Introduction to Tocharian

Ronald I. Kim
Adam Mickiewicz University
ronald.kim@yahoo.com

1. Introduction

1.1. Historical background and external history

1.1.1. Discovery and documentation

Tocharian is the conventional name for two related, extinct Indo-European (IE) languages known from documents found in the oases of the Tarim River basin north of the Taklamakan desert in Xinjiang (Chinese Turkestan).

general term | German alternatives | obsolete (see below)
---|---|---
Tocharian A (TA) | Osttocharisch (East Tocharian) | Turfanian
Tocharian B (TB) | Westtocharisch (West Tocharian) | Kuchen
A few Tocharian manuscripts were acquired by the Russian consuls Berezovskij and Petrovskij in Kašgar already in the 1890s, but the great bulk were uncovered by archaeological "missions" to Xinjiang in the years preceding World War I, led by

- Britain: Sir Aurel Stein
- Germany (Prussia): Albert von le Coq
- France: Paul Pelliot
- Japan: Count Ōtani Kōzui

In addition to a wealth of Middle Iranian documents, the expeditions brought back others in unknown languages, written in the “slanting” Brähmī script of Central Asia. In 1908, the German philologists Emil Sieg and Wilhelm Siegling conclusively identified them as non-Indo-Iranian IE languages, which they labeled “Indo-Scythian”; they also succeeded in distinguishing TA and TB. These Western missions to China ceased with the outbreak of the war; other than Langdon Warner of Harvard University (U.S.), who led an expedition in the early 1920s, no Western scholars are known to have visited Xinjiang for the next few decades.

Current distribution of Tocharian manuscripts and state of publication (for details, see Malzahn 2007a, Pinault 2007):

**Berlin**
- All TA manuscripts (except for Yanqi fragments): published in Sieg and Siegling 1921
- TB manuscripts (Buddhist and secular): published in Sieg and Siegling 1949-53

**Paris (Pinault 2007)**
- TB manuscripts, mostly medicinal and magical texts: published by Lévi and Filliozat, but unreliable without Sieg’s extensive commentary and list of corrected readings
- TB secular documents from the Mission Paul Pelliot: caravan passes and cave graffiti published in Pinault 1987; other manuscripts in Bibliothèque Nationale being published by Pinault in a long series of articles
London

- TB manuscripts in Weber, Stein, and Hoernle collections: published in Broomhead 1962 with translation, but without photographs (also no indication of provenience); new edition by Peyrot 2007, concordance in Peyrot 2007a

St. Petersburg

- Petrovskij and Berezovskij collections, now being published by Burlak and Itkin, Pinault, Klaus T. Schmidt

Tokyo

- manuscripts from the Ōtani collection, never systematically published

Ürümqi

- TA fragments from Yanqi (YQ), discovered in Xinjiang in the 1970s: published in Ji et al. 1998

The manuscripts are dated to approximately the 6th to 8th cc. AD, but further chronological precision is difficult.

- The TA records were discovered in and around Turfan and Qarašahr and are entirely of Buddhist religious content; most are translations and/or adaptations of Sanskrit originals.

- TB documents were found across the northern Silk Road from Kuča in the west to Turfan in the east; the older name “Kuchean” (Fr. koutchéen) is thus inaccurate and best avoided. Most are Buddhist in content, but there is an important Manichean bilingual fragment in TB and Old Turkic (von Gabain and Winter 1958; see §2.2) which shows that that TB speakers also played a role in the diffusion of Manicheanism. A large number of secular records indicate that TB was the vernacular of at least part of the population in these areas in the later 1st millennium: in addition to monastery records, commercial letters, there are caravan passes, cave graffiti, and a solitary (bad) love poem. On variation in TB, see §1.3.
1.1.2. Who were the Tocharians?

The speakers of Tocharian played an important role in the Buddhist civilization of pre-Islamic eastern Central Asia, but their exact identity remains unknown. The name “Tocharian” rests mainly on the form 忸ɣry in an Old Uyghur colophon, but both the reading and the identification have been challenged. It seems certain that the speakers of TA and TB were not the “Tocharians” of antiquity (Strabon’s Τόχαροι; Skt. Tukhāra-).

Among the figures in the spectacular Buddhist cave paintings of the region are some with red hair and green or blue eyes, and many have speculated that these were the Tocharians. More recently, the discovery of red-haired, “Western”-looking mummies in the Taklamakan made headlines in the Western press in the mid-1990s, and much sensationalism — with thinly veiled racial overtones — over what “white people” or “Caucasians” were “doing” in China. The entire study of this region has become sadly politicized, and results of archaeological research have been abused to support Uyghur separatism! This has led to tension between Western and Chinese historians and other scholars of Central Asia, a state of affairs decried by the late Chinese Indologist Ji Xianlin. For our purposes this debate is irrelevant: after all, we cannot be sure which language(s) the famed mummies spoke. (See also Thursday’s lecture.)

Although the exact historical context of the two Tocharian languages and their speakers is not fully known, certain conclusions may be drawn from the texts. The classic paper of Lane (1963) is still worth reading.

- TA is remarkably uniform linguistically, and a number of facts suggest that it was no longer spoken at the time of the surviving manuscripts, but served as a sort of liturgical language among speakers of TB and Old Turkic.
- Extensive differences in Buddhist terminology between TA and TB (e.g. TA märkampal vs. TB pelaikne ‘law, dharma’) indicate two different periods of Buddhist missionary activity in eastern Central Asia in early 1st millennium AD.
- Bilingual texts in TB and Sanskrit, one with inflectional forms of a noun; also TA texts with TB and Turkic glosses, TB texts with Turkic glosses
- The speakers of Tocharian (more precisely, TB) began to shift to Turkic in the later 1st millennium AD; the language was probably extinct by 1000.
1.2. History of Tocharian linguistic research

Early period (to the 1930s)

- The early pioneers in the study of Tocharian were SIEG and SIEGLING, along with Wilhelm SCHULZE in Germany; their publication of all the TA material in the Berlin collection in 1921 was followed by the classic Tocharische Grammatik (1931), still the first reference for TA descriptive grammar.
- The French Sanskritist Sylvain LÉVI (along with Antoine MEILLET) began publishing manuscripts in the Mission Paul Pelliot collection already before World War I, but their edition was rapidly surpassed in quality and quantity by the Berlin school. Along with Sieg and Siegling, Meillet drew the attention of Indo-Europeanists to a number of linguistic features of Tocharian.

Intermediate period (1930s to 1980s)

- The American Indo-Europeanist George Sherman LANE wrote several important articles on Tocharian historical phonology and morphology from the 1930s to the 1960s, and the great Danish Indo-Europeanist Holger PEDERSEN published the first historical grammar of Tocharian (Pedersen 1941).
- Prague’s own Pavel POUCHA published several studies; his Latin-language TA dictionary (Poucha 1955), although not perfect, was until recently the only lexicon of that language.
- Walter COUVREUR published the first comparative grammar of Tocharian to fully take TB into account (Couvreur 1947) and edited many Paris manuscripts, but stopped working on Tocharian after the 1960s.
- The late Werner THOMAS of Germany studied with Sieg and published literally thousands of pages of research on various philological minutiae of Tocharian for a full half century after World War II.
- Beginning in the 1950s, the late Werner WINTER published dozens of almost uniformly excellent studies of Tocharian philology and historical linguistics: most of these have now been collected in Winter 2005.
- The Belgian scholar Albert Joris VAN WINEKENS wrote innumerable articles and published an entire historical grammar and dictionary of Tocharian (van Windekens 1976-82), but his work is uncontrolled to say the least, and to be consulted only with the greatest caution.
Contemporary period (1970s to present)

Tocharian studies have experienced a remarkable upswing since 1980, and the volume of scholarly literature now rivals that on established Indo-European branches.

• Major Tocharian specialists of the past generation include Klaus T. Schmidt, †Peter Stumpf, Olav Hackstein, and Christiane Schaefer in Germany; †Jorundur Hilmarsson in Iceland; Douglas Q. Adams and Donald A. Ringe, Jr. in the U.S.; and Gerd Carling in Sweden and Svetlana Burlak in Russia.

• Georges-Jean Pinault (École Pratique des Hautes Études, Paris) has been steadily publishing manuscripts in the Fonds Pelliot Koutchéen in the Bibliothèque Nationale in Paris, along with an impressive series of of linguistic studies; see Pinault 1987 on the TB caravan passes and cave inscriptions, and Pinault 1989 and 2008, currently the best introductions to Tocharian grammar and texts.

• China’s late, revered Ji Xianlin (born a German colonial subject in Qingdao in 1911!) collaborated with Schmidt and Pinault on the publication of the TA Yanqi fragments (Ji et al. 1998).

• More recently, Melanie Malzahn (Vienna) has published many linguistic studies, including an absolutely indispensable monograph on the verb (Malzahn 2010).

• In addition to editing the Tocharian texts in the London collections (see §1.1), Michaël Peyrot has published an excellent study of variation in TB (Peyrot 2008) and a forthcoming monograph on the Tocharian subjunctive (see Peyrot 2010).

• Finally, leading Indo-Europeanists such as Jay Jasanoﬀ, Gert Klingenschmitt, Martin Kümmel, H. Craig Melchert, and Paul Widmer have made important contributions to Tocharian historical linguistics.

Krause and Thomas’s famed Tocharisches Elementarbuch (1960) contains a bibliography of all publications up to 1959; for the period 1960-1984, see Thomas 1985, with often idiosyncratic (and critical) commentary. Articles published since then are listed in the annual Bibliographie Linguistique; much, but not all recent literature is included in Pinault 2008. Malzahn 2007 is an invaluable guide to the (near-)current state of Tocharian philology and text publications.
1.3. General linguistic features

It cannot be overemphasized that TA and TB were distinct languages and certainly were not mutually intelligible. Nevertheless, TA and TB were structurally similar, characterized by right-headed constituent phrases, a system of agglutinating nominal case suffixes, and the central role of aspect and tense in verbal morphology. The two had doubtless been diverging for several centuries before the time of our documents, so that their latest reconstructible common ancestor, Proto-Tocharian (PT), must be dated to the last centuries BC.

The TB documents exhibit considerable variation on all levels.

• Based on certain phonological and morphological features, they have been divided into western, central, and eastern dialects (Winter 1955). However, vernacular TB sources (e.g. caravan passes or cave graffiti) mostly show “eastern” characteristics, so others argued that that these reflect chronological and/or sociolinguistic differences; see respectively Stumpf 1990 (originally written in 1977), Schmidt 1986.

• Peyrot (2008) is an exhaustive study of variation and change in TB, based on a much wider range of texts and improved knowledge of Tocharian paleography. It now appears that the variation in TB is primarily chronological, but naturally also conditioned by geographical and social factors. Classical TB was established as a literary language in Kucha by the sixth century, but continued to evolve under the influence of the spoken language; after the Chinese wars and destruction of the kingdom of Kucha in 648, literary activity shifted to Turfan, where scribes wrote in a language based largely on contemporary speech, and late documents form all locations show a numerous innovative features (R. Kim 2009c).

• Another source of variation is poetic: many forms in TB verse passages have been adjusted by one syllable to fit the meter, usually through optional but frequent deletion of stressed /ə́/ (§3.3.3). Also characteristic of verse is pudnäkte ‘Buddha’ for prose paňäkte.

2. Writing Systems

2.1. North Turkestan Brāhmī script

Virtually all of our Tocharian sources are written in a variety of Brāhmī script, traditionally called North Turkestan Brāhmī, which developed over the course of the 5th and 6th centuries AD from the Turkestan Gupta, itself brought from India. See the attached character

Vokale

| Einf. Vokale | a | ä | i | ü |
| Velare       | ka | ka | kha | ga | gha | ŋa |
| Palatale     | ca | cha | ja | jha | ŋa |
| Cerebrale    | ta | tha | āa | dha | na |
| Dentale      | ta | ta | tha | ān | dha | na | na |
| Labiale      | pa | pha | ba | bha | ma | ma |
| Halbvokale   | ya | r[a] | [a] | [a] | [a] |
| u. Liquiden  | ra | ra | la | la | va | wa |
| Zischlauten  | sa | s[a] | s[a] | s[a] | s[a] |
| Hauchlaut    | ha | (§ 5 Anm. 5) |
| Affrikata     | t[a] | [a] |

Notes
1. Vowel length probably not distinctive in either Tocharian language: thus ā[a] vs. a ([ə] or [ʌ]) vs. ā ([i]), with Sanskrit ā, a representing a qualitative contrast.
2. The script lacks symbols for distinctively Tocharian sounds such as labiovelar [kʷ] and (in TB) the diphthongs [ew], [ow], [aw], [ay].
3. As expected, there are signs (akṣaras) for Sanskrit and Prakrit phonemes absent in the language, e.g. v, h, and the voiced, aspirated, and retroflex obstruents; these are used almost exclusively in Indo-Aryan borrowings.

4. The most remarkable innovation of the Tocharian writing system is the creation of a series of Fremdzeichen (“foreign signs”) to represent sequences of consonant + the vowel ā, probably a high central [i]. These exist for most, but not all consonants, and are used apparently interchangeably with the normal akṣara plus two subscript dots (hence the transcriptions tā, nā, etc.).

5. A second vowel (usually u) can be combined with a ligature, e.g. kse + u, transcribed kšse; such “subscript” u’s may denote either a labiovelar or a reduced/syncopated vowel (underlyingly probably /kʷ/; §3.1.1, 1, §3.1.2, 2).

2.2. Manichean script

A single manuscript fragment in a Manichean (Aramaic-derived) script similar to that used for Manichean Sogdian was published by Werner Winter and the Turcologist Annemarie von Gabain in 1958. The language is heavily influenced by Buddhist terminology, which suggests that a (small) group of Manichean missionaries were operating within a largely Buddhist social context.

3. Phonology

3.1. Synchronic phonology

3.1.1. Phonemic inventory of Tocharian B

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels and diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/ /pʰ/ /t/</td>
<td>/i/ /u/</td>
</tr>
<tr>
<td>/c/ /k/ /kʰ/</td>
<td>/e/ /a/ /o/ /ew/ /oy/</td>
</tr>
<tr>
<td>/s/ /š/</td>
<td>/a/ /aw/ /ay/</td>
</tr>
<tr>
<td>/m/ /n/ /ň/</td>
<td>/l/ /ь/</td>
</tr>
<tr>
<td>/ř/</td>
<td>/w/ /y/</td>
</tr>
</tbody>
</table>
Notes


2. PT *pʰ is preserved in TB as a marginal phoneme /pʰ/ distinct from /p/; similarly for *mʰ, although I can find at most one example of TB /mʰ/. See §3.3.1, 2.

