

REVIEW OF JAN HENRIK HOLST, *PROPOSALS FOR THE DIACHRONIC INTERPRETATION OF TOCHARIAN*, DÜREN (SHAKER VERLAG) 2025, ISBN 978-3-8191-0155-7, 282 PAGES, 32.80 €.

CARSTEN PEUST, KONSTANZ

Among the twelve branches of the Indo-European language family, Tocharian (which is a shorthand term for two related languages Tocharian A and B) is certainly the least known and studied. This is not so much because Tocharian would be particularly irrelevant to Indo-European studies, but rather due to extralinguistic accidents: The language came to scholars' attention only late; nobody comes with Tocharian experience from home or from schooling; the texts, often translations of Buddhist content, are hard to digest for those familiar with the more classical Indo-European languages. This implies that the Tocharian evidence for reconstructing Proto-Indo-European (henceforth: PIE) is most likely underrated. In turn, this suggests that research on Tocharian should be particularly promising for providing new insights into the reconstruction of PIE. We should therefore welcome any new and accessible piece of literature—whether of scientific or didactic nature—that encourages working with this language. Such a specimen of an up-to-date, accessible linguistic contribution about Tocharian is found in the book reviewed here. This reviewer must confess that he does not read it as a Tocharian specialist, but only as a linguist interested in historical linguistics and Indo-European studies.

After an introductory chapter, Holst addresses—in this order—a number of issues of historical phonology (covering both shallow and deeper diachronic levels down to PIE), topics of morphology and syntax, and finishes with a section on the genetic and areal position of Tocharian. However, it is not a grammar book; rather, it singles out a number of more or less independent issues about which the author has something interesting to say.

Going through Holst's issues of historical phonology, I skip over his treatment of a number of consonantal correspondences between the two Tocharian varieties A and B, moving on directly to the so-called “dental problem” discussed in section 3.1, which, as we will see, has a considerable relevance for Indo-European studies in general. PIE is usually reconstructed with three series of stops, traditionally described as voiceless, voiced and voiced-aspirate, e.g. */k/, */g/, */gh/. It is also known that these three series merge into one in Tocharian at least for the velar and labial places of articulation. The only potential exception is */d/. While both */t/ and */dh/ clearly merge into Toch. /t/, the outcome of */d/ has been subject to dispute. Holst singles out and evaluates all

the different and contradictory proposals that have been made on the issue and begins with a newly compiled dataset of 87 etymologies that is large and as reliable as possible at the same time. The clear conclusion emerges that a development PIE */d/ > Toch. Ø (zero), initially proposed by W. Winter in 1962 (though without a satisfying explanation), is the one to be regarded as correct. One example out of many is Toch. A *ākār* B *akrūna* (pl.) = Greek *δάκρυ* “tear(s)”.

Competing with the traditional reconstruction of the PIE stop system (*/t/, */d/, */dh/), there is the “glottalic theory”, which reconstructs the same stops as */t/, */t'/ and */d/ respectively, mainly for reasons of typological naturalness. Holst sympathizes with this theory, which happens to align well with the zero representation of the aforementioned dental. According to the glottalic interpretation, the lost sound is */t'/ rather than */d/, which amounts to a process known as debuccalization. This interpretation also allows for grouping the surviving dentals (now conceptualized as */t/ and */d/) into a natural class “non-glottalic”. As acknowledged by Holst, the linking of the change */t'/ > Ø with the glottalic theory was briefly hinted at already by M. Peyrot in 2022. However, it is only in Holst’s book that the argument is exposed in full.

A second issue surrounding the “dental problem” is the frequent shift of the PIE dentals to /ts/ or /c/. Holst confirms a previous view according to which there were two stages of palatalization of the dental stops before front vowels, which led to *ts* (first palatalization) and *c* (second palatalization), respectively. For example, in Toch. A *pats* B *pets* = Greek *πόσις* (< **pótis*) “spouse”, the first palatalization *t* > *ts* occurred before the front vowel *-i-, which was later lost in Toch. In A *pācar* B *pācer* = Greek *πατήρ* “father”, by contrast, the first palatalization was suppressed by analogy with inflected forms of the type **patr*-, in which no front vowel was present. The dental stop therefore survived as such for a while before undergoing the second palatalization, *t* > *c*, at a later stage. This chronology should open an opportunity to also integrate the relative dating of various vowel shifts into this network, which still largely remains a task for the future.

