

Some Implications of the Carbon-14 Dating of Tocharian Manuscripts

Douglas Q. Adams
University of Idaho

Tatsushi Tamai has recently published a preliminary report on paleographic and carbon-14 data of Tocharian manuscripts (Tamai, 2005¹). While, as yet, few Tocharian manuscripts have been subject to carbon-14 dating (and those only from the Berlin collection²), even the small number of results have significant implications for our understanding of the history of the Tocharian languages. Nine Tocharian B manuscripts and four Tocharian A manuscripts were dated by carbon-14. Below are given the manuscript number, where it was found, the C-14 date, and Tamai's paleographic date.

MS	Location	C-14 Date	Paleographic Date
B-333	Ming-öy Qizil	AD 394-473	AD 400-500
B-240	Ming-öy Qizil	AD 428-524	AD 650-750
B-483	Ming-öy Qizil	AD 770-888	AD 900-1000
B-352	Qumtura	AD 679-776	AD 600-700
B-601	Kucha	AD 669-780	AD 600-700
B-558	Shorchuq	AD 339-424	AD 800-900
B-367	Murtuq	AD 737-773	AD 700-800
B-178	Sängim	AD 697-716	AD 600-700
B-296	Dakianus [Xocho]	AD 1178-1255	uncertain
A-694	Shorchuq	AD 669-780	AD 600-700
A-1018	Bäzäklik	AD 938-1022	AD 700-800
A-1030	Sängim	AD 985-1022	AD 700-800
A-1068	Dakianus [Xocho]	AD 925-998	AD 700-800

Two of Tamai's manuscripts need special discussion. He makes clear that the early carbon-14 date for B-558 found at

¹I am most grateful to Victor Mair for bringing Tamai's publication to my attention. The carbon-14 dating was done by P. M. Grootes at the Leibniz Labor für Altersbestimmung und Isotopen-forschung at Christian-Albrechts-Universität, Kiel.

²Only those manuscripts which had a sufficient unwritten area on an edge were possible candidates for carbon-14 testing.

Shorchuq is misleading. B-558 is a palimpsest. Originally used for the writing of a Chinese text, it was reused for the Tocharian B text. In this case the paleographic date is clearly superior for dating the Tocharian B text.³ On the other hand, the late dated B-296 shows errors of a sort that suggest to Tamai that it was written or copied by a non-native speaker of Tocharian B. Paleographically it is a one-of-a-kind manuscript whose ductus is unlike any other surviving Tocharian manuscript.

Leaving aside B-558 and B-296, the carbon-14 and paleographical dates of the Tocharian B manuscripts are largely concordant. The two that are discrepant have paleographical dates that are one or two centuries later than the carbon-14 dates. The situation with regard to the Tocharian A manuscripts is quite otherwise. Three out the four paleographical dates do not agree with the carbon-14 dates and they disagree by being one to two centuries *earlier* than the carbon-14 dates. Tamai suggests that the copyists of the Tocharian A manuscripts were graphically conservative vis-à-vis their Tocharian B colleagues. That explanation seems unlikely to me but I have no better suggestion. In any case, where they differ, I put more confidence on the carbon-14 dates than in paleographic ones.

As already mentioned, Tamai's findings have a number of implications for us. First, it is clear that, on the basis of the carbon-14 dates, Tocharian B was an attested written language from ca. 400 AD to ca. 1200 AD. Since the earliest manuscript (B-274) datable on paleographic grounds is paleographically older than the earliest manuscript datable by carbon-14⁴, we probably have written attestations from at the latest the mid 4th century. Furthermore, the earliest manuscripts we have show a completely formed writing system, fully adapted to Tocharian B. This fullness of development strongly suggests that the manuscripts we have do not give us a window into the earliest period of Tocharian B writing. Thus there must have been some period of prior development that preceded our earliest attested manuscripts. How long this prior period may have been is unknowable on the basis of current evidence, but I think we can assume, with a fair amount of confidence, that

³The 400 hundred years or so that passed between the original writing of the Chinese document and its reuse for a Tocharian B document is striking.