3. The phonemes /a/ and /ə̆/ had stress-conditioned allophones from classical TB onwards (§1.3), respectively stressed ā [a] ~ unstressed a [ə̆], and stressed a [ə̆] ~ unstressed ā [i] (closed syllables), Θ (open syllables). See §3.3.3 for examples.

4. The vowels /i/ and /u/ are underlingly preconsonantal /ə̆y/, /ə̆w/, and alternate with prevocalic iy, y /ə̆y/, uw, w /ə̆w/. Examples: TB pres. Cl. III lyuketə̆r ‘shines’, Cl. VIII lyukšə̆m /lə̆wɔk-/ ~ pret. lyauksa /lə̆wɔk-/ ‘illuminated’.

5. The diphthong /ə̆w/ merges with /aw/ in classical and late TB texts, e.g. present thematic 1sg. -eə̆ ~ -au /-ə̆w/ or kłyesə́ ~ kłyausə́ /kłə̆wsə́ / ‘hear’. In addition, word-final -oə̆ /-ə̆wə̆/ merges with -au in classical and late TB, e.g. tærkoə̆, tærkau /tə̆rkə̆wə̆ / ‘having left’. However, word-final -eə̆ /-ə̆wə̆/ is maintained in all dialects in e.g. keə̆ /kə̆wə̆ / ‘cow’. Note sporadic instances of e, o for ai, au in late and colloquial texts, suggesting incipient monophthongization.

3.1.2. Phonemic inventory of Tocharian A

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels and diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/ /t/ /c/ /k/ /kʰ/</td>
<td>/i/ /u/</td>
</tr>
<tr>
<td>/tʰ/ /ʃ/ /e/ /a/ [ə̆] /o/</td>
<td></td>
</tr>
<tr>
<td>/s/ /ʃ/ /ə̆/ [a]</td>
<td></td>
</tr>
<tr>
<td>/m/ /n/ /ŋ/ /n/</td>
<td></td>
</tr>
<tr>
<td>/l/ /ʃ/</td>
<td></td>
</tr>
<tr>
<td>/ɾ/</td>
<td></td>
</tr>
<tr>
<td>/w/ /y/</td>
<td></td>
</tr>
</tbody>
</table>

Notes

1. As a result of the pre-TA epenthesis of *ā in consonant clusters and syncope of *ā in open syllables, *ā (< PT *ə̆) is no longer phonemic: its appearance is entirely predictable from the
underlying consonants. This was first noticed by Jasanoff (1987; also Adams 1988), but its effects have not been properly appreciated. See below, §3.3.2, 8.

2. On the question whether to assume a labiovelar phoneme /kʷ/ or a reduced rounded vowel /ũ/, see Ringe 1998, R. Kim 1999. Note that if TA has no reduced “schwa” phoneme /ä/, it is extremely unlikely to have had a marked reduced vowel /ũ/. I therefore prefer to analyze p₂kāḷ, p₂kul, (1x) pkul and p₂kāṣ, p₂kuṣ ‘all (abl.)’ as underlying /pkʷl/, /pkʷ-ṣ/.  

3. The vowels [i], [u] are underlyingly /y/, /w/ between consonants (cf. TB /ɔy/, /ɔw/).

3.2. Historical phonology: Proto-Indo-European to Proto-Tocharian

Recent research has elucidated most of the principal phonological developments from Proto-Indo-European (PIE) to PT and the two Tocharian languages, although several minor problems remain. See especially Ringe 1996, with detailed discussion of the changes and their relative chronology (reliable with few exceptions, e.g. the section on PT *ō < post-PIE *-øy/ø-).

3.2.1. Phonemic inventory of Proto-Indo-European (PIE)

Consonants

<table>
<thead>
<tr>
<th>*p</th>
<th>*t</th>
<th>*k̂</th>
<th>*k̂</th>
<th>*m</th>
<th>*n</th>
</tr>
</thead>
<tbody>
<tr>
<td>*b</td>
<td>*d</td>
<td>*g̃</td>
<td>*g</td>
<td>*gʷ</td>
<td>*l</td>
</tr>
<tr>
<td>*bʰ</td>
<td>*dʰ</td>
<td>*gʰ</td>
<td>*gʰ</td>
<td>*gʰ</td>
<td>*r</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>s (</em>[z])</td>
<td></td>
<td></td>
<td>*w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*h₁, *h₂, *h₃</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vowels and diphthongs

<table>
<thead>
<tr>
<th>*i</th>
<th>*u</th>
<th>*m̥</th>
<th>*n̥</th>
</tr>
</thead>
<tbody>
<tr>
<td>*e</td>
<td>*o</td>
<td>*ey</td>
<td>*oy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*î  *ũ)

(*ē  *ō  *ẽ  *ōy  *éy  *ōy  (*ēw *ōw)
  *ā  (*āy  (*āw)

Notes

1. PIE definitely had three contrasting series of dorsal stops, which remain distinct in Luvian and Lycian: PIE *kʰe̞y- > Luv. ziy-ar(i) [t'-] ‘is lying (down)’, PIE *kes- > Luv. kišā(i)- ‘comb’, PIE *kʰid > Luv. kwiḍ [kʰ-] ‘what’ (Melchert 1987). Conventionally labeled “palatal”, “velar”, and “labiovelar”; more probably prevelar, postvelar, and labio-postvelar, as it is crosslinguistically rare for palatal stops to become velars.

2. On PIE “thorn” see Schindler 1977 and the discussion in Mayrhofer. TA tkam ‘earth’ [TB kem], like Hitt. tēkan, loc. tagān, tāgan, could have been influenced by nom./acc. *dʰêgʰ-ôm. TB taktsāntsa ‘expert, capable one’ has been connected with Ved. tākṣā, Gr. τέκτων ‘builder’ < *te-tkʰ-ôn-, but the equation is uncertain.

3. The phonetics of the three laryngeals remain unclear, though I find it extremely probable that *h₁ was a glottal stop [ʔ]. Possibly *h₂ [h], *h₃ [ʕ] (voicing, famously in *pi-ph₁-e/-o- > *pi-b-e/-o- > Ved. pibati, Old Irish ibid; rounding because of Greek reflex o).

4. The PIE sonorants (including the glides *y and *w) had syllabic allophones when followed by a word boundary or nonsyllabic; the rule operates iteratively from right to left, e.g. gen. *kʰun-ēs —> Ved. śūnas, Gr. κυνός vs. instr. pl. *kʰnθ-bʰís > Ved. śvabhís to */kʰnθ- ‘dog’. Numerous exceptions (e.g. nasal infix always nonsyllabic *[n], suffixes *-io- vs. *-yo-) suggest that *i, *n̥, etc. and *y, *n, etc. were distinctive already in PIE.

5. PIE had a vowel *a alongside (much) more common *e, in words such as *albʰós ‘white’ (Lat. albus; Hitt. albaš ‘cloud’, Gr. ἄλφος ‘white leprosy’); note also *ā ~ *a ablaut in *nās- ~ *nas- ‘nose’, *wāst-u ~ *wāst-u ‘settlement’. (Otherwise the Leiden school, e.g. Beekes 1995.)
3.2.2. Phonemic inventory of Proto-Tocharian (PT)

Consonants

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels and diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>*p *pʼ *t</td>
<td>*c *k *kʼ(*i) *ə (*u)</td>
</tr>
<tr>
<td>*tʼ</td>
<td>*o</td>
</tr>
<tr>
<td>*s *š *ś</td>
<td>*e *ē *o *oy *ew *ēw</td>
</tr>
<tr>
<td>*m *mʼ *n</td>
<td>*ē</td>
</tr>
<tr>
<td>*l *lʼ</td>
<td></td>
</tr>
<tr>
<td>*r</td>
<td></td>
</tr>
<tr>
<td>*w *wʼ</td>
<td>*y</td>
</tr>
</tbody>
</table>

3.2.3. Major changes from PIE to PT

3.2.3.1. Consonants

1. The PIE series of voiceless, voiced, and voiced aspirate stops have famously merged, except that *t, *dʰ > PT *t remained distinct from *d > PT *tʼ (Winter 1962).

   PIE (*demh₂- ~) *dmh₂- ‘build’ > PT *tʼōma- ‘grow’ > TB /tʼoma-/ (pres. Cl. III 3pl. tsmentār), TA ts(ā)mā- (pret. ptcp. tsmo)

   PIE (*der- ~) *dr- ‘split’ --> *tʼər-a- ‘be separated’ > TB /tʼəra-/ (pres. Cl. III 1sg. tsremar), TA ts(ā)rā- (pret. ptcp. tsro)

Apparent exceptions are due to one of two combinatory changes:

a. *ty > *tʼ

   PIE *pótis, *pōṭy- > PT *pētʼə > TB petso (metrical, for pets* /petʼə?), TA pats

   PIE loc. *-dʰyey (beside dat. *-dʰyōy > Ved. -dhyāi, Av. -dīāi, Umbrian -fet, -fī /-fyē/) > PT *-tʼoy > TB, TA inf. -tśi (usually connected with Lithuanian -t̑i, OCS -t̑i and PIE action nouns in *-tī-, but no case form in *-ty- in classical proterokinetic paradigm!)

b. Grassmann’s Law: *d < *dʰ before an aspirated stop (Winter 1962:24-5)

   PIE *dʰegʷh- ‘burn’ > *degʷh- > PT *tʼək- > TB pres. Cl. VIII tsakṣṭā, TA pres. Cl. X tsāknāṣṭār ‘burns (intr.)’
Probably voiced aspirates were then devoiced (*dʰ > *tʰ), and *d [d] was affricated to *d’. Cf. PT *tə™a- ‘put, lay’ (TB subj. tattam) ←— *dʰe-dʰh₁- to *dʰēh₁-. Then the now voiceless aspirates and the voiced stops fell together with voiceless stops.


PIE *dwi tô- > *wito- (or *dwuto->) *wuto-> PT *wətẽ > TB wate, TA wāt ‘second’
PIE m. *dwoh₁ > *dwũ > *wũ > PT *wəw *[wu] > TA m. wu ‘two’
PIE *pod- —> *pod-yo- > *poyo- > PT *pēyẽ > TB paǐyye, TA pe ‘foot’
Also in ‘wood’; see §3.2.3.2, 4.

2. PIE palatals and velars merged in PT, but labiovelars and sequences of palatal/velar + *w remained distinct. (For a parallel from Italic, cf. Oscan pis, pid < *kʷis, *kʷid vs. acc. sg. fangvam ‘tongue’ < *dʰŋwām ← PIE *dŋ̥wēh₂-m.) For details, see R. Kim 1999.

Labiovelars unrounded next to PIE *o:

PIE acc. *wókʷm > PT *wēkɔ > TB wek /wekɔ/, TA wak ‘voice’;
PIE acc. *h₃ékʷm > PT *ēkɔ > TB ek /ekɔ/, TA ak ‘eye’;
PIE *kʷólós ‘axis, turning’ (Gr. πόλος) > PT *kēlē > TB kele ‘navel’.

New *kʷ then arose from palatal/velar + *w:

PIE acc. *kʷónm̥ > PT *kʷenɔ > TB kwem /kʷenɔ/, TA kom ‘dog’.

3. Palatalization before front vowels created new allophones which then became phonemic, and gave rise to a number of morphologically conditioned alternations. (No clear trace of PT *r̥, which merged with *r in all contexts in both languages.)
4. Auslautgesetze: PIE *-s, *-n, *-t/d [d] > *-Ø; PIE *-nt, *-ns, *-nts > *-Ø
PIE o-stem nom. sg. *-os, acc. *-om > PT *-ë > TB -e, TA -Ø, e.g. in *ékʷ-os, *-om >
PT *yakʷ-ë > TB yakwe /yakʷe/ ‘horse’, TA yuk
PIE *tod ‘that’ > PT *të > TB te ‘this’
PIE sg. masc. nom. *pānts, neu. nom./acc.*pānt (Gr. παντ-) > PT *po > TB po ‘all’

As a result, note that PT forms (and TB underlying forms) can only end in a vowel or a liquid
*r, *l, with two exceptions:

a. pres./subj. 3pl. *-ən < *-ənt < PIE *-ṇi, thematic *-ën < *-ont < PIE *-onti. Variable
loss of word-final *-i (under conditions no longer recoverable) led to variants which
survived into PT and marginally in TA (usually -iñc, -eñc, rarely -i, -e). After
apocope of *-i, the now final *-t was dropped, leaving *-n which was not lost.

b. acc. sg. of animate nouns: PT *-nə > TB /-nə/, reanalyzed as /-n/, e.g. saswe /søswē/
‘lord’, acc. sáswe /søswē-nə/ vs. innovative sáswe /søswē-n/, with same surface
stress as nom.
### 3.2.3.2. Vowels and diphthongs

Main (default) correspondences

<table>
<thead>
<tr>
<th>PIE</th>
<th>PT</th>
<th>TB</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>*a</td>
<td>*a</td>
<td>/a/</td>
<td>ā/a by weakening; §3.3.2, 5</td>
</tr>
<tr>
<td>*e</td>
<td>*e</td>
<td>/ə/</td>
<td>ā/Ø</td>
</tr>
<tr>
<td>*i</td>
<td>*i</td>
<td>/ə/</td>
<td>ā/Ø</td>
</tr>
<tr>
<td>*o</td>
<td>*e</td>
<td>e</td>
<td>a</td>
</tr>
<tr>
<td>*u</td>
<td>*e</td>
<td>/ə/</td>
<td>ā/Ø</td>
</tr>
</tbody>
</table>

*ā*  | *o  | o | a (o next to *p, *m; §3.3.2, 3) |
*ē*  | *e  | e | a (o next to *p, *m; §3.3.2, 3) |
*ī*  | *i  | i | i/y |
*ō*  | *a  | /a/  | ā/a by weakening; §3.3.2, 5 |
*ū*  | *u  | u | u/w |

1. Tocharian “great vowel shift”: distinctive vowel length lost, pre-PT *e, *i, *u merge as *ə
   - *e, *i palatalized preceding consonant, *u rounds adjacent velar
   - initial offglide: *e-, *i- > PT *yə-, *u- > PT *wə-
   - Note that *wi > *wu > *wə (not *wə), e.g. in ‘second’ (§3.2.3.1, 1c; R. Kim 2000a)

2. PIE word-final *-ā > PT *-a, e.g. in fem. nom. sg. of adjectives and nouns (unless *-a spread from PIE *seh₂ > PT *sa to adjs., then to all nouns?)