Holst now finds that PIE */t'/ (= traditionally /d/), though regularly yielding zero, could be affected by the first palatalization (as in AB *tsār*- = Greek *δέρω* “to separate”), but not by the second: */t'/ never yields *c*. He explains this observation with the following chronological scenario: (1) In an early stage of Toch., when PIE */t'/ was still a stop, it could undergo the first palatalization. (2) Subsequently, the reflex of PIE */t'/ was lost. (3) The second palatalization took place after that, which is why */t'/ could no longer be affected. Careful handling of relative chronology and clear separation of successive diachronic stages of development and logically deriving one from another is one of the methodological strengths found here and elsewhere in this book.

As Holst says, “the glottalic theory fits the interpretation of the data in an excellent way. However, of course the decision whether the glottalic theory should be adopted for PIE should not be based solely on Tocharian and its dental problem” (p. 71). So he still remains cautious about the extent to which this evidence should be used to draw far-reaching conclusions. I personally would be a little braver and claim that Toch. makes a very strong case for the glottalic theory being preferable to the traditional reconstruction of PIE. It is almost impossible to imagine how a language could shift *d* to Ø, but *dh* to *t*, in all positions including word-initially, which is what the traditional reconstruction of PIE presupposes. Only in light of the glottalic theory does the loss of

this specific dental become a plausible sound change. When Winter discovered the development PIE */d/ (= /t'/) > Toch. Ø in 1962, the glottalic theory had not yet been proposed, which is why his idea did not gain widespread acceptance.

Tocharian offers yet another piece of evidence in support of the glottalic theory that was not seen or not claimed by Holst. He says (p. 69): “Why among the ejectives precisely the dental was affected in Tocharian in such a way [i.e. was lost], and not a different plosive, is difficult to tell.” An answer to this question will be proposed here. Departing from the glottalic interpretation of PIE, there was a system of ejectives including *t' and *k' (leaving aside the problem of the number of velars), but no *p', a position that, as is well known, seems to have been a gap within the PIE consonant system. In Toch., the first palatalization added a *ts', but while *t' was lost, *ts' was not. It probably coincided with the non-glottalized ts only much later. Thus, we have to reconstruct a stage at which Toch. possessed ts' and (one or more) k', but no *p' or *t'. This is consistent with a well-known preference of ejectives for back places of articulation. Cross-linguistically, k' is more common than t', and t' is more common than p'. (Note that voiced stops, on the contrary, have a preference for front places of articulation, which is why some languages, such as Arabic, have b and d, but no g.) Just as Toch. in the early stage just mentioned, Hausa is a language that has k' and ts', but no *t' or *p'. I believe this tendency is the reason why specifically t' (and not ts' or k') was debuccalized. Conversely, the debuccalization specifically of a dental stop provides even more support for the glottalic interpretation of early Toch. and, consequently, PIE. With a voiced series, a *g would have been more likely to be lost or to be shifted to something else. We can see that Toch., somewhat surprisingly, helps bring one of the core problems of PIE phonology close to a definite conclusion.

While several individual components surrounding the “dental problem” had been proposed by one or the other scholar before, we have to credit Holst for weighing all the evidence based on a large set of etymologies coupled with typological plausibility, sorting out the numerous misleading views, and, in this reviewer's opinion, having finally reached solid ground concerning the various issues that have plagued this problem in the past. And as I have tried to show, the conclusions have implications not only for the narrow field of Tocharian, but far beyond.

In the subsequent section 3.2, Holst examines a number of issues related to the history of Tocharian vocalism, a field which is still poorly understood in several respects. The two varieties A and B often diverge in their vowels, which is a pain and a blessing at the same time. Both inherited apophony and younger umlaut phenomena complicate the picture. Holst focuses on cases in which previous scholars have diagnosed different vocalic outputs from a single PIE source. The common denominator is that he links the different outputs to the presence or absence of PIE word accent. This is again accomplished by assembling a number of relevant cognate sets.

