

# Dictionary and Thesaurus of Tocharian A

Volume 1: A–J

Compiled by Gerd Carling

in Collaboration with Georges-Jean Pinault  
and Werner Winter

2009

Harrassowitz Verlag · Wiesbaden

ISBN 978-3-447-05814-8

## Contents

Acknowledgements . . . . .	VII
Funding . . . . .	VII
Introduction to Volume 1. Letters a–j . . . . .	IX
1. The project – a short history . . . . .	IX
2. Source material . . . . .	IX
3. Notes on general policy . . . . .	XI
4. Aspects of collaboration between the authors . . . . .	XIII
List of symbols . . . . .	XV
Abbreviations . . . . .	XVII
1. Grammatical abbreviations . . . . .	XVII
2. Abbreviations of languages . . . . .	XVIII
3. Abbreviations of manuscripts/texts . . . . .	XVIII
4. Other abbreviations . . . . .	XIX
Literature . . . . .	XXI
1. Commonly quoted dictionaries or publications . . . . .	XXI
2. References and source literature for grammatical identifications, restorations, corrections, translations, and derivations . . . . .	XXII
3. Source literature for parallel texts in non-Tocharian languages . . . . .	XXXV
Dictionary . . . . .	1–204

## Acknowledgements

First, I would like to thank my collaborators, Prof. Werner Winter and Prof. Georges-Jean Pinault, for a good co-operation. I would also thank Prof. Jan-Olof Svantesson, Lund University, for hosting and supporting the project, Prof. Folke Josephson and Dr. Judith Josephson, Gothenburg University, for proof-reading and commenting the manuscript, and the seminars of the Department of linguistics, Lund University, and The Swedish Collegium for Advanced Study, Uppsala, for valuable remarks. I would like to thank my husband Tomas, my parents Hans and Aina and my family for practical assistance. Finally, I would like to dedicate the book to my grandfather, Associate Prof. Per Erlanson, who took much interest in the work but who passed away before it came to publication.

Lund, March 16 2009

*Gerd Carling*

## Funding

This project has been worked out within a three-year fellowship for Gerd Carling from the Bank of Sweden Tercentenary Foundation (RJ) and a one-year fellowship from the Swedish Collegium for Advanced Study (SCAS), Uppsala. Funds for technical assistance, equipment, and travelling have been granted by the Craafoord Foundation (Lund), Elisabeth Rausing Foundation (Lund), Fil.Dr. Uno Otterstedt Foundation (Lund), New Society of Letters (Lund), Unité Mixte de Recherche 7528 (Centre National de la Recherche Scientifique/Université de Paris III/ École Pratique des Hautes Études) Mondes iranien et indien (Paris), and The Swedish Collegium for Advanced Study (Uppsala). Funds for the publication has been granted by the Bank of Sweden Tercentenary Foundation.

For this volume, Gerd Carling has received an award by the Fondation Colette Caillat de l'Institut de France in the year 2008.

## Introduction to Volume 1. Letters a–j

### 1. The project – a short history

The idea of producing a dictionary and thesaurus of Tocharian A, based on a revised and extended version of Pavel Poucha's *Thesaurus Linguae Tocharicae Dialecti A* (1955) and the wordlist of Wolfgang Krause and Werner Thomas, in the *Tocharisches Elementarbuch II* (1964), and extended by the Yanqi manuscripts, came up in 2000, first as a proposal by Prof. Dr. Werner Winter to Dr. Gerd Carling and Prof. Georges-Jean Pinault. As Tocharian philologists, we immediately agreed upon the importance and need for this project to be realised as soon as possible.

In its first phase, the project aimed at covering the basic work to clear the ground for making a dictionary of this kind. First, a careful review of the thesaurus of Poucha had to be made by inserting corrections and additions found in later publications, in particular reviews, and by inserting corrections for errors detected by the current authors. Second, it was necessary to collect bibliographical references to the texts systematically and to digitize the Tocharian A text corpus in transcribed form, including restorations and corrections. This work resulted in a preliminary version of the complete dictionary, focusing on problematic entries in earlier dictionaries, and a text/ reference database. Thereupon, a more elaborate and complete version of the dictionary was prepared, including systematic references to Tocharian B and Uighur/Sanskrit equivalents, indication of phraseologic usage, and complete references to forms and occurrences.

