The writing system:

- Variant of the Indic Brāhmī script.
- Gupta script.
- Tocharian addition: “Fremdzeichen”, in transcription marked as ä.
- A number of graphemes are used only for Sanskrit loans.
<table>
<thead>
<tr>
<th>Based on Sander &amp; Malzahn</th>
<th>a</th>
<th>tha</th>
<th>ma</th>
<th>ra</th>
<th>la</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tarim Gupta</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
<tr>
<td>Tarim Brahmi A</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
<tr>
<td>Tarim Brahmi B</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
<tr>
<td>Tarim Brahmi North 1</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
<tr>
<td>Tarim Brahmi North 2</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
<tr>
<td>Tarim Brahmi South</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
<td>ऐ</td>
</tr>
</tbody>
</table>
Dating of Tocharian texts

- Precise dating in texts from the first half of 7th century CE.
- C14 dating.
- Chronological dating: MQ-writing.
Tocharian B text from the King Swarnatepe (Pinault 1987)

PK Bois B 25

a1  [i̯a̯]m ṣe kṣumntsa piṅkce [mem]ne  •  [i] /// during the 21st year, in the 5th month ...

a2  orocepi lānte swarnatepi kṣ(u)m[n](e) /// in the year of the great king Swarnatepe

a3  [r]e [ā]k(ṣ)a  •  ///
C14 dating dates texts:
- Tocharian B 400-900
- Tocharian A 700-1000
  (Tamai 2005)

THT 333 (=TB, Vinayavibhaṅga), by C14 dated between 394-473
  (Malzahn 2011-2018)(Tamai 2005)
Image source: idp.bl.uk
Most archaic texts in Tocharian B are much less uniform with respect to language and ductus.

MQ-writing (Malzahn 2007).

THT 1233, containing MQ features (Malzahn 2011-2018)

Image source: idp.bl.uk
The Tocharian Corpus

- Manuscripts and wallpaintings, from 4-1100 AD, with Buddhist content.
- Found in cave monasteries and in the desert sand in Eastern Turkestan.
- Tocharian A: basically literary texts.
- Tocharian B: all kinds of texts.
- Total number of fragments: around 7600.
- Many very fragmentary!
Locations of documents

6. Qizil.
7. Urumchi (Regional Museum of the province of Xinjiang): Yanqi manuscripts of MSN.
Tocharian A and B:
- Monastic disciplines (Vinaya)
- Doctrinal texts (Sūtra)
- Law reflections (Abhidharma)
- Narrative texts
- Stories from the births and past lives of Buddha (Jātaka, Avadāna)
- Drama (Nāṭaka)
- Poetry, epic (Kāvya), Lyric (Stotra)
- Science and technique
- Grammar, astronomy, medicine, dream interpretation, writing exercises
- Manichaean literature

Parallel texts in:
- Sanskrit/Buddhist Sanskrit (Vinaya, Sūtra, Abhidharma, medical texts, Jātaka, Avadāna, Kāvya, Stotra, grammar, astronomy, medicine, dream interpretation)
- Uighur (MSN)
- Chinese
- Khotanese
- Tibetan
DOCUMENTS: PROFANE CONTENTS

- Historic documents
- Economic, administrative, juridical documents
- Monastery letters, business letters, processes, caravan passports
- Graffiti and inscriptions
- Glosses (basically in Sanskrit texts)
TEXT SAMPLE 1: THE ’LOVE LETTER’

... (ya)-
a1 -ltse pikala watāṃ (·)e (−) ci(ʃʃ)e /// /// (− −) (·)nts(·) p(o)
ask(a)sk(au) (mā) ŋ(i ci)sa noș ŋom ŋ(ε)m
a2 (wno)lme (l)āre tāka mā ra postaṃ cisa lāre māsketār-ŋ : cišše
laraumŋe cišše ārtaŋye pelke kaltta-
a3 -(r)r šolāmpa şse mā t(e) stālle šol wārnai : 2 t'aiysu pālskanoym
sanai šaryo-
a4 -mpa šāyau karttse(š) šaulu-wārnai snai tserekwa snai tā – :
a5 yāmor-ŋikte še cau ŋi α palskaŋe šarsa tusa ysal= e-
a6 -rsate cišy= araś ŋi sālkāte : wāya ci lau ke tseyrā ŋiš
a7 wetke «k»ly«†»autk-ŋ pāke poi lāklernt«nt»s cišše tsārwo sampā-
a8 -(te-ŋ) − − − − − (−) · šol pαls(k)= araš ŋi kom-kom mi-

THT 496, Literary, Tocharian B, from Qizil Miŋ-Ŏy.
(Malzahn 2011-2018)
Image source: idp.bl.uk
TEXT SAMPLE 2: THE STORY OF THE MECHANICAL DOLL

A5 (THT 638), Literary, Tocharian B, from Qizil Miŋ-Öy. (Malzahn 2011-2018) Image source: idp.bl.uk