3. Diphthongs behave like combinations of their elements, e.g. PIE *oy, *ey > PT *ēy, *ōy.

4. Umlaut:
   - pre-PT u-umlaut *o > *o before *u, *ū. Only three secure examples:
     PIE *oktōw ‘eight’ > *oktū (> *oktəw?) —> PT *oktə (influenced by ‘seven’, ‘nine’, ‘ten’) —> TB okt, TA okät
     PIE *dör-u ~ *dér-u- ‘wood’ —> *dör-u ~ *dr-ėw- (Ved. dāru, drós) > *dör-u ~ *rew- (§3.2.3.1, 1c) —> PT *orə > TB, TA or
PIE *suy-u-s ‘son’ (Gr. νιύς), gen. *suy-ew-s ~ *-ow-s > *soy-us ~ *soy-ows > *soy-us ~ *søy-ows > PT *soyə, gen. *søyəw > TB soy, gen. sey-i (-i from pātri, mātri, etc.), TA se, sey-o (for *sayo)

b. a-umlaut: unstressed *ē > *a / __ C₁a (extended to stressed *é in pre-TB; §3.3.1, 5)
PIE *n-ένεh₁-tih₂ > *éknat’a > PT *aknat’a > TB aknatsa /aknát’a/, TA āknats ‘ignorant’
PIE *mógh₂- (~ *mégh₂-) ‘multitude’ > pre-PT *mēka- > PT *maka- > TB māka, TA māk ‘much, many’ (Widmer 2004:159-60)
pre-PT pret. ptcp. *pē-paśṣ-əәw > PT *pa-paśṣ-əәw > TB papāṣṣu, —> TA pāpṣu; similarly for other pret. ptcps. to roots with internal *a
pre-PT priv. *ē-kaka-ttē > PT *a-kaka-ttē > TB akākatte ‘not to be called’; similarly for other privatives to roots with internal *a

c. o-umlaut: (unstressed?) *ē > *o / __ C₁o

5. Sources of the ubiquitous PT *
   a. *R > *əR (except word-initial *R > *ēR-)
   b. epenthesis in triple consonant clusters *kst, *pst
      PIE *swēkstós —> *sekstos > *sekastos > PT *śekastē > TB śkaste, TA śkāst ‘sixth’
      PIE *pståsēnos > *pståstēnos > PT *pāscänē, dual *pāścänēnē —> *pāścänē (haplology) >
         TB pāścane, TA pāśsāṁ ‘breast’
   c. palatalization then becomes phonemic: *e, *i, *u merge as *ə

6. Contraction across *w (Þórhallsdóttir 1988, Winter 1988; Ringe 1996:155-6), especially in preterite participles and *-went- adjectives:
   *ēwē > PT *ē    *ewə > PT *ə
   *owē > PT *o    *ēwə, *əwē > PT *o
   *awē > PT *a    *awə > PT *o
pre-PT pl. *të-tëm-wëša ‘born’ > PT *tëtëmoš > TB tetemos [TA tattmuš after sg. tatemu]
pre-PT pl. *törko-wëša ‘(having) let go’ > PT *törkoš > TB tärkoš, TA tärkoš
pre-PT acc. *përnëwënta ‘worthy’ > PT *përnënt > TB pernen [TA parnont after sg. parno]
PIE *g"ihw-o-mh₃no- ‘living’ > *šawēmanē > PT *šamanē > TB šamāne, TA šāma;
pre-PT *pə-wēnño > PT *poña > TB, TA poñ ‘say’!

7. Ablaut: zero-grades to roots of the shape CeyC and CewC were remodeled in productive ablaut alternations, i.e. in forms which stood in living paradigmatic or derivational relationship with other ablaut grades (Adams 1978:446-8; see Ringe 1996:135-7). On the analogy of CeC and CeRC roots, the regular outcomes of zero-grade *CiC, *CuC, i.e. *C’əC, *CəC, were replaced by *CəyC, *CəwC:

\[
\begin{align*}
*\text{CeC} & > \quad \text{PT} \ *\text{C}’\text{əC} \\
*\text{CoC} & > \quad \text{PT} \ *\text{CēC} \\
*\text{Ce}_\text{e} \text{C} & > \quad \text{PT} \ *\text{CəC} \\
*\text{CeyC} & > \quad \text{PT} \ *\text{C}’\text{əyC} \\
*\text{CoyC} & > \quad \text{PT} \ *\text{CēyC} \\
*\text{CiC} & > \quad *\text{C}’\text{əC} \rightarrow \quad \text{PT} \ *\text{CəyC} \\
\end{align*}
\]

\[
\begin{align*}
& \quad *\text{CeRC} > \quad \text{PT} \ *\text{C}’\text{əRC} \\
& \quad *\text{CoRC} > \quad \text{PT} \ *\text{CēRC} \\
& \quad *\text{CR}_\text{e} \text{C} > \quad \text{PT} \ *\text{CəRC} \\
& \quad *\text{CewC} > \quad \text{PT} \ *\text{CəwC} \\
& \quad *\text{CowC} > \quad \text{PT} \ *\text{CēwC} \\
& \quad *\text{CuC} > \quad *\text{CəC} \rightarrow \quad \text{PT} \ *\text{CəwC} \\
\end{align*}
\]

Examples:
PIE *luk- (zero-grade to *lewk- ‘become light, shine’) > pre-PT *lək’- → PT *ləwk- > Cl. VIII pres. TB lukšām, TA pl. lukšēn ‘illuminate, make shine’ (cf. full-grade *lewk- > PT *l’əwk- > TB lyke ‘light, brightness’), but

PIE *lip- (zero-grade to *leyp- ‘leave behind’) > pre-PT *ləp’- → PT *ləyp- > Cl. III pres. TB lipetrā ‘remains, is left over’ (cf. full-grade *leyp- > PT *l’əyp- > TB, TA lypār ‘rest, remainder’), but
PIE *lip-* > PT *l̥̄p-* in fossilized Cl. III pret. ptcp. *l̥̄p- * l̥̄p-ora > TA lyalypu ‘karma; action, deed’ (vs. absolutive lyalypur-āṣ ‘having left behind’ with analogical -i-).

### 3.2.3. Laryngeal developments

1. All three laryngeals become *a between nonsyllabic segments (as in most other IE branches, except Anatolian and Indo-Iranian), whence PT *a.

   PIE *ph₁tēr ‘father’ > PT *pacer > TB pācer, TA pācar
   PIE ptcp. *-mh₁nos (Av. -mna-, Gr. -μενος, Lith. -mas) > PT *-manē > TB */-mane/, TA -mām, e.g. *h₂e-g-o-mh₁nos > PT *ak-ē-manē > TB akemane, TA ākmām ‘leading’

   See also §5.2.2 on PIE seṭ roots as the source of Tocharian roots ending in *-a- (“a-roots”).

2. Word-initial laryngeals are lost before a nonsyllabic segment, as in all IE branches other than Anatolian, Greek, and Armenian.

   PIE *h₁nēh₁mn ‘name’ > PT *nem > TB ņem, TA ņom (§3.3.2, 3)
   PIE *h₁rud¹rō- ‘red’ > PT *rɔtrë > TB ratre, TA rtär
   PIE *h₂ster- ‘star’ > PT *scɔr-, pl. *scɔr(ə)ŋɔ > TB ściriŋ, TA śreiŋ
   PIE *h₃b₁ruH- ‘eyebrow’ ——> PT du. *pərwa-nē > TB pārwāne, TA pārwām

3. As usual, tautosyllabic laryngeals after vowels are lost with compensatory lengthening, i.e. *VHC > *V:C. The resulting long vowels develop like inherited *ē, *ā, *ō.

4. Laryngeals are lost after syllabic sonorants, as in Germanic; impossible to tell if *RHC first passed through a stage *R₁:C (as in Balto-Slavic) en route to merging with *R₃C.

Special developments

5. It is possible that the laryngeals develop differently when word-initial before a syllabic sonorant (cf. Rix’s Law for Greek and Latin), e.g. *h₂RC- > *arC- in PIE *h₂ŋ-t-bʱō- > PT

6. PIE *ih₂ > PT *ya at least in final position: in addition to feminine adjectives (see §4.2.1), cf. the famous PIE *bʰér-ont-ih₂ > PT *pərēnt’a > TB prentsa ‘pregnant’ (so the usual translation, but cf. now Pinault 2012:184-6). This of course recalls the Greek development to *-ya in e.g. *φέροντ-ja, *φιδή-ja > φέρουσα ‘carrying’, ἰδού ‘knowing’. It has been argued that *ih₁ > *ye word-finally (cf. PIE du. *h₁ék”-ih₁ > *ok”ye —> PT *ēśo-nē > TB ešane, TA ašām ‘eyes’), but all agree that *ih₁ > *ī in the optative suffix. Interestingly, PIE *uh₂ appears to have become *ū: cf. TB akrūna, TA ākrun ‘tears’ — PT *akrōw < PIE *(d)ákruh₂ (PT *akrōwa should have remained as such in TB ākruwa). See Hackstein 1995:16-37, Ringe 1996:7-36.

7. Perhaps *Th₂V > *TʰV, i.e. *h₂ aspirated a preceding stop (Schmidt 1989:307-8)? Only possible to tell when C = *d: thus PIE root aor. *(s)kēdh₂-t > PT pret. *kāta (with *dh₂ > *dʰ > *tʰ > PT *t), extended to PIE pres. *(s)kēd-n-h₂-tōr (Hom. (σ)κιόναται) —> PT *kōt-na-tōr > TB kātnātär*, TA pl. knāntrā. (But also possible that affrication of *d > *d’t (> PT *t’) occurred only in syllable onset before vowel, so that pres. PT *kāt-na- would be regular, whence pret. *kāta-.)

8. Schmidt (1988, 1989:308-11, 1992:103-5, 1995:275ff.) has also claimed that *h₂ > *k after a syllabic sonorant, but his main examples are doubtful:

PT *kōt-k- ‘pass over, surpass, commit (a sin)’ < *gʷdʰ-ₙ₂-h₂- to *dʰegʷh₂- (Gr. φθάνω < *φθάνϝ-, Ved. daghnoti*);

TA pres. trānk- ‘say’, tārnā- ‘let go’ < *trₙ₂-h₂-V-, *trₙ₂-h₂(0)-C- to PIE *ter(K)h₂- ‘let (in, out)’ (aor. PT *tarka- — act. *cēra < *tērh₂-t × mid. *tarka-tē < *tṛh₂-tó?)

Summary of laryngeal developments:

<table>
<thead>
<tr>
<th>PIE *CHC &gt; pre-PT *CaC</th>
<th>PIE *#HC- &gt; pre-PT *#C-</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIE *VHC &gt; pre-PT *V:C</td>
<td>PIE *#H,R₁ &gt; pre-PT *#V,R-?</td>
</tr>
<tr>
<td>PIE *RH C &gt; pre-PT *R:C</td>
<td>PIE *ih₂(#) &gt; pre-PT *ya(#)</td>
</tr>
</tbody>
</table>
3.3. Major phonological developments from PT to Tocharian B and A

3.3.1. PT to Tocharian B

Consonants

1. *ẉ and *y merge as y: PT *ẉentē ‘wind’, *ẉəəsa ‘gold’ > TB yente, yasa [TA want, wās]

2. *p̣ and *ṃ merge respectively with *p and *m in most contexts, but following *əə > i

PIE *médʰu > PT *ṃəətə > TB mit/maytə/ ‘mead’
PIE (*mēms- ~) *mēms- > PT *məənsa > TB misa ‘meat’ (nasal lost after i in TB, cf. ‘five’ below)
PIE *pēnkʷe > PT *p̣əəṇsə > TB piš ‘five’, but PT *p̣əəṇsáka > TB p̣ṣāka, pišāka/p̣ṣəəka/ ‘fifty’

3. *š replaced by *ṭ in all productive morphological alternations; survives only in šak ‘10’ < PT *šakə < PIE *dék̑m [TA šāk] and possibly šuke ‘juice’ < PT *šə̃kə < *dewko- < PIE *dewko- or *dʰewgʰo- (to *dewk- ‘lead’ viz. *dʰewgʰ- ‘draw (milk)’)

Vowels and diphthongs

4. *ē and *e merge unconditionally as e; see examples passim.

5. a-umlaut extended to stressed *ē (in contrast to TA). First discovered by Cowgill (1967) on the basis of Class V subjunctives, e.g.

(post-)PIE *ste-stóṃbʰH- ~ *ste-stmbʰH-’ —> PT *stə-stéma- ~ *stə-stáma- > TB /stáma/- ~ /stáma-/ (TA štama- ~ štma-, without a-umlaut; see §5.2.1). Cf. also
PIE *próti-h₃,kʷ-om (Ved. prātikam, Gr. πρότιος ‘face, etc.’) > PT *prēṭak-a > *prāṭaka > TB pratsāka/praṭāka/, but > *pratsæk > TA pratsak ‘chest’ (no a-umlaut).

6. Syncope of unstressed *ə in open syllables; development of stress-conditioned allophones of /a/ and /ə/ by the classical TB period (see §3.3.3).
3.3.2. PT to Tocharian A

Consonants
1. *PT wʰ and *w merge as w: PT *wʰentē ‘wind’ > TA want
2. PT *pʰ and *mʰ merge respectively with *p and *m (unconditioned)

Vowels and diphthongs
3. PT *ë > a; PT *e, *o > a, but o when next to nonpalatalized labial (*p or *m). Hence
   PIE *gómboś ‘row of teeth’ > PT *kēmē (*mbʰ > *m) > TB keme, TA kam ‘tooth’ vs.
   PIE *h₁nēh₁mn ‘name’ > PT *ṅemā > TB ṅem, TA ṅom ‘name’;
   Iran. *pəræt (cf. Oss. færet) —> PT *peretā > TB peret, TA porat ‘ax’;
   PIE *bhāgu- ‘arm’ (Ved. bāhū-, Gr. πῆχυς) > PT *pok-o > TB acc. pokai, TA poke

4. Word-final vowels *-a, *-ë, *-e, *-o lost: see examples passim.

5. Vowel weakening. Normally PT *a > pre-TA *ā, but
   a. *ā in second syllable > *a after “full vowel” *ā, *a, *e, or *o in the first syllable; and
   b. *a (including *a from by first rule) > *ā when first and third syllables contain “full vowel”
      (after apocope, no. 4 above). These changes are responsible for numerous alternations such as

   ešant ‘giving’          pl. acc. ešāntās < *ešant-ās
   pekant ‘writing’        acc. pekāntām < *pekant-ān
   āknats ‘ignorant’ < *āknats̩  pl. āknsān < *āknats̩-ān
   šāmām ‘monk’ < *šāmān     pl. šāmānān < *šāmān-ān

   After *ā was eliminated as a phoneme, the alternations became morphologically conditioned,
   as in the verbal examples below.