Holst adopts a minority view brought up by N. van Brock in 1977, who proposed that PIE *ū yields u under stress but /ə/ elsewhere in Tocharian (where /ə/ is represented as ä in A, ä or a in B, or may become zero in certain environments). It is clarified that the stress of PIE, rather than historical Tocharian, is meant; the latter has shifted in a number of cases, as far as we can tell. (Toch. stress is not marked directly in writing but can be inferred to some extent from the vowel

correspondences between the language varieties A and B.) Holst then extends the same idea to the development of PIE **ī*. I will quote only one of his examples each:

- stressed **ī̇*: A *trit* B *trite* = Greek *τρίτος* “third”
 unstressed **ī̇*: A *wäs* B *wase* (i.e. Proto-Toch. **wāsæ*) = Sanskrit *viṣá-* “poison”
 stressed **ū̇*: A *-us* B *-usa* = Greek *-υῖα* (< **-úsja*) (participle suffix)
 unstressed **ū̇*: A *rtär* B *ratre* = Greek *ἐρυθρός* “red” (i.e. Proto-Toch. **rətræ* with *ə*,
 reflected as *a* in B and metathesized in A.)

From a typological perspective, it seems to make a lot of sense for former closed vowels /ī/ and /ū/ to be reduced to /ə/ when unstressed. Similar sound laws are well known from earlier Slavonic, Ethio-Semitic, Arabic dialects and many other languages. In a similar way, the reflex of PIE **ā*, which has been much disputed, is suggested by Holst to be A *a* B *o* under stress but AB *ā* elsewhere. Holst provides eight examples for each case, from which I quote one each:

- A *pracar* B *procer* = Sanskrit *bhrātā* = Greek *φράτηρ* “brother”
 A *mācar* B *mācer* = Sanskrit *mātā* = Greek *μήτηρ* “mother”

Note that “brother” may be almost the only PIE word whose stress can be confirmed from no less than four different branches: (Vedic) Sanskrit, Greek, Germanic (through the non-application of Verner’s law), and—as we learn now—Tocharian. As for “mother”, the evidence for the PIE stress is conflicting, but I side with Holst in assuming that the Toch. form likely goes back to a predecessor **mātēr*. I skip over Holst’s attempt to explain also the diverging reflexes of PIE **ē* based on the accent position, where the evidence is more fragmentary. The most relevant result for the general public is that, through the application of Holst’s vowel laws, Tocharian can provide evidence for the PIE stress position of many lexical items. This is very welcome since PIE stress has only been retained in a few branches. Vedic offers the fullest attestation, followed by Greek (with neutralizations particularly in verbs). In the third position in terms of relevance, one has hitherto seen Germanic, where the time-honoured Verner’s law offers indications of stress for a limited number of words, though marred by some conflicting “Verner doublets”. Tocharian has now jumped, as it seems to this reviewer, into a position of approximately the same importance as Germanic, which is an unexpectedly high rank. Holst himself has barely begun to exploit the Toch. evidence for the purpose of reconstructing PIE, which should be a promising field for future research. The other branches (briefly mentioned by Holst on p. 104) contribute little to nothing to our knowledge of the PIE word accent. (His Armenian evidence is interesting but not overly informative, since according to Holst, Armenian forms a phylogenetic subgroup with Greek, a language that retains a much better representation of the accent.)

The following chapters, 4 and 5, cover a number of original suggestions regarding grammaticalization processes in Tocharian. Explanations are provided for the etymology of various grammatical items including the conjunction B *wai* “and”, the comitative suffix B *-mpa*, the verb A *nas-* B *nes-* “to be” and the verbs of position AB *šām-* “to sit” and A *štām-* B *stām-* “to stand”.