TREATMENT 1: THE STORY OF THE MECHANICAL DOLL

a1 krośavati tātṛā«sku»s tāṣ tām praṣṭaṃ puk kāryapāṃtwāśśi nkalune māskatār puk pārkōwāntwāśśi śkaṃ sātkā(lu)-
a2 -ne māskatār tāmyo knānmune wrasaśśi pukaṃ pruccamo puk kāswoven«t»wāśśi tsmār nāṃtsu || tāṣ penu (wra)-
a3 -som arāmpātyo kaknu lkātsi kāswē knānmune ats mā tāṣ-ām tāpreṃ ats pākkās (m)ām(t ne)
a4 tsekeṣi pekeṣi pat arāmpāt tāṣ tuṅ= kāpēṃune arāssī cāmpās pārko yatsi nu mā (cā-
a5 -mpā)ṣ: māṃt ne ālākṣyāṃ praṣṭaṃ ālāk saṃ pekant yāntācāres waṣṭā lokit yeṣ tmās sām yāmtrā(cā-
a6 -re) (pe)kāntām puk wākānā yēmurāś oṣēnī pāktāk waṣṭām lake raksā-ṃ kūntcī(t)šim şā(lyp-
b1 -aṣṣāl) (orṣ)śāṃ yāntārśinām śomināṃ śkaṃ la(p)āś liyālma-ṃ sām cami śla wāktasūrñē oki kāwält(une yā-
b2 -rkā) yāmāl yo tsaraṃ eṃtsus paṃ āpā-ṃ tām nu mānt wākānā || śāmnernāṃ || proṣmīṃn oki sām tṣru sāt lkā(mām)
b3 śāri lyāk : mātṛsatsi oki mā plāc craṅkās mā karyā : śla tuṅk oki nu rāsāṃ poke (paṃ)
b4 ypamāṃ : pontsāṃ kapšīnāṃ omlyi wiksā pekāntāp: 1 || tmās sām pekant ākntsu(ne)-
b5 -yo tām orśīnām śominānāṃ yneś śomiṃśiṃ ime kālpo āṁmaśl-ākk ats pāltśāṅkās o(te tā)-
b6 -preṃ kāwältune ote tāpreṃ kuleṇcīṃ māsratsuneyo proṣmune
Tocharian – Introduction, Part 3

Sources for Studying Tocharian: Electronic Corpora

A Comprehensive Edition of Tocharian Manuscripts

Tocharian and the Tocharians

The Silk Road, the famous network of trade routes connecting China to the West, was essential not only for the transfer of material goods, but also for the exchange of language and culture. Along with many other Central Asian civilizations, Tocharian A and Tocharian B could come to flourish as literary languages in the wake of the spread of Buddhism in the middle of the first millennium CE.

Although Tocharian was spoken in Kucha and Turfan along the northern edge of the Tarim Basin in North-West China (= present-day Xinjiang), it is not related to Chinese. Instead, Tocharian A and Tocharian B constitute a separate branch of the so-called Indo-European language family, which comprises ancient languages such as Latin, Greek and Sanskrit, as well as modern languages like English and Spanish.

Towards the end of the first millennium, Tocharian became extinct. It is now only known from documents that could be preserved over a period of more than 1000 years thanks to the arid climate of the Taklamakan Desert. The documents were discovered during a series of archaeological expeditions undertaken from the end of the nineteenth century onwards, and transferred to museums in China, Japan, and Europe.

Through the decipherment of these thousands of manuscripts, precious insights into the civilization of the speakers of Tocharian can be gained. On the one hand, prose and documents give information about everyday life; on the other hand, Buddhist religious and scientific texts are evidence of a highly developed literary culture.

It is the aim of our project to make all Tocharian texts available to everyone interested, by providing photographs, text transcriptions, and English translations with a commentary on the respective linguistic, philological, and cultural aspects. The textual material is made accessible through a database with various search options.

Tocharian Manuscripts from the Berlin Turfan Collection

This is a preliminary internal edition of the Tocharian manuscripts that are preserved in the Vorderasiatisches Museum (Asia Minor Department) as part of the Turfan collection of the Kyoto University Kansai (Glottolog: Vorderasiatisches Museum im Berlin).

The data is available from the server through more than 3000 image and text transcriptions was prepared by Christine Weidinger, transcribed by Dominique Temes, digitizing and transcribing by Katharina Engel. For the time being, only parts of the collection can be made available.

Tocharian, text with commercial license (licensed)

Tocharian, text with commercial license (commercial)

Tocharian texts within the TITUS Corpus

- Textual data is available as XML files.
- All texts are available as PDF files.
- All transcriptions are available as HTML files.
- All photographs are available as JPEG files.
- All comments are available as HTML files.

Corrections of transcriptions

Technical description

Tocharian – Introduction, Part 3
Sources for Studying Tocharian: Texts and Dictionaries Tocharian A

CeDICT: A Comprehensive e-Dictionary of Tocharian

About

About the project

The project aims at producing a comprehensive e-dictionary of Tocharian, including a Dictionary and Thesaurus of Tocharian A and an etymological Dictionary of Common Tocharian. Tocharian is an extinct language, belonging to the Indo-European language family tree. The language was spoken in the region of Eastern Central Asia up to the 17th c. AD. Tocharian is a difficult language with no living daughter languages, in which dictionary resources traditionally has been a problem. Renewed interest in historical linguistics, using computational dictionary methods, has resulted in a demand for reliable and electronically available dictionary resources for extinct languages such as Tocharian. The project is a continuation of an earlier project, which resulted in the publication of the first volume of a dictionary of Tocharian A. The current project aims at completing this project, as well as make available a reliable etymological dictionary of Common Tocharian (Tocharian A and B), both resources in the form of paper and electronic dictionaries. The project, which is performed by a core team of leading tochariologists, is done in collaboration with an international project, which aims at publishing online all Tocharian texts (5,000). Furthermore, the project is also based on a number of previously unpublished data collections.
SOURCES FOR STUDYING TOCHARIAN: DICTIONARIES TOCHARIAN B
Sources for studying Tocharian: Grammars

Chrestomathie Tokharienne
Textes et Grammaire

Tocharisches Elementarbuch
OTHER IMPORTANT RESOURCES

The Tocharian Verbal System
Melanie Malzahn

Instrumenta Tocharica
Melanie Malzahn

Tocharian – Introduction, Part 3