

Tocharian

Proto-Indo-European Language and Society

Rolf Noyer

Tocharian Language

- Remains of an Indo-European language generally referred to as Tocharian were found in the **Tarim Basin**, a large arid lowland located in present-day Xinjiang Uyghur Autonomous Republic (north of India, northeast of the Tibetan plateau, formerly called ‘Chinese Turkestan’).
- Tocharian is actually two languages: **Tocharian A** (north-east area) and **Tocharian B** (mostly south-west area)
- Tocharian texts are dated to between 500 and 840 CE. They are preserved in manuscripts (90% are Buddhist texts) and wall frescoes.
- Around 840 there was a migration of Uyghurs into the Tarim Basin and Tocharian ceased to be a spoken language.

Proto-Tocharian

- Tocharian is the most divergent branch of IE after Anatolian, suggesting that the population that spoke proto-Tocharian separated early
- It is probable that speakers of proto-Tocharian were associated with the **Afanasevo Culture**, located considerably to the east of the Pontic-Caspian, and north of the Tarim Basin.
- The Afanasevo Culture is dated to 3500-2500 BCE.
- Tocharian is a **centum** language

3

Tocharian language remains

- The Tocharian manuscripts were discovered around 1890.
- They were first deciphered in 1908 by Emil Sieg and Wilhelm Siegling
- They are written in a unique alphabet, clearly derived from the **Brahmi script**, the ancestor to modern **Devanagari** (Sanskrit and Hindi alphabet)

4

Centum-Satem Isogloss

- The centum/satem **isogloss** was originally thought to separate western from eastern PIE languages.

Satem Shift

*k̂, *ĝ, *ĝ ^h	>	*č, *j, *j ^h	>	Skt	ś, j, h
				Avestan	s, z, z
				Russian	s, z, z

*k̂leu-os ‘magical fame’ > Skt śravas, Av srauuah-, Gk kléos

*ĝomb^h-os ‘tooth’ > Skt jámbhas-, R zub, Gk gómp^hos, NE comb

*ĝ^heu- ‘pour’ > Skt homas- ‘sacrifice’, Av zaotar- ‘libator’, Gk k^héō ‘pour’

5

Tocharian Centum Examples

*k̂mtóm ‘HUNDRED’ > TA kánt, TB kante, Lith šim̃tas, Gk hekatón, L centum,
Skt śatam

*kuō ‘dog’ > TAB ku, Lith šuõ, Skt śvá, Gk kúōn

*ĝenu ‘KNEE’ > Gk génu ‘knee’, TA kanwem̃, TB kenine, Skt jānu

*ĝomb^h-os ‘tooth’ > TA kam, TB keme

*ĝ^heu- ‘pour’ > TAB ku-

*b^heh₂ĝ^hus ‘arm’ > TA poke, Gk pēk^hus, Skt bāhu, Av bāzu

6

Satem Shift: Areal Innovation

- Although Tocharian is a centum language, there is no evidence to prove that the *absence* of the Satem Shift characterized at any time a subgroup of PIE speakers.
- Rather, the Satem Shift was most likely an innovation that occurred within the dialect area which contained speakers of what would become Proto-Indo-Iranian
- The Satem Shift is virtually exceptionless in Indo-Iranian
- However, there are a few exceptions in Slavic, and even more in Baltic.

7

Satem Shift Lexical Diffusion

- Some words in Balto-Slavic may have been borrowed from Germanic
*ǵeus- ‘taste, enjoy’ Gmc. *kausjan ‘to taste, get to know’ > R kúšatŭ ‘to eat’, vkus ‘taste’
*b^herǵ^h ‘high’ > Gmc. *bergaz ‘hill’ > R béreg ‘bank, shore’
R kupítŭ ‘to buy’, cf. NHG kaufen

Other cases are less obvious:

- *h₂eǵ ‘sharp’ > R ostryj, Lith aštrus ‘id.’ (regular),
but R kámenŭ ‘stone’, Lith akmuo
- *ǵer- ‘horn’ > ? R koróva ‘cow’, Lith karvė
- *ǵes- ‘cut’ > R kosá ‘scythe’, Skt śástram- ‘knife’

8

Common Late Vocabulary

- Tocharian also lacks most of the ‘late’ (or borrowed) technological vocabulary that many of the other languages share, esp. B-Sl. and Gmc.
‘silver’ → Proto-Sl. *sъrebro, R serebró, Gothic silubr (B-Sl.& Gmc. only)
‘kitchen’ → NE kitchen, R kúxnja ‘id.’
‘trough, tub’ OE ealdop ‘trough, tub’, R lód-ka ‘boat, rowboat’
‘serving tray’ Gothic biuþs ‘table’, R bljúdo ‘dish, course of a meal’
*b^hāg-o-s ‘beech tree’ > Gmc. *bōkō > R búkva ‘letter of the alphabet’,
NE book (< writing runes on beechwood staff)
‘physician’ → Go lēkeis ‘physician’, R lekástvo ‘medicine’
‘onion’ → Gmc *lauk- (NE leek), R lúk ‘onion’

9

Borrowings into Tocharian

- Late borrowings into Tocharian are principally from Eastern Indo-Iranian languages such as Saka
- There are a few old Chinese borrowings (Proto-Sinitic *gləw > *klu* ‘rice’, units of measurement, *ʃak*, *tow*)
- TB *pito* ‘price’ = Saka *pīha*
- TA *pare*, TB *peri* ‘price’ = Skt *pīra* ‘that which must be paid’
- T *kertte* ‘sword’ < IIr., cf. Avesta *kərəta* ‘dagger’
- One possibly very early borrowing:
*gordeb^hós ‘ass, onager’ > Skt *gardab^há*, TB *kercapo* — very archaic borrowing from Indo-Iranian (before 2000 BC)

10

Indo-Iranian and Tocharian timeline

- c 4000-3000 Proto-Indo-European developed without Anatolian; beginning of separation of dialects (→ "branches")
- c 3500-2500 Afanasevo culture (poss. ancestral to Tocharians) in the extreme east
- c 3000-2000 Proto-Indo-Iranian period ?
- c 2300-1000 Andronovo culture probably associated with Indo-Iranian speakers
- c 2100 Oxus Civilization; Bactria-Margiana Archeological Complex with unclear connection to PIE

11

- c 2000 Separation of Iranian from Indic
- c 1900 Indo-Aryans' migration in to Pakistan and Western India; end of the Indus Valley Civilization
- c 1800 Tarim Mummies in the Tarim Basin probably left by nomadic speakers of the easternmost dialect of PIE
- c 1500 Mitanni inscriptions contain Indic royalty names and technical terms relating to horses; possibly showing an Indic-language-speaking aristocracy
- c 1500- Nomadic Iranians in southern central Asia (modern Turkmenistan, Uzbekistan, Tajikistan) and northern Afghanistan; probable composition of Old Avestan texts
- c 1500-1200 Composition of the Rig Veda in India

12