   Cl. VI pres.      kārans < *kārpnās to kārpa- ‘go down’ vs. kārnsnā to kārsnā- ‘know’
   kotnas < *kotnās to kota- ‘split’                 tārnās to tārka- ‘let go’
Cl. V subj.  ārtatār < *ārtātār to ārta- ‘praise’

kalkāṣ < *kalkāṣ to kālkā- ‘go’  vs.  2pl. kālkāc, abstr. kālkālūne

Cl. I pret.  2sg. tākaṣt < *tākaṣṭ to tāka- ‘be’  vs.  2sg. kālkāṣt to kālkā- ‘go’

pekat < *pekāt to peka- ‘write’  kālpāt to kālpā- ‘attain’

   PIE *ōmsos > PT *ansē > *āys > TA es ‘shoulder’ [TA āntse]
   PT gen. *yōkāënsē > *yākāëns > TA yukes ‘of the horse’ [TB yākwentse]
   PIE 3pl. *-nti > PT *-oñcọ ~ *-on > *-āñcā ~ *-ān > *-āñcā ~ *-āy —> TA -iñc ~ -i;
   similarly thematic PIE *-o-nti > PT *-oñcọ ~ *-en > *-añcā ~ *-an > *-añcā ~ *-ay —>
   > TA -eñc ~ -e (see §3.2.3.1, 4)
   PIE *wiHḵmtih₁ > PT *wîkën (*wîyêkën) > *wāykān > *wāykāy > wiki ‘twenty’ [TB ikām]
   PT *wostōññë > *wastāññ > *wastāyñ > TA wāstem ‘domestic, belonging to the house’
   [TB ostaññe]
   PT fem. *kłômāññëa > *kłomāññ > *kłomāyñ > TA kłomiñ ‘noble’ [TA kłomāa]

7. Monophthongization of PT *Vy, *Vw > *e, *o

8. Fate of PT *o > pre-PT *ā, and elimination of phonemic *ä
   a. epenthesis of *ä in consonant clusters created by apocope
   b. syncope of *ä in open syllables

3.3.3. Accent and stress

The TB stress system has been investigated by Marggraf 1970; see also Ringe 1987.
Main characteristics:
• The usual realization rules for /a/ and /ə/ apply in classical and late TB (§3.3.1, 6).
• Stress on underlying word-final syllables is retracted one syllable leftwards. This is an
   exceptionless, fully productive rule in all periods of TB, and is responsible for alternations
   such as
• Stressed *ə is often syncopated in metrical texts, almost always in open syllables (Thomas 1978, Winter 1990): cf. antpi, olypo, pärkre, wärpnätär beside antápi ‘both’, olyápo ‘now’, pärkare ‘long’, wärpánatär ‘enjoys’. Phonetically this probably involved a shift of (at least some) prosodic elements of stress to an adjacent syllable, with automatic deletion of /ə/. A vowel /a/ to the left was usually written ä, but /ə/ was usually written ā, not a; see Winter 1990, Pronk 2009. Most examples of syncopated /ə/ in closed syllables involve reduction of a geminate, e.g. pres. yāmsäm ‘does, makes’, caus. stāmsäm ‘stands (tr.)’, tānmsäm ‘bears, gives birth’ beside yamāssäm, stāmāssäm, tānmāssäm; verbal noun almost always -l(y)ñe beside -lläññe (<— gerundive -lle).

• Alternations such as puwar ‘fire’, pl. pwāra or metrical kewye, kästwer for kewiye ‘cow’s’, kästuwer ‘at night’ demonstrate that prevocalic iy, uw are in fact underlying /y/, /w/.

• Clitic personal pronouns (see §4.3.1) count as part of the domain for stress computation, “allowing” underlying final stress on a preceding verb to surface on that syllable.

On the other hand, the secondary case endings are NOT included in the domain for stress computation, i.e. they had not yet been fully “univerbated” with the acc. of the governed noun (§4.1.3).
• Final /-ə/ may be realized optionally as -o in metrical texts when an extra syllable is required (“bewegliches o”, e.g. pernewo for pernew /pernéwa/ ‘shining’, gen. pl. pontso for ponts /po-ntə/ ‘of all’, pres. 2sg. aksasto /aks-əs-tə/ ‘you announce’.

Although there are no clear indications of stress in TA, weakening of *ā > *a, *ā in the second syllable (§3.3.2, 5) suggests that principal stress fell on the first “full vowel”, with secondary stress two syllables after. Thus in words containing ā, a, e, or o in the first syllable, the stress pattern was *V´ V V´.

The reconstructed PT stress, like the underlying stress in TB, usually falls on the second syllable, but not always.

• Cf. TB pratsäka ‘chest’ vs. TA pratsak (no a-umlaut in first syllable!), pointing to PT *prëťako ← PIE *prórīkʰom (§3.3.1, 5), with retention of initial stress at least into PT.
• A rightward shift from the first to the second syllable (“accent throwing”; Ringe 1987:258ff.) may account in part for the overwhelming preponderance of underlying second-syllable stress in TB
• If initial-stressed Class I, V subjunctives go back to reduplicated PT preforms (§5.2.1), all finite verb forms had stress on second syllable. Might this possibly reflect cliticization in main clauses, as attested in Vedic and indirectly (via recessive stress) in ancient Greek? Cf. Slovenia, where default initial stress in underlyingly unstressed Proto-Slavic forms shifted in most dialects and the standard language to the second syllable, e.g. PSl. acc. sg. *gōrǭ ‘mountain’, *nà gōrǭ ‘to the mountain’ > gōrǭ, na gōro (vs. Russ. gōru, nà goru, SC gōru, nà goru). See now R. Kim 2012a.

4. Nominal morphology

4.1. Nouns

4.1.1. Gender

The noun distinguishes two genders, masculine and feminine, plus a class of nouns of “alternating” gender which take masculine agreement in the singular and feminine in the plural. Historically, the latter class goes back to PIE neuters: o-stem nom./acc. sg. *-om merged with
masc. nom. sg. *-os, acc. *-om, while the PIE collective in *-eh₂ was generalized to the pl. of all feminine adjectives.

<table>
<thead>
<tr>
<th>TB</th>
<th>masculine</th>
<th>feminine</th>
<th>neuter</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom. sg.</td>
<td>astare enkwe</td>
<td>astarya klyiye</td>
<td>astare palsko</td>
</tr>
<tr>
<td>acc.</td>
<td>astarem enkwem</td>
<td>astaryai klañ</td>
<td>astarem palsko</td>
</tr>
<tr>
<td>nom. pl.</td>
<td>astari enkwi</td>
<td>astarona klaina</td>
<td>astarona pälskonta</td>
</tr>
<tr>
<td>acc.</td>
<td>astarem enkwem</td>
<td>astarona klaina</td>
<td>astarona pälskonta</td>
</tr>
</tbody>
</table>

**4.1.2. Number**

Nouns and adjectives inflect for singular, dual, and plural, as does the verb in the 3rd person, though the plural is often used for casual pairs.

- Pace Krause 1954 and Krause and Thomas 1960, Winter (1962) has shown that there is no contrast between “dual” (casual pairs) and “paral” (natural pairs, e.g. paired body parts).
- The productive dual ending is PT *-në, cf. TB ešáne, TA ešäm ‘two eyes’; cf. fossilized eš in eš lmau ‘blind’ (lit. ‘(with) the eyes sat (upon)’), yn-eš ‘obviously’. On the history of the various dual endings, see Hilmarsson 1989.

Tocharian also has a “numerative”, which functions as a count plural for collectives: TB ost ‘house, Buddhist monastery’, ostuwa ‘houses’, ostuw-aïwenta ‘several sets of houses’; pl. tantum mîsa ‘meat’, mis-aïwenta ‘several pieces of meat’; eš-aïwenta ‘(many) individual pairs of eyes’ (see above on eš). The ending *aïwë-nta goes back to PIE *oywo- ‘(one and) only’ (cf. Av. aēuua- ‘one’, Gr. οἶος ‘only’) plus the same individualizing pl. *-nta is found with many other neuter nouns, and in Anatolian count plurals (§4.1.4, 1): cf. especially Cuneiform Luvian tāwa ‘eyes’ vs. tāwanta ‘(many) pairs of eyes’.
4.1.3. Case

Nouns inflect for nine cases in each language, but only the three “primary” cases, nominative, accusative (“oblique”), and genitive, are of PIE date.

- Most non-feminine (= masculine and neuter) nouns have identical forms for nom. and acc. sg., derived from the PIE accusative. Cf. reflexes of PIE consonant-stems such as TB ek, TA ak ‘eye’ < PT *ēkw < PIE *h₂ekʷ-m₇ or TB wek, TA wak ‘voice’ < PT *wēkw < PIE *wókʷ-m₇ (see above, §3.2.3.1). Archaisms: TB nom. pācer ‘father’, mācer ‘mother’, procer ‘brother’, tkācer ‘daughter’ vs. acc. pātār, mātār, protār, tkātār (TA nom./obl. pācar, mācar, pracar, ckācar); nom. TB, TA ku vs. acc. TB kwem, TA koṃ ‘dog’; nom. TB walo, TA wāl vs. acc. TB, TA wḷānt ‘king’ (see §4.1.4).
- However, masculine nouns denoting rational beings have secondarily created a distinct acc. sg. in PT *-nō > TB -m, TA -(a)m. This ending, which has also been extended to most acc. sg. adjective forms, goes back to n-stems formed with the PIE individualizing suffix *-on (cf. the development of the Germanic weak adjective inflection). A good parallel comes from Old High German, where acc. -n was extended to all proper names, e.g. Petrusan ‘Peter’ (preserved in Yiddish, cf. (ix ze) Mojše-n ‘I see M.’).
- Neuter nom./acc. plurals all continue PIE collectives in *-h₂, often with metanalysis of a preceding consonantal suffix: hence PT *-a, *-wa, *-na, *-mna, and especially *-nta (see below). All non-neuter accusatives end in -m in TB, but -s in TA; both apparently from PT *-ns < PIE *-ms.
- The genitive continues the PIE genitive and dative functionally and formally: cf. nt-stem TB -e, TA -O < PT *-ē < PIE *-os (e.g. TB wḷānte, TA wḷānt ‘of, to the king’), r-stem TB, TA -i < PT *-iyo < PIE dat. *-ey (e.g. TB pātri, TA pācri ‘of, to the father’ < PIE *ph₃tr-ey). The origin of the feminine genitive in TA -e and its relationship to TB acc. -ai is hotly debated; see most recently Peyrot 2012a. Other genitive endings are of obscure origin: sg. PT *-ntē (*-nsē?) > TB -ntse, TA -s and pl. TB -nts, -ntś, -ts /-(n)tēə/, TA -śši.

The remaining “secondary” case suffixes are agglutinative, added to the accusative of singular, dual, and plural alike, and attached only to the last element of a noun phrase, e.g. TA kuklas yukas onkālmās-yo ‘with chariots, horses, and elephants’, TB kektseñ reki palsko-sa
‘with/by body, word, and thought’. This Gruppenflexion may be compared typologically with that in Turkic languages or Japanese and Korean (see also §6). With genitives we find variation, e.g. TB kreñcepi onolmentse ‘of the good (gen.) being (gen.)’ or krent onolmentse ‘of the good (acc.) being (gen.)’.

- Although their functions in TA and TB largely coincide, not all the secondary suffixes are clearly cognate. TA has a separate instrumental ending -yo (denoting e.g. agent by which), which is written separately and surely identical in origin to yo ‘and (also)’; in TB the instrumental is expressed by the perlative /-sa/. The TB “causal” is found with only a few abstract nouns, e.g. läkle-ñ ‘because of suffering, out of suffering’, and best treated as a fossilized case form.

- Note that the secondary case endings in TB still do not form part of the accentual domain: hence gen. ņakténtse /ŋaktént’e/ ‘of the god’ vs. perl. ņaktesa /ŋakté/ + /-sa/ ‘on, by the god’, all. ņakteš(c) /ŋakté/ + /-š(c)ə/ ‘to the god’. Exception: the ablative usually does belong to the accentual domain, hence ņaktéməŋ /ŋakté-men/ ‘from the god’ (Pinault 2006). (Contrast the personal pronouns in §4.3.1, which do “attract” stress, i.e. allow stress to surface on the final syllable of the base.)

- In TA, the case endings have become entirely univerbated (except for instr. -yo, see above), and the secondary case forms are subject to vowel weakening (§3.3.2, 5): cf. oṅkas ‘men’, perl. oṅks-ā, all. oṅks-ac, comit. oṅks-aśsāl, loc. oṅks-aṃ (whence analogical abl. oṅks-āṣ).


