will be enough to draw a definitive conclusion on the formative *-(X)k²-* in the verb forms mentioned above. On the other hand, we should not be surprised that the corpus of this category in Old Turkic is not so rich. There are, namely, such formatives in Old Turkic, for which one can find only a few examples; such as the case of the medial formative *-d-*, for which Erdal was able to point out just seven transparent examples in Old Turkic.

Now – before we draw a definitive conclusion from the verb form *kärk-* – let us take a brief look at the etymologies of the verbs in question in order to understand the character of the formative *-(X)k²-* in Old Turkic better. It is a common characteristic of the verbs in this group that they are all transitive verbs, but not all the bases of them are attested in Old Turkic. They are all transparent.

yak- (< **ya-k-*) 'to ignite, burn (something *Acc.*)'. The verb is a derivative of the (transitive?) base **ya-* 'to burn, set on fire' unattested in Old Turkic. This base, however, can also be seen in the reflexive verb *yan-* 'to burn, blaze up' in Old Turkic, which is quite regularly analyzed as **ya-n-*, where the formative *-(X)n-* is the well-known reflexive formative. It is clear that it should be analyzed as **ya-k- > yak-*.

kırk- (< *kır-k-*) 'to shear (sheep and the like *Acc.*)'. The verb is a derivative of the transitive base *kır-* 'to scrape, strip (hair)' well attested in Old Turkic. Thus, it should be analyzed as *kır-k- > kır-*.

irk- (< **ir-k-*) 'to collect or assemble (things *Acc.*)'. It is a derivative of a verb base **ir-* *'to collect, bring together, round up (and the like)' unattested in Old Turkic. This base, however, can be shown as a transitive verb in the form *ir-* 'to bring together, to round/herd up' in Kara-Kirghiz Turkic.[13] It is also attested in Mongolian as an intransitive verb, i.e. *ir-* 'to fill up; to pack; to heap up; to crowd'.[14] Thus, it should obviously be analyzed as **ir-k- > irk-*.

alk- (< **al-k-*) 'to use up, finish, come to the end of (something *Acc.*)'. It is a derivative of a verb base **al-* *'to end up, finish (and the like)' unattested in Old Turkic. It is, however, attested in Mongolian as *ala-* 'to kill, murder; to butcher'.[15] It should, thus, be analyzed as **al-k- > alk-*.

sik- (< **sī-k-*) '(of the male only) to copulate (with a female *Acc.*)'. It is a derivative of a verb **sī-*[16] *'to urinate, discharge spermatic fluid', which

---

13 See Radloff I:1456.

14 See Lessing 1960:412b.

15 See Lessing 1960:26a.

16 The length of the vowel of the reconstructed verb **sī-* was stated first by T. Tekin on the basis of the DLT *sīd-* 'to uninate' (< **sī-d-*) and Yakut (Sakha) *īk* 'urine' (< **sī-k*) (s. Tekin 1995:183).

is unattested in Old Turkic, but can be shown in Anatolian as a dialectical variation in the form DS *sī-* 'to urinate; to leak; (sheep and goats) to discharge spermatic fluid'. The base also has other regular derivatives in Old Turkic (i.e. *sīd-* 'to urinate' < **sī-d-* and *sik* 'penis' < **sī-k* > Yakut *īk* 'urine'). Contrary to some opinions, however, I am of the view that the meaning 'to copulate' of the verb is not primary, but has emerged secondarily (i.e. **'to discharge spermatic fluid' → 'to copulate'). So, we have to analyze it as **sī-k- > *sīk- > sik-*.

sık-   (< **sı-k-*) 'to squeeze, press, compress' and metaphorically 'to distress, depress (someone)'. It is obviously a derivative of a verb **sı-* **'to squeeze, press, compress (and the like)', because we also have a noun *sık* (< **sı-*) 'shallow, scanty, tight, dense' in Old and modern Turkic languages. It should, thus, be analyzed as **sı-k- > sık-*.