⁴B-333, dated 394-473.

Tocharian B was written at the latest shortly after the beginning of the 4th century AD.⁵ Our (admittedly fewer) carbon-14 dates for Tocharian A reveal a shorter period of attestation, ca. AD 700 to ca. AD 1000.⁶

Our “outlier” Tocharian B manuscript B-296 (carbon-14 dated to AD 1178-1255), while not compelling evidence that Tocharian B was a spoken language in the 12th and 13th centuries, *is* evidence that there was a “market” for Tocharian B manuscripts in the 12th and 13th centuries, whether the market was composed of Tocharian B speakers or speakers of other languages who knew Tocharian B only as a language of scholarship or liturgy.

Another surprise of the carbon-14 dates is the surprisingly late dates of the Tocharian A manuscripts. The Tocharian A manuscript discovered at Shorchuq is contemporary with the Tocharian B manuscript found there, but the three found in Turfan are later, by a couple hundred years, than the Tocharian B manuscripts found there (excepting, as always, the odd B-296). In the Turfan region we find Tocharian B attested in the eighth century, Tocharian A in the tenth and possibly eleventh centuries, and Tocharian B again (B-296) in the twelfth or thirteenth century; there is no evidence that the two languages ever overlapped.

The relationship between the two Tocharian languages has been the subject of considerable discussion. The evidence has often been read as saying that attested Tocharian A was a “dead” liturgical language in competition as a language of liturgy and/or scholarship with Tocharian B (so, for instance, Lane, 1966). The latter, by anyone’s account, was the spoken and administrative language of Kucha. However, the carbon-14 dating of the manuscripts do not support a simplistic explanation on the order of Tocharian A, the older language, replaced by Tocharian B, the younger language.⁷ Certainly

⁵Even if one takes the latest possible date for B-333, writing in Tocharian B would have started no later than the beginning of the 5th century and would be roughly contemporary with Gothic. In any case our earliest attested Tocharian B texts are contemporary with Kucha’s most famous literary native son, Kumārajīva (AD 344-413). The son of an Indian father and a Kuchean princess, Kumārajīva was well-known as a translator of Buddhist works into Chinese.

⁶And thus a century or so earlier than Old English or Old Irish.

⁷This is Lane’s position (1966:226-232). He supports his position by noting that there are several Tocharian A manuscripts from the Turfan area that are glossed with Uigur words and one Tocharian A manuscript that is glossed by

Tocharian A contains a large number of obvious borrowings from Tocharian B, some of them quite ordinary words such as *epe* ‘or.’⁸ It seems, therefore, that as a literary language Tocharian A grew up “in the shadow” so to speak of Tocharian B.⁹ As always in this discussion, our understanding of things is very much hampered by having no documents from the Qarashahr area except from Shorchuq. But the evidence we do have suggest a very different scenario from Lane’s: Tocharian A reduced to written form on the model of Tocharian B some half a millennium after Tocharian B became a written language, used side by side with Tocharian B in the Qarashahr area¹⁰ and replacing Tocharian B, at least for a while, in Turfan. That Tocharian A shows very little internal variation (evidence to some of its status of a dead language) may simply be, as suggested above, the result of its short period of attestation.

It is interesting to look at the distribution of carbon-14 datings in conjunction with the distribution of Winter’s (1955) dialect classification of Tocharian B texts. Winter examined the orthographic variation to be found in Tocharian B texts and in an elegant way was able to divide Tocharian B texts into three groups on the basis of their orthographic characteristics. On this basis Winter divides Tocharian B texts into three varieties. One, labeled “Central,” is found everywhere one finds

Tocharian B words. From Shorchuq come two Tocharian A manuscripts that show certain orthographic peculiarities otherwise associated with Tocharian B, i.e., <î> and <û> for stressed *-i-* and *-u-* rather than the usual <i> and <u> and final <ñc> rather than the usual <ñs>. Certainly these phenomena suggest that in both the Turfan and Qarashahr areas Tocharian A was used side by side with Tocharian B (and in Turfan side by side also with Uigur) but not that it was necessarily a dead language. One should note too that the Tocharian A manuscript glossed by Tocharian B in Turfan tells us that there were at least some Tocharian B speakers/users in Turfan during the “Tocharian A ascendancy.”