### 2. Source material

The dictionary is intended to include the complete material of Tocharian A (for the unpublished material see below) as attested from manuscripts discovered in sites of the Eastern area of the Tarim Basin: Yanqi, Shorchuq, Khocho (Gaochang), Yar-khoto, Toyuq, Murtuq, Sängim, and Bāzāqliq. Most of the manuscripts are assembled in three different collections: 1) Turfan Collection, Berlin, 2) Collection of the Xinjiang Museum, Urumchi, and 3) Collection Pelliot, Paris. In addition, a few fragments are kept in the St. Petersburg Collection and the London Collection.

The largest collection is preserved in the Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, as part of the Turfan Collection of the Berlin-Brandenburgische Akademie der Wissenschaften. 467 of these manuscripts were published in transliteration by Emil Sieg and Wilhelm Siegling in *Tocharische Sprachreste* (1921). Passages from fragments not included in this publication were quoted elsewhere, especially in the *Tocharische Grammatik* (1931) published by Emil Sieg and Wilhelm Siegling, with the collaboration of Wilhelm Schulze. Many documents of the Turfan Collection were destroyed during World

War II or have been lost otherwise, but the remainder, including older photos of the manuscripts (published in the volume of plates of the *Tocharische Sprachreste*), have been digitized and are now partly available under the acronym THT (see below the abbreviations of manuscripts) as part of the project TITUS:Tocharica, directed by Jost Gippert, Frankfurt a. M. University.

The texts published by Sieg and Siegling (1921) are included completely in the present Thesaurus. In the mass of unpublished manuscripts, those in Tocharian A (around 620) have been identified and transcribed by Georges-Jean Pinault in 2004–2005: most of them are very fragmentary, but it has proved possible to find in this material words that had been quoted by Sieg and Siegling in comments to their edition (1921) and in their grammar (1931). The present Thesaurus does not intend to cover the totality of this material, which is still waiting for scientific publication: only safely identified items and interesting passages have been included. Finally, there are Tocharian (A and B) interlinear glosses in Sanskrit manuscripts of the Berlin collection (published under the acronym SHT). Thanks to the compilation of these data by Melanie Malzahn (2007), published in the book *Instrumenta Tocharica*, it has been possible to insert in the Thesaurus the Tocharian A items attested as glosses. The readings of those glosses are due mainly to K.T. Schmidt and partly to Melanie Malzahn.

Complete translations for a number of the texts included in this material were published by, among others, Emil Sieg, Werner Thomas, K.T. Schmidt, and Georges-Jean Pinault (see *Source literature*). However, for most of the texts, no complete text editions, including translation and commentary, are available.

A smaller set of texts was unexpectedly discovered in the Yanqi district in 1974. This collection, consisting of 44 [43] fragments of leaves from one text, a Tocharian A version of the Maitreyasamiti-Nāṭaka, is stored in the Xinjiang Museum, Urumchi. The material was first published in parts by Ji Xianlin and commented on by, among others, Werner Thomas, Georges-Jean Pinault, and K.T. Schmidt, and later published in complete form, including translation, commentary and photos of the manuscripts, by Ji Xianlin, Werner Winter and Georges-Jean Pinault in *Fragments of the Tocharian A Maitreyasamiti-Nāṭaka* (1998).

In the Pelliot Collection of the Bibliothèque Nationale de France, Paris, there is only one Tocharian A manuscript, consisting of six leaves, of which three can still be read without any technological implement: Pelliot Koutchéen N[ouvelle] S[érie] 1–3. This material has been included completely in the Thesaurus. So has the text MG1 from the Musée Guimet, which probably also belongs to the findings of the Pelliot Mission in Central Asia, but which was not included in the collection of the Bibliothèque Nationale de France.

Still, there are other Tocharian A fragments in various collections, which have been neither identified nor published. Consequently, the corpus covered by this dictionary should not be considered as complete. However, in principle we include all the material that has been effectively published or is otherwise available to the authors of the dictionary. The intention of the dictionary is to reflect the present state of knowledge of Tocharian A lexicography. We do intend to include all newly available material in the forthcoming volumes.

Accompanying this text material mentioned before, there is a rich literature dealing with all kind of issues, phonological, morphological, syntactic, lexical or philological. Restorations,

corrections, translations, commentaries or discussions on individual problems of a particular text, might be found anywhere within this literature. Therefore, the compilation of the present database in which bibliographical references to individual texts are systematically collected, has been of great importance for the project.

Parallel texts in non-Tocharian languages, particularly in Sanskrit and Uighur, but also in Chinese, Tibetan, Khotanese, etc., are of great help in understanding and translating complicated texts and passages properly. In recent years, a number of texts that are either directly translated from or into Tocharian or have a content equivalent to the Tocharian texts, have been published. The knowledge gained from comparison with these parallel texts has been an important source in the re-interpretation of lexemes, passages and texts in the dictionary. For this reason, parallels as for lexemes, constructions, or passages, have been introduced in the entries, beside the Tocharian B equivalents.