As seen from the examples given above, all the four attested bases (i.e. *alk-*, *irk-*, *sik-*, *kırk-*) are transitive. Thus, we may assume that the bases **ya-* 'to burn, to set on fire' of *yak-* and **sı-* **'to squeeze, press, compress (and the like)' of *sık-* were transitive as well. In Old Turkic there are, of course, also other such derivatives of the formative *-(X)k²-* (e.g. *ük-* 'to heap up, accumulate *Acc.*' < Turkic dialectical *ǖ-* 'to accumulate', etc.), which have hitherto been unrecognized. However, I will discuss them in a separate study.

Based on the etymologies discussed above we can now draw a definitive conclusion on the status of the formative *-(X)k²-* in Old Turkic: (1) It is obvious that in contrast to the anti-transitive formative *-(X)k¹-*, which forms intransitive and inchoative verbs, the formative *-(X)k²-* forms only transitive verbs (see *yak-*, *alk-*, *irk-*, *kırk-*, *sik-* and *sık-*). (2) While the formative *-(X)k¹-* forms verbs from transitive or intransitive bases, the formative *-(X)k²-* forms verbs from transitive ones only. (3) The transparent *-(X)k¹-* verbs are not attested before the 11th century. Among the *-(X)k²-* verbs, however, two (*kırk-* and *alk-*) are attested much earlier. (4) The *-(X)k²-* seems to be attached to simplex bases. The final word on this formative *-(X)k²-* will, of course, be provided by the results of further research.

As for the semantic function of the formative *-(X)k²-*, the situation is not so easy and clear. In Old Turkic we had so far only been aware of one formative, i.e. *-(X)k¹-* (in Erdal *-(X)k-*). We now, however, have another *-(X)k²-*. The first *-(X)k¹-* has been suggested in earlier literature to be an "emphatic" or "intensive" formative.[17] Erdal is, of course, right when he writes that in the case of the *-(X)k¹-* he "can see no base for this contention".[18] This is absolutely true for the anti-transitive *-(X)k¹-* discussed by Erdal in his book. For the time being, however, I think as a working hypothesis that the second, i.e. the transitive formative

---

17   See Gabain 1950, §160.
18   Erdal 1991: 650.

-(X)k²- being discussed here, may perhaps have had an emphatic or intensive content. This content cannot, however, be seen in the derivatives of the *-(X)k²-* at all. So, its primary semantic function remains uncertain and necessitates further research. It should be noted, though, that the primary semantic function of the transitive formative *-(X)k²-* is an internal Turkic problem and absolutely irrelevant for the question of the Tocharian B *kärk-* and Turkic *kärk-* discussed here.

After clarifying the status and derivatives of the *-(X)k²-* in Old Turkic, we can lastly return to the main question, i.e. to the matching of the Tocharian B *kärk-* and Turkic *kärk-*. As it has been said above, the base of the Turkic *kärk-* is preserved as *kir-* (< Proto-Turkic **kär-*) in Yakut (Sakha). Thus, we have a transitive base to analyze the Turkic *kärk-* as *kär-k-*, where the *-k-* is evidently the transitive formative *-(X)k²-* described above. As for the Proto-Turkic base **kär-* itself, this raises the question as to whether the Proto-Turkic **kär-* 'to cut, to hack (up); to notch; *metaphorically* to bite, gnaw' is a borrowing from PIE, i.e. whether it goes back to the PIE **(s)ker-* 'to cut' or not. This is possible, but ascertaining whether it is the case or not cannot be the task of this study because (1) whether it is determined or not, is irrelevant as regards the question of the Tocharian B *kärk-* and Turkic *kärk-* discussed here; (2) on the other hand, ascertaining this would raise many and very difficult questions both chronologically and theoretically. (3) To ascertain this, we also, namely, need many criteria, that we do not, however, possess for the time being. I will return to this question again after finishing the examination of all the Tocharian-Turkic lexical correspondences. I think we will have a relatively large number of criteria for making a decision on this question.