⁸Certainly there are borrowings in the opposite direction, but they are distinctly fewer. For the classical discussions of inter-Tocharian borrowings, see Winter 1961 and 1962a.

⁹That Tocharian A has a very different vocabulary of fundamental Buddhist contents than that of Tocharian B (Lane, 1966) certainly argues a good deal of cultural and linguistic independence of the two Tocharian-speaking communities during the formative period of the Buddhist missionary effort in Central Asia. Differences in Buddhist terminology that Lane noted include A *lyalypu*, B *yāmor* (Skt *karman-*), A *kārme*, B *empreṃ* (Skt *satya-*), A *kloṣ*, B *lakle* (Skt *duḥkha-*), A *pñi*, B *yarpo* (Skt *puṇya-*), A *mārkampal*, B *pelaikne* (Skt *dharm-*), A *śka-tampeyum*, B *śka-maiyya* (Skt *daśa-bala-*), A *plyaskeṃ*, B *ompalskoñne* (Skt *dhyāna-*, *samādhi-*).

¹⁰At least at Shorchuq.

Tocharian manuscripts and is the only kind found in the center, the Kingdom of Agni (that is, the Qarashahr area, more particularly Shorchuq), a second, the “Western,” in the Kingdom of Kucha (Kucha, Ming-öy Qizil, Qumtura), and an “Eastern,” in the Turfan Basin (Sängim, Murtuq, Dakianus [= Xocho])—and also in the west. In the absence of a way to reliably date any of the manuscripts at the time of his investigation, Winter saw his three varieties as geographical dialects.

However, when we look at the carbon-14 dates an interesting pattern emerges. The two earliest manuscripts Tamai discusses (B-274, B-333) are both written in what Winter labels the “western” dialect. The bulk of Tamai’s carbon-14 dated manuscripts are in Winter’s “central” dialect. Tamai’s evidence is compatible with seeing at least the western and central varieties as being chronological rather than geographical variants. The “western” variety would be older than the “central” one. That make sense, since the “western” one is characterized by what are, on anyone’s account, linguistic archaisms vis-à-vis the “central” variety (e.g., non-merger of the diphthongs *-eu-*, *-au-*, and *-ou-* [= uniform “central” *-au-*], preservation of stressed *-ä-* [= “central” *-a-*], preservation of unstressed *-ā-* [= “central” *-a-*]).

Unfortunately of those manuscripts showing Eastern characteristics only the problematic B-296 has been carbon-14 dated (AD 1178-1255). With much less assurance (‘one swallow does not a summer make’), one might make the case that “eastern” is another chronological division rather than a geographical one. Certainly the Winter’s “eastern” characteristics look to be linguistically more advanced than those that characterize the “central” variety.¹¹ If these chronological suppositions are correct, one might see the following “time and area” scenario:

¹¹For example “eastern” manuscripts show a number of consonant cluster simplifications such as *štw-* > *šw-*, *-nts* > *-(n)s*, *-ñc* > *-ś*, *-kt* > *-k*, *-stsi* > *-ssi*, *-śtsi* > *śsi*, *-lñe* > *-ñe*, they also show *-i-* in place of etymologically older *-ä-* in palatal environments and *-aun-* > *-om-* (Winter, 1955:217-222). Also we have examples of *-ni-* > *-ñi-* (Winter, 1962b:119). In all these cases the “eastern” variant is demonstrably innovative.