- TA has reanalyzed the thematic vowel as part of the allative, comitative, and locative ending: cf. PT *ňaktē, loc. *ňaktē-nē > pre-TA *ňakt, *ňaktan —> TA ņkät, ņkät-aṃ.
- On the other hand, TB has generalized final *-s from the acc. pl. of animate nouns in the perlative and allative, e.g. PT *ňaktēns, *ňaktēns-a > pre-TB *ňakten, *ňaktensa —> TB ņaktem, ņaktem-sa.
Sample paradigms of TB ńakte, TA ńkät ‘god’:

<table>
<thead>
<tr>
<th>TB</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom.</td>
<td>ńakte</td>
</tr>
<tr>
<td>acc.</td>
<td>ńakte</td>
</tr>
<tr>
<td>gen.</td>
<td>ńäktents</td>
</tr>
<tr>
<td>instr.</td>
<td>—</td>
</tr>
<tr>
<td>perl.</td>
<td>ńäkteś(c)</td>
</tr>
<tr>
<td>all.</td>
<td>ńäkteś(c)</td>
</tr>
<tr>
<td>comit.</td>
<td>ńäktëmpa</td>
</tr>
<tr>
<td>abl.</td>
<td>ńäktëmeṇ</td>
</tr>
<tr>
<td>loc.</td>
<td>ńäktene</td>
</tr>
</tbody>
</table>

4.1.4. Main inflectional classes

1. Neuter (nom. = acc.)

<table>
<thead>
<tr>
<th>PT</th>
<th>TB</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg. *-R</td>
<td>*pʰokʰəl</td>
<td>pikul</td>
</tr>
<tr>
<td>pl. *-R-a</td>
<td>*pʰokʰəl-a</td>
<td>pikwal-a</td>
</tr>
</tbody>
</table>

The most archaic type, with *-a < PIE *-eh₂ (properly *-e-h₂ to thematic nouns). Note that final -ā has been retained in TA, as the plural morpheme.
These go back to old u-stems (PIE wāst-u) whose pl. *-əwa was reinterpreted as *-ə-wa; the new ending *-wa was extened to other nouns ending in a consonant. Note that ‘house’ is the only exact word equation between TA and TB! Most other pls. of this type in TA end in -wā, with *-ā retained as above, e.g. kursār ‘vehicle, mile’, pl. kursār-wā [TB kwārsā, pl. kwārsar-wa, kursar-wa].

These go back to PIE n-stems, though most have been remade in TA. The ending *-na became the default fem. pl. ending in adjectives: TB orotsts-ana ‘great’, astār-ona ‘pure’.

Apparently from PIE action nouns in *-m(e)n- (cf. TA wākām ‘difference’, pl. wākmant vs. TB wāki, pl. wakana), but many details are unclear. Virtually absent from TA, which has pl. -ant(u) for most nouns with TB cognates in -anma.

Melchert (2000) has correctly compared these pls. with the long misunderstood Anatolian “individualizing” -ant- (e.g. hōamehōant- ‘spring (the season), a particular spring’ vs. hōamehōa- ‘springtime’, ērmant- ‘sick’ to ērmān- ‘illness’). Cf. Hitt. 1-ant- [ānt-] ‘one set, amount’, and the suffixation of -ant- to numerals modifying collectives in -a < PIE *-eh₂ to express a count plural, i.e. ‘so many (units of) X’. See also above under the “plurative”, §4.1.2.

The last two are the only fully productive neuter pl. formations in TB. Virtually all Indo-Aryan borrowings other than proper names are assimilated to these types, depending on whether they
consist of two or three syllables: TB šlok /šlok/ ‘verse’, šlokánma; kleša /kleš/ ‘doubt’, klešánma; akšár /akšár/, aksáränta (also aksáránma); akált /akálk/ ‘wish’, akáltän; samudtär /samówtär/ ‘ocean’, samutärnta (see Lane 1969). In TA an extended variant -ant-u became the most common pl. ending of neuter nouns: -ant —> -ant, secondary cases -antw —> -antu, secondary cases -antw-.

2. Masculine

<table>
<thead>
<tr>
<th>Case</th>
<th>Stem</th>
<th>Form</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom.</td>
<td>*-ē</td>
<td>*yəkʷē</td>
<td>yake</td>
</tr>
<tr>
<td>acc.</td>
<td>*-ē(ə)</td>
<td>*yəkʷē</td>
<td>yake</td>
</tr>
<tr>
<td>gen.</td>
<td>*-ēnt'ē</td>
<td>*yəkʷēntsə</td>
<td>yäkwentse</td>
</tr>
<tr>
<td>nom. pl.</td>
<td>*-ē</td>
<td>*yəkʷē</td>
<td>yake</td>
</tr>
<tr>
<td>acc.</td>
<td>*-ēns</td>
<td>*yəkʷēns</td>
<td>yakwen</td>
</tr>
</tbody>
</table>

The good old thematic o-stems in *-os, with nom. pl. PT *-ē < PIE pronominal *-øy (extended to nouns as in many other IE languages). This vowel regularly palatalized a preceding consonant; the palatalization has been undone in most cases by analogy to the sg., but survives in e.g. TB kercci ‘swords’, kokalyi ‘chariots’ to kertte, kokele. TA has replaced many of these pls. with -əñ, acc. -as, but many are retained in the adjective, e.g. TB trici, TA trice (with palatalization) to TB trite, TA trit ‘third’.

2a. Masculine/feminine

<table>
<thead>
<tr>
<th>Case</th>
<th>Stem</th>
<th>Form</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom.</td>
<td>*-ē</td>
<td>*meñe</td>
<td>meñe</td>
</tr>
<tr>
<td>acc.</td>
<td>*-ō</td>
<td>*meño</td>
<td>meñ</td>
</tr>
<tr>
<td>gen.</td>
<td>*-ōnt'ē</td>
<td>*meñōnt'ē</td>
<td>meñantse</td>
</tr>
<tr>
<td>nom. pl.</td>
<td>*-ōy</td>
<td>*meñōy</td>
<td>meñi</td>
</tr>
</tbody>
</table>

Often said to go back to hysterokinetic n- and s-stems with nom. sg. *-ēn, *-ēs (cf. Gr. ὑμίν ‘membrane’, Lat. lien ‘spleen’; Lith. méνuo, gen. méνes-io ‘moon’), but many details remain to be clarified. Note the type in TB -iye, A -e of TB kälymiye, acc. kalymi (pl. kālymiñ, kālymiñų), TA kālyme (pl. kālymeyantu) ‘direction’.
3. Feminine

nom. *-a  *šōna  šōna  śām ‘woman’
acc. *-o  *šono  šono  śām

The old type, confined to exactly three nouns: ‘woman’; TB lānta ‘queen’ (cf. ‘king’ below); and TB sārya ‘dear one’ (rather ‘lady’), clearly a substantivized adjective. Probably PIE final *-eh₂ > PT *-a (see §3.2.3.2, 2), **-eh₂m > PIE *-ām (Stang’s Law) > PT *-o. See R. Kim 2009a, 2009b:78-81.

nom. *-o  *kantwo  kantwo  kāntu ‘tongue’
acc. *-a  *kantwa  kantwa  kāntu

Peters (1990) argues that at least some of these nouns go back to PIE hysterokinetic nouns in *-ēh₂s, e.g. *dn̥gʰwēh₂s —> *gn̥dʰwās > PT *kantwo, or pl. *-eh₂s in *dʰoHneh₂-es (cf. Lith. sg. dūona ‘bread’) > PT *tano, but what of the others? See now Malzahn 2011.

nom. *-o  *prosko  prosko, proskiye  praski
acc. *-ai  *proskai  praskai

The largest feminine type, with acc. *-ai and in TB, pl. nom. -aiñ, acc. -aim /-ái-ña, -ái-na/ for disyllabics and -a, -am /'-a-nə, -'ə-nə/ for longer stems: pyāyo ‘flower’, acc. pyāpyai, pl. pyapyaíñ, pyapyaím vs. oṅkolma ‘she-elephant’, acc. oṅkolmaí, pl. oṅkolmañ, -am*. The division reflects a pre-PT sound change of posttonic *ai > *a / __ N#; see Winter 1989. TA has generalized -ān, acc. -ās in pyāpyāñ, pyāpyās, but other secondary developments have also taken place. The nom. sg. variants in -iye are surely taken over from the kālymiye type above.

Smaller classes and irregular nouns

‘father’  TB  TA  PT  PIE
nom. sg.  pācer  pācar  *pacer  *ph₂: tér
acc.  pātār  pātār  *patār  *ph₂: tér-m₁
gen.  pātri  pācri  *patrāy  dat. *ph₂:tr-éy
4.2. Adjectives

The adjective exhibits numerous complexities and irregularities, which have still not been fully described, but the main PIE types are well represented.

4.2.1. Thematic stems

Distinction in TB between mono- and disyllabic stems in *-ro- in the masc. pl., cf. lareñ vs. astari. The latter inflection is followed by most other thematic adjectives, including the very productive ones in -tte (privatives), -lle (gerundives), -sse (basic), and -nñe (appurtenance), except that the acc. sg. masculine is not marked by an additional -m/-n(ə)/, and adjectives in -sse and -nñe form the nom./acc. pl. in -ssana, -nñana (in late TB, also gerundive -lyana).

<table>
<thead>
<tr>
<th>Case</th>
<th>‘dear’</th>
<th>‘pure’</th>
<th>‘non faciendus’</th>
<th>‘doable’</th>
<th>‘of the house’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nsg. masc.</td>
<td>lare</td>
<td>astare</td>
<td>ayāmātte</td>
<td>yamālle</td>
<td>ostaśše</td>
</tr>
<tr>
<td>A</td>
<td>larēm</td>
<td>astarem</td>
<td>ayāmāccci</td>
<td>yamalyi</td>
<td>ostaśski</td>
</tr>
<tr>
<td>Npl.</td>
<td>larēn</td>
<td>astari</td>
<td>ayāmāccem</td>
<td>yamalyem</td>
<td>ostaśsem</td>
</tr>
<tr>
<td>A</td>
<td>larē(nā)m</td>
<td>astarem</td>
<td>ayāmāccem</td>
<td>yamalyem</td>
<td>ostaśşa</td>
</tr>
<tr>
<td>Nsg. fem.</td>
<td>lariya</td>
<td>astarya</td>
<td>ayāmācca</td>
<td>yamalya</td>
<td>ostaśšai</td>
</tr>
<tr>
<td>Asg.</td>
<td>lariyai</td>
<td>astaryai</td>
<td>ayāmāccai</td>
<td>yamalyai</td>
<td>ostaśšai</td>
</tr>
<tr>
<td>N/Apl.</td>
<td>laróna</td>
<td>astarona</td>
<td>ayāmāttona</td>
<td>yamallona</td>
<td>ostaśšana</td>
</tr>
</tbody>
</table>
Surprisingly, the feminine suffix continues *-ih₂-, NOT *-eh₂-! This fact has recently received much attention, along with the reflexes of abstract and collective *-h₂- formations in Tocharian: see R. Kim 2009a, forthcoming b, Hackstein 2012, Pinault 2012.

In TA, most thematic adjectives have masc. pl. nom. -e, acc. -es < PT *-e, *-ēns and fem. -i < PT *-āya (acc. -āṃ with overt ending, e.g. āṭāryāṃ, yāmlyāṃ). Continuous recharacterization of the endings has produced a wealth of variants especially in the -ṣi adjectives (< PT *-ṣəә yë, TB -ṣṣ e): masc. acc. -ṣi, -ṣim, -ṣināṃ, fem. acc. -ṣi, -ṣim, -ṣināṃ, -ṣsāṃ, -ṣyāṃ.

4.2.2. Consonant stems

The three most important consonant-stem classes are nt-stems, almost all continuing PIE *-went-; preterite participles in PIE *-wos- ~ *-us-; and n-stems. TA has mostly conflated the first two into a single inflectional type, along with the masculine of n-stems; later TB replaces -n- with -nt- in acc. klyomōṃ —→ klyomont. Paradigms below for ‘worthy’, ‘(having) done, made’, and ‘famous’.

<table>
<thead>
<tr>
<th>TB</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nsg. masc.</td>
<td>perne, yāmu klyomo parno yāmu klyom</td>
</tr>
<tr>
<td>A</td>
<td>pernet yāmos klyomōṃ parnont yāmunt klyomānt</td>
</tr>
<tr>
<td>Npl.</td>
<td>perneñc yāmos klyomoñ parnoñc yāμuñc yāmuñc</td>
</tr>
<tr>
<td>A</td>
<td>pernetāṃ yāmosāṃ klyomōṃ parnoñcāṃ yāmuñcāṃ klyomāncāṃ</td>
</tr>
<tr>
<td>Nsg. fem.</td>
<td>pernauntsa yāmusā klyomāṅa parnōnts yāmus klyomīm</td>
</tr>
<tr>
<td>Asg.</td>
<td>pernauntsai yāmusai klyomāṅai parnomtsāṃ yāmusāṃ klyomīnāṃ</td>
</tr>
<tr>
<td>N/Apl.</td>
<td>pernenta yāmuwa klyomānā parnont yāmunt klyomīnān, -ān</td>
</tr>
</tbody>
</table>

Note also TB, TA pont- ‘all’: TB has an invariant sg. po (both genders and cases), and TA nom. sg. puk (acc. m. poṅcāṃ, f. pontsāṃ).

4.2.3. Other features

The productive genitive singular ending is TB -epi, TA -(y)āp, of obscure origin, though the -p- may be related to the adverbal *-bh- that was grammaticalized in (post-)PIE in oblique case endings.
Tocharian has a number of suppletive adjectives.

<table>
<thead>
<tr>
<th></th>
<th>‘good’</th>
<th>‘big’</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB</td>
<td>TB</td>
<td>TA</td>
</tr>
<tr>
<td>Nsg. masc.</td>
<td>kartse</td>
<td>kāsu</td>
</tr>
<tr>
<td>A</td>
<td>krent</td>
<td>krant *krēnto</td>
</tr>
<tr>
<td>Nsg. fem.</td>
<td>kartsa</td>
<td>krāntsāṃ</td>
</tr>
<tr>
<td>A</td>
<td>kartsai</td>
<td>krāntsāṃ</td>
</tr>
<tr>
<td>Npl. masc.</td>
<td>kreñc</td>
<td>kraṃś *krēñco</td>
</tr>
<tr>
<td>A</td>
<td>krentām</td>
<td>krañcās *krēntōns</td>
</tr>
<tr>
<td>N/Apl. fem.</td>
<td>krenta</td>
<td>krant *krēnta</td>
</tr>
</tbody>
</table>

4.3. Pronouns

The prehistory of the personal and demonstrative pronouns contains a number of unsolved problems. I discuss the major types here.

