

Dictionary and Thesaurus
of Tocharian A

Gerd Carling and Georges-Jean Pinault

Dictionary and Thesaurus of Tocharian A

2023

Harrassowitz Verlag · Wiesbaden

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet
über <https://dnb.dnb.de> abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the Internet
at <https://dnb.dnb.de>

For further information about our publishing program consult our
website <https://www.harrassowitz-verlag.de>

© Otto Harrassowitz GmbH & Co. KG, Wiesbaden 2023
This work, including all of its parts, is protected by copyright.
Any use beyond the limits of copyright law without the permission
of the publisher is forbidden and subject to penalty. This applies
particularly to reproductions, translations, microfilms and storage
and processing in electronic systems.
Printed on permanent/durable paper.
Printing and binding: Memminger MedienCentrum AG
Printed in Germany
ISBN 978-3-447-12002-9

Contents

Acknowledgements	VII
Introduction	IX
1. The project – a brief history	IX
2. Source material	X
3. Differences between DTA (2009) and the current volume	XII
4. Notes on general policy	XIII
5. Overview of contributions to the current volume	XV
List of symbols.....	XIX
Abbreviations.....	XXI
1. Grammatical abbreviations.....	XXI
2. Abbreviations of languages	XXII
3. Abbreviations of manuscripts/texts.....	XXIII
4. Other abbreviations	XXIV
Literature	XXV
Dictionary	1

Acknowledgements

The completion of this book has been a long process and many people ought to be acknowledged for their contribution. First, I acknowledge my co-author, Georges-Jean Pinault, for all his effort and patience with this project. I also thank the team in Paris, Timothée Chamot-Rooke, Athanaric Huard, Véronique M. Kremmer, and Boris Petipas, for their various contributions—without them the project would not have been completed. I also thank members of my team in Lund, who worked for several years with data feeding and controlling, including Filip Larsson, Victor Bogren Svensson, and Olof Lundgren. I thank Rob Verhoeven and Edin Kuckovic for handling the technology and constructing the database. Further, I thank Melanie Malzahn, Hannes Fellner, Michaël Peyrot, Bernhard Koller and Fanny Meunier, for various contributions and for setting up the project in the framework of the CeTOM database. I thank Olav Hackstein for donating the Nachlass of Rudolf Dietz and for additional contributions to the dictionary. I also thank my colleagues at the Department of Linguistics at Lund University, where I have been working with this project since 2003. Finally, I would like to thank my family for their support.

Lund, December 13, 2022

Gerd Carling

Funding

The letters A-J were funded by the Bank of Sweden Tercentenary Foundation (RJ) and The Swedish Collegium for Advanced Study (SCAS), Uppsala. This material was published as *A Dictionary and Thesaurus of Tocharian A. Volume 1: A-J*, Compiled by Gerd Carling in Collaboration with Georges-Jean Pinault and Werner Winter (Harrassowitz Verlag, 2009).

The work on the complete the dictionary, including an update of the initial letters A-J, was funded by the Swedish Research Council, Grant 2013-01097_VR.

Additional funding sources

Funding for the electronic publication, organizing a workshop, and travelling was granted by the Elisabeth Rausing Foundation (Faculties of Humanities and Theology, Lund University).

The work on the texts of the Pelliot Collection (Paris) in Tocharian A and the investigation of the phraseology and terminology of Tocharian Buddhism based on parallel texts from Central Asia belong to two focused axes of the HisTochText project (History of the Tocharian Texts of the Pelliot Collection), ERC Advanced Grant, directed by Georges-Jean Pinault (EPHE, PSL). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 788205).

Introduction

1. The project – a brief 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), the wordlist of Wolfgang Krause and Werner Thomas in the *Tocharisches Elementarbuch II* (1964), and extended by the Yanqi manuscripts (Ji – Winter – Pinault 1998), emerged in the year 2000. The project was granted funding by the Bank of Sweden Tercentenary Foundation for 2003-2006 (prolonged until 2007) and by Swedish Collegium for Advanced Study, Uppsala for 2004-2005.