These suggestions all seem plausible to me, though not always strictly provable, and are underpinned with a number of typological or areal parallels in each case. I would like to highlight

two subchapters here, beginning with section 4.3, which deals with “The gender distinction in the 1st pers. sg.” Tocharian A is one of the few languages worldwide that distinguish gender in the 1st sg. personal pronoun. The forms are masc. *nās*, fem. *ñuk*. Tocharian B has a common form *ñās* for both genders. The etymology of the feminine form *ñuk* has remained disputed. Building upon a suggestion made by J. Schindler in 1967, Holst derives this pronoun from a pronominal base **mene-*, otherwise only found in Slavonic and Indo-Iranian (a fact that this reviewer would like to consider as a hint as to the genetic position of Tocharian), combined with a relative pronoun **k^wis* (masc.) ~ **k^wā* (fem.). He explains this as having evolved from a former cleft sentence “(it is) me who ...”, which evolved into an emphatic (focalizing) personal pronoun and hence into the regular personal pronoun. As a support of this theory, I would like to mention that a very close parallel to what Holst reconstructs for Tocharian has been suggested for Laal, a language spoken in Chad. Like Tocharian, Laal is among the few languages that possess a gendered 1st sg. personal pronoun: *já* masc., *jí* fem. This pronoun is homonymous with the relative marker in Laal. R. Voigt concluded from this fact in *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 136 (1986), p. 185 that the personal pronoun arose from a periphrasis *v-já* / *v-jí* “(it is) me who ...”, where the original pronominal stem was a vowel (*v-*) that was lost through phonetic erosion, so that all that remained from the cleft construction was the former relative marker.

As a further typological parallel, I would like to draw attention to the Proto-Semitic personal pronouns **anāku* “I” and **anta* masc., **anti* fem. “you”, which evidently suggest a segmentation into *an-āku*, *an-ta*, *an-ti*. The element *an-* must originally have been a focalizing morpheme, as revealed by Ancient Egyptian, where *ʔn-* is still the productive focalizer before any noun phrase and is also contained in the personal pronoun *ʔnk* “I” (presumably $\approx$ *anāku*, vowels are not written in this language). In fact, there are many more languages in which some or all of the personal pronouns share a common element, which at least in some of the cases may indicate a history of grammaticalization from focalizers or cleft periphrases. I would like to briefly suggest some potential candidates with speculative segmentation, admitting that I have not conducted detailed research in each case. All of the following examples cite the pronouns “I” and “you (thou)” respectively: Sanskrit *ah-am*, *tv-am*; Welsh *f-i*, *t-i*; Finnish *mi-nä*, *si-nä*; Turkish *b-en*, *s-en*; Chechen *s-o*, *h-o*; Kannada (Dravidic) *nā-nu*, *nī-nu*; Somali *an-iga*, *ad-iga*; Margi (Chadic) *nà-yà*, *nà-gà*; Bemba (Bantu) *i-né*, *i-wé*; Zulu (Bantu) *mì-ná*, *wè-ná*.

The last grammaticalization topic to be discussed here is located on the borderline between morphology and syntax. It is Holst’s explanation for the genesis of the infinitive (section 4.6), whose marker is AB *-tsi* (= Balto-Slavic *-ti* with the first Tocharian palatalization applied, see above). This suffix attaches to the present stem of the verb in A, but to the subjunctive stem in B. The latter formation is typologically unexpected and has thus far defied explanation. Holst suggests that at an earlier stage of Tocharian, speakers employed a periphrasis built upon a finite subjunctive verb as in Albanian *ai do të shkojë* “he wants to go”, lit. “he wants that he go (subjunctive)”. Holst describes how such a construction could have given rise to a verbal form interpretable as an infinitive but retaining traces of subjunctive morphology. This can be the background of the Toch. B peculiarity, whereas Toch. A has restored its infinitive formation towards a

more common typological state. This scenario, though by necessity somewhat speculative, seems convincing to the reviewer. Holst also reminds us (p. 144) that PIE may not have possessed an infinitive at all, noting that no common infinitive morphology can be reconstructed for the proto-language. The idea that PIE may have employed a finite construction similarly to modern Balkan languages is a remarkable assumption about PIE syntax that probably no one has ever considered. It is striking how a neglected language such as Tocharian can help to establish far-reaching hypotheses about PIE syntax.