4.3.1. Personal pronouns

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>TB</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom. sg.</td>
<td>űnas, űni</td>
<td>nās, f. ūnik</td>
<td>twe (tuwe)</td>
<td>tu</td>
</tr>
<tr>
<td>acc.</td>
<td>űnas, ūni</td>
<td>nās, f. ūnik</td>
<td>ci</td>
<td>cu</td>
</tr>
<tr>
<td>gen.</td>
<td>ūni</td>
<td>ūni, f. nāni</td>
<td>tañ</td>
<td>tīñ</td>
</tr>
<tr>
<td>nom./acc. du.</td>
<td>wene</td>
<td>yene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nom./acc. pl.</td>
<td>wes, was</td>
<td>yes</td>
<td>yas</td>
<td></td>
</tr>
<tr>
<td>gen.</td>
<td>wes, wesāñ</td>
<td>wasāṃ</td>
<td>yesi, yesāñ</td>
<td>yasāṃ</td>
</tr>
</tbody>
</table>

TB twe, TA tu < PT *təwē < PIE *tuH-om (Ved. tuvām) and TB ci, TA cu < PT *cəwə < PIE *tewe, but the etymology of the others remains largely obscure. Noteworthy is the existence of separate masculine and feminine forms for ‘I’ in TA. For an intriguing suggestion, see Jasanoff 1989.
Clitic personal pronoun

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>PT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg.</td>
<td>-ŋ /-ão/</td>
<td>-ñi</td>
<td>*-ñayə?</td>
</tr>
<tr>
<td>2</td>
<td>-c /-cə/</td>
<td>-ci</td>
<td>*-cəyə?</td>
</tr>
<tr>
<td>3</td>
<td>-ne</td>
<td>-m</td>
<td>*-nə</td>
</tr>
<tr>
<td>1-3pl.</td>
<td>-me</td>
<td>-m</td>
<td>*-mə</td>
</tr>
</tbody>
</table>

4.3.2. Demonstrative pronouns and adjectives

These continue PT *së, *sa, *tē < PIE *so, *seh₂, *tod, with various suffixes; the unsuffixed forms survive only in TB. On the prehistory of the suffixed demonstratives, see Stumpf 1971, Pinault 2009.

<table>
<thead>
<tr>
<th>TB</th>
<th>TA</th>
<th>‘this’</th>
<th>(Skt. ayam)</th>
</tr>
</thead>
<tbody>
<tr>
<td>se, sā, te</td>
<td>sās, sās, tāś</td>
<td>‘this’</td>
<td>(Skt. ayam)</td>
</tr>
<tr>
<td>su, sām, tu</td>
<td>sām, sām, tām</td>
<td>‘that, he/she/it’</td>
<td>(Skt. sa-, ta-)</td>
</tr>
<tr>
<td>sem, sām, tem</td>
<td>sām, sām, tām</td>
<td>‘this here’</td>
<td>(Skt. eṣa-)</td>
</tr>
<tr>
<td>sam(p), som(p), tam(p)</td>
<td></td>
<td>‘that there’</td>
<td>(Skt. asau)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>masc.</td>
<td>sē</td>
<td>se</td>
<td>sē</td>
<td>sa</td>
<td>sē</td>
<td>sa</td>
</tr>
<tr>
<td>fem.</td>
<td>sā</td>
<td>sā</td>
<td>tā</td>
<td>sā</td>
<td>tā</td>
<td>sā</td>
</tr>
<tr>
<td>neut.</td>
<td>te</td>
<td>te</td>
<td>te</td>
<td>te</td>
<td>te</td>
<td>te</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nsg.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Npl.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
These pronominal stems form the basis for several adverbs, including:

**TB** tesa, tusa (perl. of te viz. tu), TA tāmyo (instr. of tām) ‘thereby, therefore, thus’

**TA** tśam, tšā (loc. viz. all. of tāṣ) ‘here’

**TB** tu, TA tām ‘then, in that case’

**TB** tumem, TA tmāṣ (abl. of TB tu, TA tām; TA also tmā) ‘thereupon, then’

**TB** tune, TA tmaṃ (loc. of TB tu, TA tām) ‘therein, there’

---

<table>
<thead>
<tr>
<th>Case</th>
<th>Nsg.</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>su</td>
<td>ceₙ, cau</td>
<td>cwi ~ cpi</td>
</tr>
<tr>
<td></td>
<td>sāₙ</td>
<td>tāₙ</td>
<td>tāy tuntse</td>
</tr>
<tr>
<td></td>
<td>tu</td>
<td>tu</td>
<td>cam tām tām</td>
</tr>
<tr>
<td>Ndu.</td>
<td></td>
<td>tai</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>tim</td>
<td></td>
</tr>
<tr>
<td>Npl.</td>
<td></td>
<td></td>
<td>tainaisi</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>cey ~ cai</td>
<td>cey ~ cai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>toṃ</td>
<td>to(y)na</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cem</td>
<td>cem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>toṃ</td>
<td>to(y)na</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cesām tosām</td>
<td>cesām tosām</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cesmi tosmāssī</td>
<td>cesni</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>~ cainamts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>samp</td>
<td>ceymp ~ caimp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>comp</td>
<td>toym</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cwimp</td>
<td>toym</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cemp</td>
<td>cemp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cemp</td>
<td>~ cainamts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TA tāmne, tamne ‘in this/that way, thus’ (on this -ne see §4.3.3)
TB tane ‘here, hither’ (stem /təə-/)
TB taisa, taisu, tais, taisem ‘thus, so; in this/that way’
TB tot, TA tāprem ‘so much, so many, so far’

4.3.3. Interrogative and relative pronouns

Interrogative ‘who, which?’

<table>
<thead>
<tr>
<th>TB</th>
<th>TA</th>
<th>PT</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom.</td>
<td>k̑se</td>
<td>kus</td>
<td>*k̑wośē</td>
</tr>
<tr>
<td>acc.</td>
<td>k̑ce</td>
<td>kuc</td>
<td>*k̑wōcē</td>
</tr>
<tr>
<td>gen.</td>
<td>ket(e)</td>
<td>ke</td>
<td>?</td>
</tr>
</tbody>
</table>

The corresponding relative pronoun is TB k̑se, acc. k̑ce, TA kus ne, acc. kuc ne. Late TB reduces k̑se, k̑ce to se, ce. The relative particle -ne also surfaces in other relative pronouns, e.g.

TB mākte    TA mānt    ‘how?’ (Skt. katham)
TB mākte    TA māntne  ‘as; so that’ (Skt. yathā)
TB mant     TA tāmne, tamne ‘so, thus’ (Skt. tathā)

TB kos      TA kos, kosprem ‘how much? how long?’
TB kos      TA kosne, kospremne ‘as much as, as long as’
TB tot      TA tāprem    ‘so much, so many, so far’

4.4. Numerals

Most of the numerals were recognized as IE already in the earliest years of Tocharian studies, though the details are often open to debate:

TB m. ʂe ‘1’ [TA sas] < PT *ʂe < PIE *sēm (Hilmarsson 1984), acc. TB ʂeme, TA ʂom,
pl. TB ʂemi, TA ʂome ‘some’ < PT *şemē, *şemē < thematized *sēm-o-
TB f. sana, obl. samo, sanai, TA sām, pl. TB somona, TA ʂomām < PT *sōna, *somo-
TA m. wu ‘2’ < PT *wōw *[wu] < PIE masc. *dwōh₁
TA f. we, TB m./f. wi ‘2’ < PT *we < PIE fem./neut. *dwoyh₁
TB m. trey ~ trai, TA m. tre ‘3’ < PT *trēy(ə) ← PIE *tréyes (*tróyes?)
TB f. tarya, TA f. tri ‘3’ < PT *təə(y)ə < *tr'əya < PIE *trih₂ (Ved. trī, Av. thṛī, Hom. trīa, OCS tri)
TB m. stwer, m./f. TA ștwar ‘4’ < PT *ștwēr(y)ə < PIE *kʷetwōres (Ved. catvāraḥ, Gr. Doric τέτορες)
TB f. ștwāra [TA f. ștwar by sound change or generalized masculine?]
TB piś, TA pān ‘5’ < PT *pʰənəə < PIE *pēnkʷe
TB skas, TA sāk ‘6’ < PT *șəkə (TB skas rebuilt to śkaste ‘sixth’)
TB sūkt, TA śpāt ‘7’ < PT *ṣapto < PIE *septī
TB okt, TA okāt ‘8’ < PT *oktə ← *oktəw [*oktu] < PIE *oktōw
TB, TA ūu ‘9’ < PT *ņəwə < PIE *h₁néwn₁
TB șak, TA sāk ‘10’ < PT *șəkə < PIE *dēkm₁

TB ikāṃ, TA wiki ‘20’ < PT *wɨkən < PIE *wɨkṁtih₁ (< **dwi-dkṁt-ih₁)
TB tāryāka, TA taryāk [with a from śtarāk, for *tāryāk] ‘30’ < PT *tər(y)əkə < PIE *trih₂ (d)kōmt (< **-omt-h₂, cf. Gr. τριάκοντα; Schindler 1967)
TB śtvārka, TA śtarāk ‘40’
TB pśāka, piśāka /pʰśāka/, TA pān ‘50’
TB śkaska, TA sāksāk [<*sāksāk] ‘60’
TB śuktānka, TA śāptuk ‘70’
TB oktaṇka, TA oktuk ‘80’
TB ūnuka, TA nmuk ‘90’

TB kante, TA kānt ‘100’ < PT *kənt̪e < PIE *kim tôm
TB yaltse, TA wältse ‘1000’ < PT *w̥olt̪e to PIE *weld̪-? (cf. Lith. tūkstantis, OCS tysostí, PGmc. *þusundi- ‘strong hundred’)
TB tmāne, tumane, TA tmām ‘10,000’ (cf. Iranian and Turkic cognates; Eurasian Wanderwort)
5. Verbal morphology

The Tocharian verb exhibits numerous idiosyncratic developments alongside a wealth of interesting and archaic features, and has played an increasingly prominent role in the reconstruction of the PIE verbal system.

5.1. Inflectional categories

Both languages have the same morphological categories and system of stem derivation. Each verb has five “principal parts”, from which all other forms may be derived:

- imperfective (present) stem:
  present and imperfect (= non-past and past)
  present participle active and middle
  gerundive I (denoting obligation, Latin *-ndus*)
  infinitive (TA only!)
  agent nouns in TB -enta, -uki, other mostly unproductive suffixes, e.g. *kaus-enta* ‘murderer’, *kālpāss-uki* ‘thief’

- perfective (subjunctive) stem:
  subjunctive/future and optative (= non-past and past of the perfective stem)
  gerundive II (possibility, Latin *-bilis*)
  verbal noun or “abstract”, almost always built to gerundive II
  infinitive (TB only!)
  TB privative in -tte (TA relics in -t)

- imperative stem: imperative

- preterite stem: preterite

- preterite participle stem: preterite participle, “absolutive” (deverbal noun, cf. Skt. *-tvā, -ya*)
### imperfective (present) | perfective (subjunctive) | imperative | preterite | pret. ptcp.
--- | --- | --- | --- | ---
finite | present | subjunctive | imperative | preterite
imperfect | imperfect | optative | imperative | preterite
nonfinite | pres. nt-ptcp. | pres. m-ptcp. | pret. ptcp. | gerundive 1
(gerundive) | gerundive 2 | verbal noun | infinitive (ta) | infinitive (tb)
agent nouns | privative | absolutive |

Note that the infinitive is built to the imperfective (present) stem in TA, but to the perfective (subjunctive) stem in TB. The imperfect is formally divergent in TA and TB (except for ‘be’ and ‘go’), but is always formed to the present stem in TB with suffix -i-, and usually in TA with suffix -ā-. (On exceptional TA lyāk ‘was seeing’, pārat ‘was taking’, pl. šārsar ‘knew’, crahkās ‘was saying’, etc., see now Peyrot 2012b, R. Kim forthcoming.)

In practice, most verbs do not distinguish five separate stems. Especially in TA, most verbs have a single stem for the perfective, imperative, preterite, and pret. ptcp., so much so that Sieg, Siegling, and Schulze (1931) classify subjunctive forms under the preterite stem. TB has gone slightly further than TA in creating innovative subjunctive and/or present stems, so more verbs in that language contrast three stems for imperfective, perfective, and preterite.

The preterite participle of most verbs can be predicted from the preterite, including all verbs ending in -a-. In general, verbs not ending in -a- exhibit more irregularities and greater unpredictability in their stem formation; see §5.2.3 below for more discussion. As might be expected, a few verbs form irregular imperatives (§§5.2.1, 5.3).

#### 5.1.1. Aspect and tense

See Thomas 1957 on the usage of the various past-tense forms: imperfect, preterite, and periphrastic perfects. The preterite participle and gerundive II build a range of periphrastic constructions, e.g. resultative perfect, past irrealis, etc.; note especially that mā ‘not’ + gerundive II is the standard way to express negative futurity.
The designation of the present and subjunctive stems as “imperfective” and “perfective” is due to Winter (e.g. 1982:9, 1994b:286), and finds a neat parallel in North Slavic languages, where the present of the perfective has future reference in main clauses (cf. Russ. pišet, Cz. píše vs. napišet, napše). However, Peyrot (2010:155-327) has carried out an exhaustive study of uses of the subjunctive in both TA and TB, with reference to Old Turkic parallels of TA texts, and concludes that the subjunctive does not have any perfective value synchronically, but rather expresses uncertainty and hence (often in main clauses) futurity. If the Tocharian contrast of present and subjunctive goes back to an earlier aspectual distinction, as I believe, this must have broken down by the time of our documents.

5.1.2. Voice

The inherited voice distinction of active and mediopassive is robustly preserved. Most mediopassive verbs are either deponent (medium tantum) or denote middle voice (reflexive, self-interest, etc.), but passive examples are also found, e.g. TB sem pšākka še cakanma ok taum yap piṅkce ikām še tāŋktsi šavāte ‘this 51 cāks [and] 80 taus of wheat was eaten [from] the fifth to the 21st [of the month]’ (B461.3). For a detailed study of the middle in Tocharian, see Schmidt 1974.

5.1.3. “Grundverb” vs. “Kausativ”

A famous feature of the Tocharian verb is the widespread suffixation of PT *-skē ~ *-ssə ~ *-s- (> TB /-ske/- /-ssə-/ ~ /-s/-; −→ TA -sa- ~ -s-) to derive transitives to intransitive roots and causatives to many, but not all, transitive roots. Krause and Thomas set up a fundamental divide between “Grundverb” and “Kausativ”, but not every verb forms both, and some “causatives” have the same meaning as the basic verbs from which they are derived. The label is useful, but we should rather think of them as marked transitives. On the formation of the causative, see §5.2.1.