In its initial form, the project aimed at covering the basic work to prepare a complete dictionary of the Tocharian A language. First, a careful review of the thesaurus of Poucha (1955) had to be made by inserting corrections and additions found in later publications, in particular reviews of Poucha's book, and by inserting corrections for errors. Second, it was necessary to collect all bibliographical references referring to the texts and to digitize the Tocharian A corpus in transcribed form, including restorations and corrections. This work resulted in three digital materials, 1) a draft version of the complete dictionary, focusing on problematic entries in earlier dictionaries, 2) a reference database with text numbers and references to occurrences in the literature, and 3) a text corpus of Tocharian A, including restorations and partial translations. All three components were incomplete at the end of the initial project (2003-2007). Thereupon, we prepared a more elaborated and complete version of the dictionary, including systematic references to Tocharian B and Uighur/Sanskrit equivalents, indication of phraseology, and complete references to forms and occurrences. This volume included 1/3 of the Tocharian dictionary and was published in 2009, with the title *Dictionary and Thesaurus of Tocharian A. Volume 1: A-J*, by Gerd Carling in collaboration with Georges-Jean Pinault and Werner Winter (DTA 2009). The text and reference corpora that formed the basis for this publication, in the published volume referred to as TAL: Text and Reference Database of the Tocharian A Language (with a DOI at Lund University), was in 2012 moved to the CEToM database in Vienna. The material has been available from this platform since 2012 (<https://www.univie.ac.at/tocharian/>). The DOI at Lund University referred to in DTA (2009) is no longer valid. Since 2012, the texts, references, editions, and translations of Tocharian A have been continuously improved in the CEToM database, but the prospected second and third volume of the Tocharian A dictionary were halted due to lack of funding. Meanwhile, the collaborator Werner Winter passed away in 2010.

The project of completing the Tocharian A dictionary received renewed funding from the Swedish Research Council for 2014-2016 and was prolonged until 2017, due to other grants and commissions by the project leader Gerd Carling. Within this project, we constructed a complete lexical database at Lund University by the name *CeDICT: A Comprehensive e-Dictionary of Tocharian* (<https://projekt.ht.lu.se/cedict/>) for the purpose of producing the complete dictionary. The digital version of the Dictionary and Thesaurus of Tocharian A (DTA 2009), which was automatically transferred to CEToM, was also automatically transferred to the new database. The CeDICT database contained an interface for feeding and viewing lexical lemma and was built to reflect the layout of the dictionary (DTA 2009). However, an important innovation of the CeDICT database was the automatic linkage of occurrences in the database to the corresponding manuscripts in CEToM.

The current dictionary is extracted from this database. It is published in its complete form as a monograph. This means that the first section of the alphabet (A-J), which was published in DTA (2009)

has been carefully reviewed and updated (see 3. below). The electronic version of the dictionary, including all the search functions of the database, is published with another publishing house, Brill, within the series Indo-European Etymological Dictionaries Online (<https://brill.com/view/db/iedo>).

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 Central and 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. Most of these texts, including transcription and a digitized version of the manuscripts, are available on CEToM, either public or behind password. Complete translations for several texts in Tocharian A have been published by various authors (see *Source literature*). These translations form the basis for the reference database and the translations in DTA (2009) and the Tocharian A translations in CEToM. At the time of the publication of DTA (2009), the CEToM database was not available, and therefore this volume contained a substantial corpus of text samples, including restorations and translations. However, since this material is now available via the CEToM database, the current volume contains fewer text samples and translations.

We distinguish several text collections, explaining their status and how they have been treated in this dictionary.

Berlin texts A 1-467 (in the dictionary given as numbers only, without a preceding “A”). The largest collection of Tocharian A texts is preserved in the Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, as part of the Turfan Collection of the Berlin-Brandenburgische Akademie der Wissenschaften (now Berlin-Brandenburg Academy of Sciences and Humanities). 467 of these manuscripts were published in transliteration by Emil Sieg and Wilhelm Siegling in *Tocharische Sprachreste* (1921). The dictionary authored by Poucha (1955) was based on this edition as well as on the *Tocharische Grammatik* (1931) by Emil Sieg (1866-1951) and Wilhelm Siegling (1880-1946) in collaboration with Wilhelm Schulze (1863-1935). In the current dictionary, an additional source is the personal and annotated copy of the *Tocharische Sprachreste* by Wilhelm Siegling, which was scanned by Douglas Q. Adams and communicated to Melanie Malzahn (Vienna), who transmitted it to Georges-Jean Pinault, as well as to a few other scholars. This copy contains numerous notes and corrections of readings which are extremely valuable, reflecting the high expertise of Siegling, since notes were based on manuscripts which have later been lost. These notes by Siegling were written after the publication of the *Tocharische Grammatik* (1931). By chance, several important and large manuscripts had been photographed and included in the volume of plates of the *Tocharische Sprachreste* (1921). For the present dictionary as well as for DTA (2009), all problematic or difficult passages have been controlled again by examination of all available photographs and—in critical cases—of the original manuscripts kept in the Turfan Collection, Staatsbibliothek zu Berlin–Preussischer Kulturbesitz, Orientabteilung.