To summarize what has been said, on the basis of the results obtained here, mainly on the basis of the Turkic and Tocharian morphological criteria, the Tocharian B *kärk-* seems to be a borrowing from Turkic. However, in order to say something definitive on the *real* and *relative chronology* (and perhaps also on the geography) of the Tocharian B borrowing, it may help us to know when the Tocharian B *kärk-* was first attested, i.e. in early or late Tocharian texts. Regarding this, however, I have no information for the time being. This, namely, would be a determinative factor with regard to the real chronology of the borrowing. But the real chronology can be neglected, of course, if we have a substantial reason for this. Of the questions posed above, I think I have answered the first one, i.e. the basic question as to whether or not the word *kärk-* is a borrowing in Turkic or in Tocharian. The question as to "how, when, where and from which Turkic language" it was borrowed, is, however, difficult to answer for the time being. As I have said above, definitive answers to these questions can only be given if we have examined all the Tocharian-Turkic lexical correspondences, especially the verb borrowings.

Abbreviations

Acc.   *Accusative*

PIE   Proto-Indo-European

References

Adams 1999

Adams, D. Q.: *A Dictionary of Tocharian B*. Amsterdam-Atlanta. 1999.

Adams 2013

Adams, D. Q.: *A Dictionary of Tocharian B: Revised and Greatly Enlarged.* New York–Amsterdam, 2013

Alt. Baskakov, N. A. / Toščakova, T. M.: *Oyrotsko-russkiy slovar*'. Moskva, 1947.

AltGND. Gürsoy-Naskali, E. / Duranlı, M.: *Altayca-Türkçe Sözlük*. Ankara, 1999.

Aydemir 2009

Aydemir, H.: Bemerkungen zu den tocharisch-türkischen und tocharisch-uigurischen Beziehungen: türk. *twqry*, sogd. *twγr'k*, pers. *t(u)ġr(a)k* [Toharca-Türkçe ve Toharca-Uygurca ilişkileri üzerine notlar: Türkçe *twqry*, Soğutça *twγr'k*, Farsça *t(u)ġr(a)k*]. In: *Journal of Oriental and African Studies* 18. 2009:159-180.

Clauson 1972

Clauson, Sir G.: *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford, 1972.

DLT Dankoff, R. / Kelly, J.: *Maḥmūd al-Kāšġarī. Compendium of the Turkic Dialects (Dīwān Luγāt at-Turk).* 1-3. Edited and Translated with Introduction and Indices by R. Dankoff in collaboration with J. Kelly. Sources of Oriental Languages and Literatures. 7. Edited by Ş. Tekin / G. A. Tekin. Harvard University Printing Office. Cambridge, Mass., 1982-1985.

DS *Türkiye'de Halk Ağzından Derleme Sözlüğü*. 1-12. 2. Baskı, Ankara, 1993.

Erdal 1991

Erdal, M.: *Old Turkic Word Formation. A Functional Approach to the Lexicon*. 1-2. Wiesbaden, 1991.

Gabain 1950

Gabain, A. von: *Alttürkische Grammatik*. 2. verbesserte Auflage. Leipzig, 1950.

Hilmarsson 1996

Hilmarsson, J.: *Materials for a Tocharian Historical and Etymological Dictionary*. Reykjavík, 1996.

Lessing 1960

Lessing, F. D.: *Mongolian-English Dictionary*. 1960.

Radloff I

Radloff, W.: *Versuch eines Wörterbuches der Türk-Dialecte*. Band I. S. Petersburg, 1893.

Stachowski 1993

Stachowski, M.: *Geschicte des jakutischen Vokalismus*. Kraków, 1993.

Tekin 1995

Tekin, T.: *Türk Dillerinde Birincil Uzun Ünlüler*. İstanbul, 1995.

Van Windekens 1976

Van Windekens, A. J.: *Le Tokharien confronté avec les autres langues indo-européennes. Vol. I: La phonétique et le vocabulaire*. Louvain, 1976.