### 3. Notes on general policy

The dictionary will include as much important information on Tocharian A as feasible – without being too extensive and spacious. Even if the work is based on already existing dictionaries, such as Pavel Poucha's *Thesaurus Linguae Tocharicae Dialecti A* and Wolfgang Krause and Werner Thomas' *Tocharisches Elementarbuch II*, each lemma has been expanded by important information that has been gained through later research, in particular from recent decades. It should be mentioned that, in all problematic cases, the readings of the texts have been controlled through inspection of the manuscripts, in original form or in photo, when available.

The lemma presents first the lexeme (in bold/italics) in its *basic form*, i.e., nominative singular for nouns, nominative singular masculine for pronouns, adjectives and numerals, and the stem for verbs. Thereupon, word-class is indicated (within round brackets) by one or more of a number of fixed abbreviations (cf. *Grammatical Abbreviations*), and a translation is given within single quotation marks (''). For the nouns, the gender is stated if it can be ascertained from the occurrences or from the morphological class. If the lexeme, in the form that it is referred to, has more than one usage or translation, the different usages/translations are represented by different numbers, 1), 2), 3) etc., which then recur in the **F(orms)** and **T(hesaurus)** sections (see below). Directly after the translation, information is given about parallel forms in Sanskrit or Uighur (in round brackets), if available, and the Tocharian B equivalent [in square brackets], if available. A parallel form is quoted here only if it is attested from a parallel or bilingual text, and not if it, for example, is a source of the Tocharian word, which would be given in the **D(erivation)** section.

After the lemma, the entry is divided by letters representing different pieces of information as follows:

**L = L(iterature)** concerning previous translations of the lemma.

Most references in this section are quoted in short form, using standardized abbreviations for the most common reference works. For well-known lemmata, where there is no doubt about the translation, only standard works are quoted (cf. *Source literature*). This section

does not contain any discussion concerning the meaning of a lexeme. When necessary it is provided under the heading **R** (see below).

**P = Paradigm** (of verbs)

This section occurs only for verbs and gives a complete paradigm, including also non-attested but synchronically reconstructed categories.

**F = Morphological inventory of F(orms).**

Under this heading, all grammatical forms, including different variants, of a lexeme are listed. A fixed order of categories (case, gender, etc.) is consistently used. Just as in the text examples, hyphen (-) is used to distinguish the following categories: members of compounds, verbs and clitic pronouns, nouns and the instrumental *-yo*, and particles that might be divided in the text, cf. for instance *äntsān-ne*.

**S = S(yntax) and phraseology.**

Here, common or specific phrases, formulas and constructions of a lexeme are listed. Just as in the lemma, Sanskrit and Uighur parallels, if available, are given (in round brackets).

**T = T(hesaurus).**

This section contains a complete list of the occurrences of an individual lexeme. The occurrences are separated by semicolon (;), and examples of an occurrence by comma (,). The same order as in **F(orms)** is used. A selected number of examples is quoted and translated. As in the lemma and in **S(yntax)**, Sanskrit and Uighur parallels, if available, are given (in round brackets) directly after the example. Bare text numbers (247, 125, 340, etc.) refer to Tocharian A texts of the Turfan Collection as published by Sieg – Siegling (1921). Unidentified unpublished texts of the Berlin collection are mentioned as “Berlin ms.”. Identified Berlin texts not included in Sieg – Siegling (1921) are given by their THT-number (from TITUS:Tocharica). “YQ” refers to the Yanqi texts as published by Ji – Winter – Pinault (1998). The manuscript numbers (1.29 etc.) of this collection are not given, but, instead, the numbers III.3 etc., referring to the chapter of the respective manuscript of the Maitreyasamiti-Nāṭaka, are used. Elsewhere, manuscripts are reproduced as in their respective publications (cf. *Abbreviations*). Quoted texts are reproduced in transcription only, not transliteration, which means that Fremdzeichen is not marked, i.e., the difference between *nq* and *nā* is not indicated in the text. Furthermore, virāma and superscribed <sup>a</sup> at the end of a word are omitted. Otherwise, all signs given in the texts are reproduced in the examples. The complete transliterations can be found in Sieg and Siegling’s original edition of 1921, on TITUS:Tocharica homepage, and in Ji, Winter and Pinault’s edition of the Yanqi manuscripts (see 2. *Source material*). The restorations, corrections, and translations have been taken from many different sources, which will not be given reference to for each individual translation. The basic sources are (authors in alphabetical order): Bernhard (1958), Carling (2000), Couvreur (1955/56 = COU, 1956, 1959 and passim), Dietz (1981), Geng – Laut – Pinault (2004a, 2004b), Hackstein (1995), Hilmarsson (1996) = HILM, Ji – Winter – Pinault (1998) = JWP, Knoll (1996), Kölver (1965), Lane (1947, 1948), Pinault (1989a, 1990, 1991a, 1997a, 1997b, 1999a, 2001a, 2002a, and passim), Schmidt (1974, 1983, 1987, 1989, 1994a, 1994b, 1999a, 2001a, 2002), Sieg (1944, 1952), Thomas (1952,