Berlin THT texts. Passages from fragments not included in Sieg – Siegling (1921) were quoted elsewhere, especially in the *Tocharische Grammatik* (1931) 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 remains, including older photos of the manuscripts (published in the volume of plates of the *Tocharische Sprachreste*), were made available under the acronym THT (see below the abbreviations of manuscripts) as part of the project

TITUS:Tocharica, directed by Jost Gippert, Goethe University, Frankfurt a.M. In the mass of unpublished manuscripts, those in Tocharian A were identified and transcribed for the most part by Georges-Jean Pinault in 2004-2005 and included in DTA (2009). Since then, the remaining text fragments (in Tocharian A as well as in Tocharian B) from Berlin have been made available, with transcriptions by Tatsushi Tamai, Georges-Jean Pinault, and Michaël Peyrot via the CEToM database (several of them are not yet public), and more recently by Ilya Itkin (2019). The latter treatment retrieves around 1050 fragments in Tocharian A and supersedes definitely the previous transliterations. The printed work of Itkin (2019) was received in the midyear 2020, while the final revision of the present dictionary had started. Most of these manuscripts are very fragmentary, but it has been possible to find words in this material that have been quoted by Sieg and Siegling in comments to their edition (1921) and in their grammar (1931). The present dictionary does not intend to cover the totality of this material; many readings are still uncertain and most of the material is still waiting for scientific publication. In the Thesaurus part (T) of each lemma, we have given priority to THT fragments containing new forms or alternative variants of words which were previously recorded. As a rule, the systematic transliteration of fragments containing only a few letters contain words previously recorded in Poucha (1955), and do not add anything new to Tocharian A lexicon. Several publications up to date include THT fragments, identifying Sanskrit or Uighur parallels and joining these manuscripts to better known and longer manuscripts (Itkin 2002, 2014, 2015, 2018, 2020; Burlak and Itkin 2004, 2020; Itkin and Malyshev 2016, 2019, 2021; Malyshev 2017, 2019, 2022). The results retrieved from these publications are included in the present dictionary. Since the Berlin THT material in this volume stems from various sources (TITUS:Tocharica, CEToM, Itkin 2019), and reflects different states of inclusion over the 20-year period of the production of the dictionary, there might be discrepancies in the numbering of texts and lines, as well as in readings, of this material. In principle, we follow Itkin's index verborum (2019) for numbering and identification of recto/verso and line numbering, but the dictionary may show traces of earlier versions, such as CEToM. Itkin's work (2019) mentions at times specific contributions by Sergey V. Malyshev to the decipherment of some items, but those have not been specified among the current references to THT on the basis of Itkin 2019.

Glosses. Further, there are Tocharian A interlinear glosses in Sanskrit manuscripts of the Berlin collection (published under the acronym SHT). These glosses are available on CEToM and have been published in several books and papers (see Malzahn 2007b, Peyrot 2014 and 2015a, which give further references) and are included in the current volume.

Yanqi manuscripts. A smaller set of manuscripts was unexpectedly discovered in 1974 in the Yanqi district of Xinjiang, China. 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 Klaus 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). This material is included in the CEToM database as well as in the current volume.

Paris texts. In the Pelliot Collection of the Bibliothèque Nationale de France, Paris, which contains mostly fragments in Tocharian B, there are a handful of Tocharian A manuscripts. One set consists of six leaves of the same manuscript, found in Kucha region: Pelliot Koutchéen N[ouvelle] S[érie] 1-6. Most of these pieces are quite difficult to read, partly because of superposition of lines by imprints from one leaf to another. A provisory reading had been made by Pinault during the 1990s, and the first 4 items have been published online on CEToM in August 2014. These leaves have been

investigated more deeply through imaging techniques in 2019-2020 as part of the project HisTochText (History of the Tocharian Texts of the Pelliot Collection, ERC Advanced Grant. Horizon 2020, Action Number 788205), directed by Georges-Jean Pinault. Thanks to new photographs, it proved possible to read larger parts of the text: a complete edition by Athanaric Huard, Kilian Laclavetine and Georges-Jean Pinault is currently in preparation. The material which has been obtained so far has been included in the Thesaurus. This holds also for some little fragments (from the Pelliot Koutchéen and Pelliot Sanskrit collections) which have been identified by Huard in the meantime. So has the data of the text MG1 from the Musée Guimet (edited by Pinault 2007a), 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 fully complete. However, in principle we aim at including 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, as well as to give a sound basis for further investigations.

Accompanying the text material mentioned above is a rich literature dealing with all kinds 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 database in which bibliographical references to individual texts are systematically collected, has been of great importance for the project. Since 2011, this database is incorporated with CEToM. For the current volume, we have systematically searched and included publications since 2008 for additions and improvements of the Tocharian A text material.