Contrary to the handbooks, there is no contrast between causative and noncausative Class VIII s-presents; all presents of this class (usually with Class I or II subjunctive and Class III preterite) in fact have transitivizing meaning. Cf. Cl. VIII wiksām ‘avoids’ vs. Cl. IXb wikāssaṃ ‘removes’ (‘makes disappear’) to Cl. III wiketār ‘disappears’; the two have merged in TA wikās ‘removes; avoids’. See the extensive discussion of Hackstein 1995:1-2, 147-65.
### 5.1.4. Person-number endings

#### Present/subjunctive

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>PT</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg. act.</td>
<td>-u /-əw/</td>
<td>-m</td>
<td>*-m</td>
<td>*-mi</td>
</tr>
<tr>
<td>thematic</td>
<td>-eₜ, -au /-əw/</td>
<td>-a-m</td>
<td>*-eₜ-m</td>
<td>(*-o-mi)</td>
</tr>
<tr>
<td>2</td>
<td>-t(o) /-t/</td>
<td>-t</td>
<td>*-tə</td>
<td>*-tu?</td>
</tr>
<tr>
<td>3</td>
<td>-m</td>
<td>-ʂ</td>
<td>*-ʂə ~ *-Ø</td>
<td>*-ti</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>PT</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1pl.</td>
<td>-m(o) /-mə/</td>
<td>-mäs</td>
<td>*-məsə, *-mə</td>
<td>*-mes~ *-me?</td>
</tr>
<tr>
<td>2</td>
<td>-cер (!)</td>
<td>-c</td>
<td>*-cə</td>
<td>*-te</td>
</tr>
<tr>
<td>3</td>
<td>-m</td>
<td>-inc ~ -i</td>
<td>*-əncə ~ *-ən</td>
<td>*-nti</td>
</tr>
</tbody>
</table>

#### Imperfect/optative

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>PT</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg. act.</td>
<td>-m /-mə/</td>
<td>(-m)</td>
<td>*-mə</td>
<td>(*-m)</td>
</tr>
<tr>
<td>3</td>
<td>-Ø</td>
<td>(-ʂ)</td>
<td>*-Ø</td>
<td>*-Ø</td>
</tr>
</tbody>
</table>

All others like present/subjunctive.

#### Imperative

<table>
<thead>
<tr>
<th></th>
<th>TB</th>
<th>TA</th>
<th>PT</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2sg. act.</td>
<td>-Ø</td>
<td>-Ø</td>
<td>*-Ø</td>
<td>*-Ø</td>
</tr>
<tr>
<td>pl.</td>
<td>-so</td>
<td>-su</td>
<td>*-so?</td>
<td></td>
</tr>
</tbody>
</table>
2sg. mid.  -r  -r  *-r
pl.  -t  -c  *?@

Preterite

a-preterites

1sg. act.  /-a-wa/  -ā  *-a-wa?
2  /-a-sta/  -ā-ṣt (-a-ṣt)  *-a-sta-  *-s + *-th2e?
3  /-a/  -Ø, -ā-  *-a  *-t

1pl.  /-a-mə/  -ā-m (-a-m)  *-a-mə  *-me
2  /-a-sə/  -ā-s (-a-s)  *-a-sə  ?
3  /-a-re/  -ā-r (-a-r)  *-a-rē  perf. *-r + *-ont

(pre)sigmatic

1sg.  /-ə-wa/  -u, -wā  *-ə-wa?
2  /-ə-sta/  -āst  *-ə-sta  *-s + *-th2e?
3  /-əsa/  -ās  *-əsa  *-s-t?

1pl.  /-ə-mə/  -ām  *-ə-mə  *-me
2  /-ə-sə/  -ās  *-ə-sə?
3  /-ə-rə/  -ār  *-ə-rē  perf. *-r + *-nṭ

1sg. mid.  /-a-mai/  -e  *(a)-ai  *-h2e + *-i?
2  /-a-tai/  -ā-, -(a)-te  *-a-tai  *-th2e + *-i?
3  /-a-te/  -ā-, -a-t  *-a-tē  *-to

1pl.  /-a-mte/  -ā-, -a-māt  *-a-mtē  *-medh2
2  /-a-sə/  -ā-, -ac  *-a-Xs?
3  /-a-nte/  -ā-, -a-nt  *-a-ntē  *-nto
5.2. Stem formation

5.2.1. Traditional classification

Krause and Thomas (1960) divide Tocharian stem formations into the following types.

Present
• I: athematic to root not ending in PT *-a-, with originally epenthetic Stützvokal *-ə- before ending
• II: thematic, stem vowel PT *-ə- ~ *-ë- < PIE *-e/-o-
• III: intransitives in PT *-e- (< PIE *-êh-?) to roots with internal vowel *ə (*əy, *əw)
• IV: intransitives in PT *-o- to roots with internal vowel *a (*ay, *aw)
• V: athematic to roots ending in *-a-
• VI: nasal-infixed to roots ending in *-a-, PT *-(ə)n-a-
• VII: nasal-infixed to roots not ending in *-a- (TB only); athematic! (Schmidt 1985)
• VIII: thematic s-presents: PT *-ṣə- ~ *-s- < pre-PT *-s/-ə- (N.B. “non-causative” VIIIa vs. “causative” VIIIb invalid, see §5.1.3)
• IXa: suffix PT *-ṣṣə- ~ *-s- ~ *-skə- < PIE *-skə-
• IXb: same, but causative
• Xa: suffix PT *-nə-ṣṣə- ~ *-nə-s- ~ *-nə-skə- (nasal + *-sk- suffix)
• Xb: same, but causative
• [XIa: suffix PT *-sə-ṣṣə- ~ *-sə-s- ~ *-sə-skə-; actually Cl. IX to roots ending in *-s-]
• [XII: suffix PT *-nə- *-nə-s- *-nə-skə-; actually Cl. IX to roots ending in *-s-]
• XII: suffix PT *-nə- ~ *-nə-: denominal < *-n- (originally n-stems) + PIE *-yə/-ə-, or deverbal < PIE *-nə- (< Ved. grbhāyāti beside grbhnhāti)

Subjunctive
• I: athematic to roots not ending in PT *-a- (§5.2.2); mostly initial stress in TB; roots with underlying /ə/ show ablaut between *-ə- in the subj. act. sg. and *-ə- elsewhere
• II: thematic, stem vowel PT *-ə- ~ *-ə- < PIE *-e/-o-
• III: stem vowel PT *-e- (or *-e-?)
• IV: stem vowel PT *-əy- (*[-i]) < PIE *-(e)yə/-ə- (see R. Kim 2007a)
• V: athematic to roots ending in PT *-a- (§5.2.2); many have initial stress in TB; roots with underlying /ə/ show ablaut between *-ê- (> TB ā by a-umlaut, TA a) in the subj. act. sg. and *-ə- elsewhere
• VII: (TA only) suffix PT *-ñə- ~ *-ñë- < ???
• IX: suffix PT *-ʃʃə- ~ *-ʃ- ~ *-skê- to Cl. IX causative presents
• X: suffix PT *-nə-ʃʃə- ~ *-nə-ʃ- ~ *-nə-skê- to Cl. X presents
• XII: suffix PT *-nə-şʃə- ~ *-nə-ʃ- ~ *-nə-skê- to Cl. XII presents

Preterite
• I: PT *-a-, originally final vowel of set roots, reinterpreted as suffix and extended to almost all preterites (except Class VI relics); roots with underlying /ə/ (§5.2.2) show ablaut between act. sg. *-ê-, du./pl. -ë-(!), and mid. *-ə-
• II: PT *-a- to causative verbs with root vowel *ə: originally reduplicated, preserved in TA but remodeled in TB, e.g. TA šašārs vs. TB šārsa ‘made know(n), announced’ (vs. Cl. I TB šarsa, TA šārsa ‘knew’; R. Kim 2003a)
• III: PT 3sg. *-əsa to roots not ending in *-a-, stem vowel *e < PIE *ē: continues ancestor of classical PIE sigmatic aorist (§5.4.3)?
• IV: PT *-ʃʃ-(əy)a- to causative verbs with root vowel *a (and a few with *ə)
• V: PT *-nə-(əy)a- to Class XII pres./subj.

The imperative usually patterns with the subjunctive (e.g. TB ptärka, pl. ptärkaso, TA ptark, pl. ptärkās ‘let go!’ like subj. Cl. V TB tārkam, pl. tārkam, TA tarkaš, pl. tārkeñc) or preterite (TB ptes, mid. ptāsar, TA ptaš, mid. pātsār ‘put!’ like pret. Cl. III TB tessa, TA casāš). Irregular forms include relics of PIE thematic *-e (virtual PIE *wok*-n-ye —> *pə-wēnə > PT *poññ > TB poñ, TA peñ ‘say!’), athematic *-d'i (*h₁idʰi —> *isi > *(pə)-yəsə > TB paš, TA piš ‘go!’; Jasanoff 1987), and the *-si imperative type (*klewsi —> PT *pə-kl'ewə > TB pāklyauš, TA pāklyoš; cf. Ved. šrōsi, ibid.), as well as several TB iptvs. in -e, of obscure origin.
5.2.2. “a-roots” and “non-a-roots”

The most important division within the Tocharian verb system is that between roots ending and not ending in *-a- (> TB /-a/-, TA -ā- ~ -a- ~ -∅-). Since laryngeals between nonsyllabic segments became PT *a, PIE set roots, i.e. roots ending in a laryngeal, developed into roots ending in *-a-. This is clearest from Class I preterites which continue PIE root aorists (Schmidt 1982; see below, §5.4.1):

PIE *kʷrih₂- ‘buy’ (Ved. krīṇāti, mp. krīṇīte, OIr. crenaid, aor. Myc. qi-ri-ja-to, Hom. πρίατο) > *kʷr'ya- > PT *kʷərʰya- > TB /kərya-/; pret. 1pl. kāryāmai, mp. 1sg. kāryāmi, 2 -tai, 1pl. -mtē;

PIE *kedh₂- (Hom. ᱠkə́-a-ra) > PT *ṣata- > TB pret. 2sg. ᱠtasta, 3pl. ᱠtare (w. dial., for ᱠtāṣta*, ᱠtāre*) ‘scattered, sowed’

PIE *stemb⁶H-t (Ved. āstambhbhi) > PT *ścema > pret. TB śama, ścma(-c), TA śăm ‘stopped, came to a stand’

PIE *kēlh₁-t > PT *śala > pret. TB śala, TA śāl ‘led’, mid. *klḥ₁-tō > PT *kala-tē > TB klāte, TA klāt

This root-final *-a- was extended to many roots which were not originally set, as well as many of obscure etymology. We can observe the spread of *-a- to a limited degree in TB: the verb for ‘eat’ is /sōwa-, but the pret. ptcp. šesu /še-sōwə/ and isolated subj. 1sg. sū /sōwə/ (for *sawu? Pinault 1994:170-82) preserve the earlier root *šaw- < *kyew- (OE čēawan ‘chew’, Pers. jāv-īdan). As we will see below, the morphology of roots ending in *-a- was much more regular and predictable, and so was more readily learned by children and tended to be overapplied to other verbs.

Tocharian verb roots can also be divided into those with underlying root(-internal) vowel *ə and those with *a; the latter have various sources, but in those with identifiable PIE sources, the *a appears to reflect a root-internal laryngeal (cf. PT non-pres. *(s)tk- ‘be’ < PIE *sth₂-k- to *steh₂- ‘stand’). There are thus four possible types of roots, which may be labeled as follows (Peyrot 2010:44-7):
5.2.3. Major patterns of verb stem formation

a. Roots ending in *-a-

<table>
<thead>
<tr>
<th>Present</th>
<th>Non-present</th>
</tr>
</thead>
<tbody>
<tr>
<td>intransitives in *-e- (III/IV)</td>
<td>root in *-a- (V)</td>
</tr>
<tr>
<td>transitives in *-(ə)n-a- (VI)</td>
<td>root in *-a- (V)</td>
</tr>
</tbody>
</table>

- Note the high degree of predictability: most transitive verbs take nasal-infixed presents in PT *-(ə)n-a-*, whereas intransitives form mostly deponent presents in PT *-e- or *-o- < PIE stative *-éh₁-.
- A handful of originally Cl. VI presents have been recharacterized with the *-sk- suffix to become Cl. X presents in *-na-/~əә~/. The original pres. may become subj., displacing the earlier Cl. V formation. Example: TB pres. Cl. X /kəә-rn-a-sk-/ ‘buy’, subj. Cl. VI /kəә-r-n-a-/, pret. Cl. I /kəә-rν-yα-/.
- Cl. V subjs. to transitive Cl. VI presents continue old reduplicated formations (cognate with PIE perfect) and so exhibit initial stress and paradigmatic ablaut in TB, while Cl. V subjs. to intransitive Cl. III presents are inner-Tocharian innovations, and have regular second-syllable stress and no ablaut in TB. Similarly, Cl. I preterites to the first type go back to root aorists and show ablaut, while Cl. I preterites of the second type are likewise innovative and do not ablaut. For details, see R. Kim 2003b.

b. Roots not ending in *-a-

<table>
<thead>
<tr>
<th>Present</th>
<th>Non-present</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple thematic *-²/ɛ- (II)</td>
<td>simple thematic *-²/ɛ- (II)</td>
</tr>
<tr>
<td>thematic *-š³/xe- (VIII)</td>
<td>root athematic (I)</td>
</tr>
</tbody>
</table>
| simple thematic *-²/ɛ- (II) | }
thematic *-skê- (IX)  
thematic *-nê/e- (XII)  
“causative” *-skê- (IX)  

Note that TB drops root-final /-a-/ and has initial stress in the caus. pres. and subj.:

anā-sshām /anâ-/ ‘breathes (in)’ vs. caus. 1pl. ānāskem /án-/ ‘id.’;  
kāršanam /kərs-ón-a-/ ‘knows’ vs. caus. pres./subj. šārsāsshām /šórs-/ ‘lets know, informs’;  
wiketār ‘disappears’ (root /wɔyká-/) vs. caus. pres./subj. wīkkāsshām /wɔyk-/ ‘makes disappear, removes’.

TA likewise drops root-final -ā- in the pres., but regularly retains or adds -ā- in the subj., e.g.  
pres. tsālp-ā-stār ‘releases’, subj. tsālp-ā-stār vs. TB pres./subj. tsālp-ās-tār (to TA tsālpā-,  
TB /tɔlpa-/, non-caus. pres. TB tsālpetār, TA ŝalpatār ‘is released’).