1. Early Tocharian B (ca. 400-600 AD) only in Kucha and environs
2. Middle Tocharian B (ca. 600-900 AD) everywhere in “attestation area”
3. Late Tocharian B (ca. 900-1100 AD) in Turfan and Kucha¹² and environs
4. Very Late Tocharian B (ca. 1100-1300) in Turfan (only B-296)

The chronological/geographical pattern of manuscript attestations would look like this:

	Western	Central	Eastern
early	b		
middle	b	b	b
late	b		
very late			b [only B-296]

It is important to note that the absence of late (and early?) manuscripts from the center (Agni/Qarashahr) may be accidental. We have manuscript finds only from one place in the center, Shorchuq. That monastery may simply not have been in existence or active in the early and/or later periods. It may be significant, though exactly of what is unclear, that there is one of Winter’s “eastern” manuscripts (and thus Late Tocharian B by the hypothesis suggested here), namely B-486, found in the west but which is dated to a regnal year of what is apparently a king of Agni.¹³

If we add Tocharian A carbon-14 dated manuscripts to our chronological/geographical pattern above (a = Tocharian A), we get:

¹²Manuscripts with “eastern” characteristics are to be found in the western area (Winter’s “MQ IX group”) but Winter takes them as an importation from the east (either manuscripts imported from the east or manuscripts written by easterners resident in western monasteries). Given the “chronological hypothesis” made creditable by the carbon-14 dated “eastern” manuscript, the MQ IX group simply reflects Late Tocharian B written in the west.

¹³In B-486 the king’s name is *Kemarcūne*. Presumably it is the same king whose name appears in the genitive as *Kṣemarjūni* in a graffito at Subeshi (Pinault, 1986). In other contexts the formative *-arjuna-* is restricted to names of Agnean kings, hence the supposition that this king *Kṣemarjūna* is a king of Agni as well.

	Western	Central	Eastern
early	b		
middle	b	b,a	b
late	b		a
very late			b

Again, the absence of anything (early or) late in Agni/Qarashahr may be accidental; the absence of anything earlier than “late” in Turfan is probably not.

If the chronological hypothesis is correct, it is noteworthy that Tocharian B changed quite slowly over the approximately 600 year period of its attestation.¹⁴ This slow rate of change contrasts with that of Khotanese which changed very much indeed in the same period, going from a reasonably conservative Middle Iranian type to a very early exemplar of a New Iranian type.¹⁵ As we have noted, Tocharian A changed virtually not at all over its period of attestation but, since that period is only some three hundred years long, that constancy may not be significant of much.

Aside from the increase of what we might call the “internal knowledge” of the Tocharian languages, our new ability to date Tocharian texts presents us with two wider Indo-European implications. The first is that Tocharian B is a “language with a history.” Our attested texts are probably not chronologically undifferentiated and thus not every form found therein is necessarily equally old. A case in point is the third person plural copula *stare*. Watkins (1969:90-91) has taken *stare* to reflect PIE **sth₂-ró* and thus the quasi-equivalent of Sanskrit *ásthiran*. The identification is both phonologically and morphologically plausible. However, Winter (1959:236) has shown that *stare* is restricted to “eastern” texts. But since “eastern” is probably actually late Tocharian B, it is unlikely that *stare* is of Proto-Indo-European age; rather, it must be an inner Tocharian development.¹⁶ Our ability to give relative dates to Tocharian B

¹⁴I leave out of account B-296.

¹⁵One might compare the slow-changing Welsh vs. the fast-changing Irish in much the same time period.

¹⁶Another morphological difference setting of “eastern” texts from others is that feminine plural ending *-lyana* to gerunds. In non-eastern texts we have instead *-llona* (Winter, 1955:224). This distinction is presumably another later versus earlier form. Significantly, the putatively older Tocharian B form *-llona* agrees exactly with Tocharian A *-lam* whereas the Tocharian B *-lyana* has no

texts may not have as dramatic an impact on our understanding of the language as a similar ability in Hittite, but it is sure to change that understanding somewhat.¹⁷ Not only does Tocharian have a history, but that history begins earlier than has been commonly thought. As noted above, our earliest Tocharian B texts are contemporary with Gothic or even Runic and not with Old English, say, as has been previously supposed.