The final chapters of Holst's book address the areal and genetic position of Tocharian. Holst argues that contact with an early Turkic language probably shaped some of the language's typological traits. I will not go into the details here except for one point made by Holst that I consider perhaps the weakest. This point concerns the presence of only one plosive type, which Holst relates to the (allegedly) limited distinction between voiced and voiceless plosives in many Turkic languages (section 6.2.2). But we saw above that earlier Tocharian had two plosive types, plain and ejective. This fact does not seem to have been coordinated with his statement (p. 193) that Tocharian speakers may have arrived with a contrast of the *t-d* type when they encountered a language with only one plosive type from the Turkic stock. I would like to suggest a slightly different scenario along the following lines: The Tocharians arrived in their later homeland with a contrast of plain and ejective plosives. The areal contact languages—whether Turkic or other—did not use ejectives but other types of plosive distinctions (not necessarily an absence of plosive types altogether). This could have sufficed to break down the inherited binary contrast. I am reminded of the German dialect area, where a northern zone with a voiceless–voiced distinction of plosives and a southern zone with a fortis–lenis distinction are separated by a central contact zone in which the distinction of plosive types has been abandoned altogether.

The final question regarding the genetic position is, naturally, where Tocharian fits into the Indo-European family tree. It is an enigma that, despite two centuries of intense study of the Indo-European stock, its internal subgrouping has not yet been established. Much of the earlier research was misled by a dichotomy between *kentum* and *satem* languages, which has proven to be fruitless. More recently, various highly technical approaches using sophisticated phylogenetic methods have emerged, but so far without reaching a consensus. Note also that despite their high level of sophistication, they inherently do not distinguish between common innovations and common retentions. Holst discusses the genetic position of Tocharian within the family tree in section 7.2. While remaining cautious on this point, he leans towards grouping Tocharian with Anatolian. On this occasion, I would like to advocate for a simple classification method that assembles lexical isoglosses of the basic vocabulary and turns them into fragments of an unrooted family tree. Consider the following representations of the verb “to lie down” (incidentally, this Tocharian verb is also the subject of a specific discussion in section 5.3.1 of Holst's book):

“to lie down”: Tocharian AB *lyäk-* = Gothic *ligan* ≠ Hittite *ki-* = Sanskrit *śay-*

We find one lexical representation in Tocharian + Gothic and another one in Hittite + Sanskrit. This must not be taken as straightforward evidence of two genetic subgroups { Tocharian +

Gothic } and { Hittite + Sanskrit }, because one of the two terms is most probably a shared retention, which is not probative in terms of genetic subgrouping. However, (at least) one term should represent a shared innovation, if we assume (1) that borrowing is improbable for such a basic vocabulary item and (2) that two independent lexical replacements are unlikely to produce the same outcome. This isogloss does not allow us to decide on which side to localize the innovation or the retention, but we can conclude that at least this single lexeme is a piece of counterevidence against grouping Tocharian and Hittite together (or if we do, we must argue for either borrowing or independent but converging lexical replacements). I only use this single lexeme to exemplify the method. Many more examples would, of course, be needed to build a robust argument. I took this approach in a paper titled “The subgrouping of East Chadic” in *Folia Orientalia* 55 (2018), to which any interested readers are referred.

As we have seen, Holst’s book brings up a number of solutions of long-standing riddles concerning Tocharian, based either on arguments forwarded by him for the first time, or dug out from footnotes or from highly specialized literature where they are likely to go unnoticed by general linguists. This accessible and well-written book, which is full of original ideas, makes the reader want to learn more about the underrated language that is Tocharian. The reader is made aware of how many basic questions about this language have still remained unsettled. The argument repeatedly extends into the realm of other languages, offering remarkable typological insights and suggestions of general interest. Holst’s insightful piece of work can be recommended not only to every Tocharian specialist, but also to anyone who wants to broaden their knowledge of Indo-European in general, or even as a source of inspiration for linguists working on reconstruction or typology in quite different parts of the world.