5.3. Suppletion

Approximately a dozen verbs are suppletive, second only to Old Irish among IE languages. The most important are:

‘go’

• PT pres. Cl. I *yə- (< *i- <— PIE *h₁éy- ~ *h₁i-’) > TA yāṣ, pl. yiṇc; TB 1sg. yam, 3 yam, but  
  pl. ynem (<— *yən + them. *-en, whence 1pl. ynem, m-ptcp. ynemane)  
• TB subj. = pres.; pret. sg. Cl. III masa, pl. Cl. III maitār (cf. pres. Cl. III mitetār ‘sets out,  
goes’)  
• TA non-pres. kalka- ~ kālkā-
‘do, make’

- TA ya-, ypa- (distributed exactly like *-e- vs. *-o- forms in thematic paradigms; prehistory?);
  TB /yam-²⁰₉/ /myk-/ rebuilt to subj. /yam-/  
- PT non-pres. *yam-: subj. Cl. I (thematized as Cl. II in TA), pret. Cl. III (replaced in TB with secondary *yam-əəṣṣəəya-)

‘stand (up), be standing’

- PT pres. Cl. II *kəəlʸ-²/ə- (medium tantum; < *kl'yə-²/ə- < PIE aor. subj. *kléy-²/ə-)  
- PT non-pres. *stēma- ~ *ścōma- ~ *stōma- (< PIE *stembʰH-)

‘sit (down), be sitting’

- PT pres. Cl. II *şomʸ-ə- ~ *şom-ə-  
- PT non-pres. *lēma- ~ *lōma- ~ *lōma-

‘give’

- PT pres. Cl. IX *ay-²⁰₉/šk-², subj. Cl. I *ay-  
- PT pret. Cl. I *wəsa-  
- iptv. TB pete, pl. petso, petes, TA paṣ, pl. pac

‘take’

- PT pres. Cl. II *pər⁸-(y)-ə- ~ *pər-ə- (< PIE *bʰer-²/ə-)  
- PT subj. TB /as-/, TA kām(a)-  
- PT pret. TB /kama-/, TA kām(a)-

‘lead’

- PT pres. Cl. II *aš-ə- ~ *ak-ə- (< PIE *həg-²/ə-)  
- PT non-pres. TB /waya-/, TA wāwa- (< PT *wawʰa-?)
The forms of the verb ‘be’ are also formed to several stems:

- **present**: PT *nēs-* > TB Cl. I *nes-, TA Cl. II *nas- (3sg. *naṣ, pl. *neñc)
- **imperfect**: PIE *h₁s-yēh₁- (~ *h₁s-ih₁-’) > *ṣe + productive impf./opt. suffix *(o)y- —> PT *ṣey > TB ṣey ~ ṣai, TA ṣe-ṣ (with pres. -ṣ)
- **non-present**: PIE *steh₂- ‘be standing’ > *sta-
  initial *s- preserved only in TA iptv. 2sg. pā-ṣṭāk ‘be’
  PT *ta- preserved in TA subj. 1sg. ṭāṁ, 3 tāś, 3pl. teñc (~ tākeñc)
  otherwise PT *taka- (subj. TB ṭāka; pret. TB 3sg. ṭāka, 3pl. takāre, TA 3sg. tāk, 3pl. tākar)

In addition, TB also has a copula ste, pl. skente < PIE *h₁s-skē-to, *h₂s-skō-nto, the only clear survival of the imperfect in Tocharian (Hackstein 1995:272-82, also on late TB pl. stare).

### 5.4. Evolution of verb stems and stem systems

Most inflectional categories and patterns of verbal stem derivation are of PIE date, including reflexes of nasal and stative presents and root and (pre-)sigmatic aorists.

### 5.4.1. PIE verb formations in Tocharian

The following reconstructed PIE verb categories have secure reflexes in Tocharian.

<table>
<thead>
<tr>
<th>Aspect (Aktionsart)</th>
<th>Present root (athematic)</th>
<th>Present nasal infix</th>
<th>Present simple thematic</th>
<th>Present thematic <em>(sk</em>/o-</th>
<th>Present thematic <em>(y</em>/o-</th>
<th>Present Class I presents (non-a-roots)</th>
<th>Present Class V pres.? (a-roots)</th>
<th>Present Class VII pres. (non-a-roots)</th>
<th>Present Class VI pres. (a-roots)</th>
<th>Class II pres./subj. (only three! see §5.4.3)</th>
<th>Class II pres./subj.</th>
<th>Class IX pres. (subj.)</th>
<th>[Class X pres. (subj.)]</th>
<th>Class XII pres./subj. (*n- + <em>y</em>/o-)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*-nt-, *-mh- no-participles</td>
<td>nt-, m-participles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>imperfect</td>
<td>TB copula ste, skente</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aorist</td>
<td>root (athematic)</td>
<td>thematized in Class VI preterites (non-\textit{a}-roots)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|   |   | Class I preterites (\textit{a}-roots), *
\textit{a}- reinterpreted as suffix |
|   | proto-sigmatic | Class III preterites |
|   | reduplicated | Class II preterites |
|   | (or originally imperfect?) |   |
| Perfect | (proto-)perfect | Class I subjunctive (non-\textit{a}-roots)? |
|   |   | Class V subjunctive (\textit{a}-roots)? |
|   | perfect ptcp. | preterite participle, (formerly) reduplicated |
| Mood | subjunctive | (present?) |
|   |   | Class II presents/subjunctives |
|   | aorist | Class II presents/subjunctives |
| optative |   | imperfects of ‘be’, ‘go’ |
|   |   | all optatives and imperfects (except small TA class |
|   |   | of reanalyzed preterites; R. Kim forthcoming a) |
|   |   | Class I pret. in *
\textit{a}-\textit{ya}- to Class II pres./subj. |
|   |   | Class IV pret. in *
\textit{a}-\textit{s}s\textit{a}ya- to Class IX pres./subj. |
|   |   | Class V pret. in *
\textit{a}-\textit{n}n\textit{\o}ya- to Class XII pres./subj. |
| imperative | root (athematic) in *
\textit{a}-Ø | Class I iptv. (roots in *
\textit{a}-) |
|   | thematic | ‘say!’ (see §5.2.1) |
|   | athematic *-d\textit{i} | ‘go!’ ‘give!’ |
|   | *-si ← s-subjunctive | ‘hear!’ |
5.4.2. PIE origin of Tocharian verb formations

<table>
<thead>
<tr>
<th>Tocharian</th>
<th>PIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>present</td>
<td>present (athematic, thematic; root, nasal-infixed, suffixed)</td>
</tr>
<tr>
<td>root present/aorist subjunctive</td>
<td></td>
</tr>
<tr>
<td>subjunctive</td>
<td>proto-perfect? (R. Kim 2007b)</td>
</tr>
<tr>
<td>root aorist subjunctive</td>
<td></td>
</tr>
<tr>
<td>imperfect/optative</td>
<td>optative (*-yêh₁- ~ ) *-ih₁-</td>
</tr>
<tr>
<td>imperatve</td>
<td>imperative (various types)</td>
</tr>
<tr>
<td>preterite</td>
<td>root aorist</td>
</tr>
<tr>
<td>proto-sigmatic aorist?</td>
<td></td>
</tr>
<tr>
<td>optative *-ih₁-</td>
<td></td>
</tr>
<tr>
<td>preterite participle</td>
<td>&lt;  perfect participle</td>
</tr>
</tbody>
</table>

5.4.3. Outstanding problems

- How did the morphological split between a-roots and non-a-roots become established?

6. Word formation and syntax

As the Tocharian languages are left-branching, the verb is usually clause-final in prose documents, but may be raised for various pragmatic effects. (Metrical texts not surprisingly
offer much variation.) The left-branching nature of Tocharian is also apparent in complex
derivatives: it is far from rare to find nominals with four or even more suffixes, as in the
following examples:

TB  \textit{raddhi- lak- äṣ(ś- āly)- ſe- şše}

\begin{tabular}{l}
\textit{wonder} \textit{see} \textit{CAUSATIVE GERUNDIVE VB.NOUN ADJ} \\
\end{tabular}

‘of causing to see wonders’

TA  \textit{ākār- aś- n- um- in- ſn- ac}

\begin{tabular}{l}
\textit{tear} \textit{eye DUAL ADJ FEM ACC.SG ALLATIVE} \\
\end{tabular}

‘to her of the teary eyes’, i.e. ‘to her whose eyes are both full of tears’ (Krause 1955:30-1)

Here once again we observe a typological similarity to Turkic languages or Japanese and Korean.
Along with the grammaticalization of postpositions as case markers (see §4.1.3), one also
finds the opposite phenomenon in which suffixes become phrasal clitics:

TB  \textit{[tāu ri]- şsī śāmna šemi}

\begin{tabular}{l}
\textit{this.FEM. city.SG.OBL pertaining to human.PL.NOM one.MASC.} \\
\textit{SG.OBL PL.NOM} \\
\end{tabular}

‘some people of this city’ (PK 17.10 b5; Hackstein 2004:93 with refs.)

Cf. English possessive ’s ←— Old English a-stem gen. -es in \textit{[the guy in our class who I ran into yesterday]’s dog just died}.

Nominal compounds are fairly common, although the overwhelming majority are
Sanskrit or Prakrit borrowings. Due to massive cultural influence of Buddhism over the
centuries, very few traces of PIE poetic language or naming practices survive in our texts: in
addition to the famous TB \textit{ñem-kālywe}, TA \textit{ñom-klyu} ‘(name and) fame’, cf. (male) names of
laymen such as \textit{Kwem-toko} ‘running like a dog’ (Pinault 1987).
7. The position of Tocharian within Indo-European

Early on, many Indo-Europeanists were struck by the apparent connections between Tocharian and the western IE languages, particularly (Italo-)Celtic and Germanic. However, most of the relevant features turned out to be archaisms (and so of no value for subgrouping) or trivial innovations, most famously

- the “centum” merger of PIE palatals and velars: easily repeatable (like most phonetic mergers), also occurred between Proto-Anatolian and Hittite; and
- the primary r-endings of the mediopassive: archaism, also preserved in Anatolian (and Phrygian) in addition to Italic and Celtic; replaced in Indo-Aryan, Greek, and Germanic by *-i from the active.

Adams has argued for a close connection between Tocharian and Germanic; likewise Georgiev for Tocharian and Balto-Slavic. Ringe (1990) argues persuasively that none of these hypotheses withstands closer scrutiny, and that Tocharian is not closely related to any other branch of IE. More recently, Hackstein (1998, 2005) adduces further innovations allegedly shared by Tocharian and “western IE” (including Greek and Balto-Slavic), but most of these are either possible archaisms or isolated syntagms, instances of grammaticalization, etc. involving Tocharian and a single other branch, e.g. Greek.

Today, the emerging (now general?) consensus holds that Tocharian is not closely related to any other branch, but rather was the second after Anatolian to diverge from the ancestral speech community. The remaining “Brugmannian” or Inner IE languages then underwent certain innovations in common, which may be projected back to “Proto-Inner-IE”.

Morphological evidence

- Cl. III preterite with -s- formant in 3sg. only, vs. classical PIE sigmatic aorist; cf. Hitt. hi- conj. pret. 3sg. -š (Jasanoff 1988, 2003, with older references going back to Ivanov, Watkins, etc.; otherwise Ringe 1990)
- Paucity of simple thematic presents (Ringe 2000; see §5.4.3)
• Thematic optative *-ih₁- < *-o-ih₁- with deletion of thematic vowel as in i-stem abstracts and adjectives in *-yo- to o-stem bases; remodeled in Italic and Celtic as *-ā- (*-eh₂-?), elsewhere as *-oih₁- (see now Jasanoff 2009)
• Generalization of feminine adjectives in *-ih₂- to thematic bases, vs. Inner IE *-eh₂-? See R. Kim 2009b, forthcoming b, Fellner 2012.

Lexical evidence (see Winter 1968, 1997, Schmidt 1987, 1992)
• PIE *h₁ebebʰ- ‘enter’ > PT *yop- ~ *yəp- > TB /yop-/ ~ /yəp-/ vs. Inner IE ‘f*ck’ (Ved. ýabhāṭi, Gr. οἴφω, Russ. jebū): one-way semantic development!
• PIE *wiHró- ‘young’ > TA wir vs. Inner IE ‘hero, man’ (Ved. vīrā-, Lat. vir, OIr. fer, Lith. výras): ditto
• PIE *nókʷt- ~ *nékʷt- ‘evening’ > Hitt. nekuz (meḫur) ‘evening time’, TB nekciye, TA nakcu ‘(yesterday) evening’) vs. Inner IE ‘night’
• PIE *h₂erh₃-o- ‘plow’ > PT *arē > TB āre vs. Inner IE *h₂erh₃-tro- > Gr. ἀρότρον, W. aradr (Lat. arātrum)
• PIE *(h₁)ēh₂gʷh- ~ *(h₁)ēh₂gʷh- ‘drink’ > Hitt. ēkw- ~ akw-, PT *yokʷ- > TB, TA yok- (cf. Lat. ēbrius); lost as a verb in Inner IE in favor of *peh₃- (R. Kim 2000b)

Computational models have also arrived at an optimal cladistic tree for IE with this same first-order subgrouping; see Ringe et al. 1998.
Selected Bibliography

1. Major text editions (by date of publication)


*A Comprehensive Edition of Tocharian Manuscripts (CETO)*. <http://www.univie.ac.at/tocharian/>

2. General references (by date of publication)

2.1. Grammars and general descriptions


2.2. Dictionaries


Hilmarsson, Jörundur. 1996. Materials for a Tocharian Historical and Etymological Dictionary. Ed. by Alexander Lubotsky and Guðrún Þórhallsdóttir with the assistance of Sigurður H. Pálsson. (Tocharian and Indo-European Studies Supplementary Series, Vol. 5.) Reykjavik: Málvisindastofnun Háskóla Islands. [Far from complete, but very important for the words it does contain, mostly beginning with ā and k.]


3. Specialized monographs


Malzahn, Melanie, ed. 2007a. Instrumenta Tocharica. Heidelberg: Winter. [Not a monograph, but unquestionably certainly belongs here as a vade mecum of Tocharian philology.]
—. 1985. *Die Erforschung des Tocharischen (1960-1984)*. (Schriften der Wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität, Frankfurt am Main. Geisteswissenschaftliche Reihe Nr. 5.) Stuttgart: Steiner. [Also not a monograph; more of a critically annotated bibliography.]

4. *Other works cited*

—. 1988. The sigmatic aorist in Tocharian and Indo-European. TIES 2, 52-76.


—. 2000. Tocharian class II presents and subjunctives and the reconstruction of the Proto-Indo-European verb. TIES 9, 121-42.


—. 1967. Tocharische Miszellen. IF 72, 239-49.