A second issue of interest to Indo-Europeanists in general is the date of separation of Tocharian A and B. Dating a proto-language is a notoriously uncertain and subjective undertaking. But, given the degree of divergence of attested A and B, there could be no serious disagreement that that divergence must reflect at least a minimum of 500 years before the earliest attestation.¹⁸ Since, thanks to Tamai, we now know that Tocharian B is attested at least from the fifth century AD, the latest possible date for the dissolution of the Proto-Tocharian community would be in the last centuries before Christ. However, if the rate of change seen in attested Tocharian B is roughly the same as the rate that prevailed in the preceding, unattested period, then the date of separation from Tocharian A is at least twice as long. Moreover, since the two languages were spoken side by side, that means they were subject to mutual, post-separation, influence and that influence is, if anything, going to artificially foreshorten the apparent date of separation. All things considered, a date for the dissolution of Proto-Tocharian in the mid to late second millennium BC seems the best estimate: we may think of Proto-Tocharian as the

Tocharian A counterpart. In this case, however, neither represents anything Proto-Indo-European but are both analogical developments within Tocharian, *-llona* of Proto-Tocharian date, *-lyana* a development of later Tocharian B.

¹⁷I looked for instance to see whether there might be a chronological distribution of dual forms. It turns out that verbal dual forms are restricted to “western” and “central” texts but, since there are only five verbal forms attested in Tocharian B texts *in toto*, the lack of any such in “eastern” texts is probably not significant. Certainly there are nominal dual forms in “eastern” texts, e.g., *komñiktene* ‘the two sun gods’ with characteristic “eastern”, i.e. late, phonological features (*-om-* for *-aun-* and *-i-* for *-ä-* next to a palatalized consonant). However, ‘two’ + a dual is confined to “western” and “central” texts while ‘two’ + a plural is found everywhere. Again, the number of forms (three of ‘two’ + a dual vs. thirty-eight of ‘two’ + a plural) involved may make the lack of ‘two’ + a dual in “eastern” texts not significant. Still, the totality of the evidence could be read as saying that the dual was in retreat during the late period of Tocharian B. For dual forms of the noun, see Winter, 1962b.

¹⁸Lane (1966:232) suggests the separation “might be anywhere from five hundred to a thousand years, certainly not less than five hundred.”

approximate contemporary of Mycenaean Greek.

In sum, we can say that, being such fundamental items, the presence of a few dates has a substantial importance for our understanding of Tocharian all out of proportion to their fewness. Naturally all this makes one greedy for more, and one may hope that both the carbon-14 and paleographic dating will soon be extended to the manuscript collections in London and Paris.

References

Lane, George Sherman

- 1966 On the Interrelationships of the Tocharian Dialects. *The Ancient Indo-European Dialects*, edd. Henrik Birnbaum and Jann Puhvel, pp. 213-233. Berkeley, University of California Press.

Pinault, Georges-Jean

- 1986 Epigraphie koutchéenne. In *Sites divers de la région koutchéenne*, pp. 61-196, edd. Chao Huashan, Simone Gautier, Monique Maillard, and Georges Pinault. (Mission Paul Pelliot, Documents Archéologiques, VII.) Paris, Collège de France.

Tamai, Tatsushi

- 2005 Paläographische Untersuchung und ¹⁴C-Prüfung. Digitalisierung der chinesischen, tibetischen, syrischen und Sanskrit-Texte der Berliner Turfansammlung, 02. O6. 2005 - Berlin.
<http://www.bbaw.de/bbaw/Forschung/Forschungsprojekte/turfanforschung/bilder/Tamai.pdf>

Watkins, Calvert

- 1969 *Indogermanische Grammatik, Band III.1: Formenlehre, Geschichte der Indogermanischen Verbalflexion*. Heidelberg, Winter.

Winter, Werner

- 1955 A Linguistic Classification of 'Tocharian B' Texts. *Journal of the American Oriental Society* 75:216-225.
1959 Zur Dialektgliederung von 'Tocharisch' B. *Zeitschrift für Vergleichenden Sprachforschung* 75:233-237.
1961 Lexical Interchange between Tocharian A and B. *Journal of the American Oriental Society* 81:271-280.
1962a Further Evidence of Inter-Tocharian Lexical Borrowing. *Journal of the American Oriental Society* 82:71-73.
1962b Nominal and Pronominal Dual in Tocharian. *Language* 38:111-134.