1954, 1957, 1960, 1983b, 1990, 1991 and passim), Winter (1981, 1985a, 1988, and passim, see now the index of his *Selected Writings*, 2005).

**D = D(erivation)** and etymology.

In this section the affiliation of the lexeme is discussed in brief. Because of the many problems connected with words of Indo-European origin or very ancient borrowings in Tocharian, these matters are not considered here. Words of Tocharian B that are equivalent or related to words of Tocharian A are systematically mentioned, but the reconstruction does not go beyond the Common Tocharian stage. Accordingly, the vast literature concerning Indo-European etymologies is not taken into consideration. Other dictionaries (Van Windekens 1976, Adams 1999, Hilmarsson 1996) or discussions elsewhere should be consulted on this highly complicated issue. The reason is partly that in previous literature etymological analysis often has been done too quickly, before the proper meaning of an item had been established by philological methods. Actually the number of Tocharian words for which absolutely certain Indo-European etymologies can be assumed is relatively low, and these items have been identified for a long time. By contrast, there are loanwords on each line of every Tocharian text, and most of them are not yet correctly described. Recent and obvious borrowings, such as from Sanskrit, Middle Iranian, or Chinese, are identified, referring to relevant dictionaries, such as *Sanskrit-Wörterbuch der buddhistischen Texte aus den Turfanfunden* or the like (see *Commonly quoted dictionaries or publications*). Internal reconstruction done to trace back a form to Common Tocharian or just within Tocharian A, is mentioned here, including references to literature.

**R = R(eferences)** and comments to interpretation of lexemes and passages.

Bibliographic references to discussions concerning the interpretation of a lexeme, relation to other lexemes or items, translations of passages, indication of parallels in Sanskrit, Uighur or Tocharian B and so forth, are given here. For the purpose of clarity, the sections F(orms), S(yntax) and T(hesaurus) do *not* contain any bibliographical references. Therefore, this section might offer information of relevance for issues in previous sections as well, for instance problems of translation of an individual passage under T(hesaurus), or a construction under S(yntax). The policy for references has been as follows: references that primarily discuss the meaning or form of the particular item are given preference. References that simply discuss or give translations of individual passages under the T(hesaurus) section are not given, with the exception of references to passages that are crucial to the translation of the item itself. For references to translations of text passages we refer to the TAL database (cf. below). As a general policy, more recent literature (i.e., later than TEB II) is given preference.

#### 4. Aspects of collaboration between the authors

The compilation of this dictionary is a joint work of Gerd Carling, Lund, Georges-Jean Pinault, Paris, and Werner Winter, Kiel. The work with the text/reference database TAL that has served as the main source for restorations, reconstructions, translations, and discussions has been done by Gerd Carling. The basic work with the dictionary, i.e., compil-



ation of items, identification/translation of lexemes, identification of forms, identification/restoration/ translation of passages has been worked out by Gerd Carling. Some parts have been worked out by Werner Winter and Georges-Jean Pinault. The material has then been discussed by all three authors, separately or together. Werner Winter and Georges-Jean Pinault have often proposed different or new translations of words, and these proposals have been introduced with a remark under the L section (WW or PIN). The translation given in the head is normally the translation on which there is almost general agreement between the collaborators. Further, Werner Winter and Georges-Jean Pinault have improved translations, identifications, and restorations. The reading of unpublished A texts from the Pelliot Collection and identification and reading of unpublished A texts in the Turfan Collection has been done by Georges-Jean Pinault. Werner Winter has worked separately with the verbs and developed the system of organising stems, given in the lemma and in the Paradigm and Form sections. Werner Winter has also worked on identification of unpublished THT texts. Finally, the complete volume has been revised by Georges-Jean Pinault.

We are fully aware that there might be limitations and shortcomings in the present work. But we hope that the present volume will stimulate philological research on Tocharian, about a century after the decipherment of this language as belonging to the Indo-European family. Therefore, for the next volumes we would welcome additions, corrections and improvements proposed by fellow researchers.