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. Differences between DTA (2009) and the current volume

The current volume contains an updated version of the volume DTA (2009), which has been completed by the remaining letters, covering the entire language. The volume DTA (2009) was published before the text corpus and reference database was made available via CEToM, and therefore all lemma of DTA (2009) contained substantial amounts of text samples with translations, which were extracted from the corpus. However, since the Tocharian A corpus currently is almost completely available open access (several THT fragments are still behind password) via the CEToM database, we have decided to include less samples of text and translation in the current volume.

As for the first part of the volume, which has been migrated from the text-file of DTA (2009) into the database CeDICT and then extracted into the current volume, there are some improvements:

- Reviews of DTA (2009) have been taken into consideration, in particular Itkin (2011), Malzahn (2011), and Peyrot (2011).
- Verbal roots have been standardized, including, or referring to suggestions in *The Tocharian Verbal System* (TVS).
- Forms and occurrences have been cross-checked with CEToM.
- Names of tunes have been checked against CEToM, adopting their standard for the restoration of nominative forms.

4. 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. 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 by one or more of a number of fixed abbreviations (cf. *Grammatical Abbreviations*). For the nouns, the gender is stated if it can be ascertained from the occurrences or from the morphological class. This has been thoroughly controlled by Timothée Chamot-Rooke (PhD student, EPHE, PSL) under the supervision of Georges-Jean Pinault. Translation is given within single quotation marks (‘’) under the heading Meaning. 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). After the translation, information is given about the Tocharian B equivalent, if available, and parallel forms in Sanskrit or Uighur, if available. A parallel form is quoted here only if it is attested from a parallel or bilingual text only, 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 introduction, 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 lemma, 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, and particles that might be segmented in the text, cf. for instance *āntsan-ne*. However, contrary to the policy of the 2009 volume (DTA), the casual affix *-yo* of the instrumental has been treated as the other casual affixes expressing the so-called secondary cases.

S = S(yntax) and phraseology.

Here, common or specific phrases, formulas and constructions of a lexeme are listed. 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 **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, CEToM, and Itkin 2019). “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 respective chapter (Roman number) and to the respective leaf (Indo-Arabic number) of the YQ (Yanqi) 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 *ṅa* and *nā* is not indicated in the text. Furthermore, virāma and superscribed ^ā 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, TITUS:Tocharica homepage, in Ji, Winter and Pinault’s edition of the Yanqi manuscripts and on CEToM (see 2. *Source material*). The restorations, corrections, and translations have been taken from many different sources, which are not given reference to for each individual translation.

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 [revised 2013], 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 borrowings, such as from Buddhist Sanskrit, Middle Iranian, or Chinese, are identified, referring to relevant dictionaries, such as *Sanskrit-Wörterbuch der buddhistischen Texte aus den Turfanfunden* and Edgerton’s dictionary of Buddhist Hybrid Sanskrit (see *Commonly*

quoted dictionaries or publications). Since many Sanskrit words have been obviously borrowed into Tocharian through Middle Indic (alias Prākṛit) intermediary, references have been made to Turner's dictionary (CDIAL) and to dictionaries of Pāli, as well as to other reference tools. 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, 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.

5. Overview of contributions to the current volume

The compilation of the first part of the dictionary (*a-j*, DTA 2009) was a joint work of Gerd Carling, Lund (who at the time was postdoctoral researcher), together with Georges-Jean Pinault, Paris, and Werner Winter, Kiel. The building of the text/reference database (available via CEToM) that is the main source for restorations, reconstructions, translations, and discussions was done by Gerd Carling. The basic work with the dictionary, i.e., compilation of items, identification/translation of lexemes, identification of forms, identification/ restoration/ translation of passages was also worked out by Gerd Carling. Some parts were worked out by Werner Winter and Georges-Jean Pinault. The material was discussed by all three authors, separately or all together. Werner Winter and Georges-Jean Pinault often proposed different or new translations of words, and these proposals were in the original publication of DTA (2009) introduced with a remark under the L section (WW or PIN). Further, Werner Winter and Georges-Jean Pinault 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 was done by Georges-Jean Pinault. Werner Winter worked separately with the verbs and developed the system of organizing stems, given in the lemma and in the Paradigm and Form sections. Werner Winter also worked on identification of unpublished THT texts. Finally, the complete volume was revised by Georges-Jean Pinault.

The volume of DTA (2009), reviewed and updated, is included in the current volume. For the remaining part of the alphabet (*ñ - ts*) the procedure of working out lemmas has been somewhat different compared to the first part. Werner Winter passed away in 2010, leaving behind a draft for organizing the verbal stems. This material forms the basis for the verb lemmas in the current dictionary. The project of completing the dictionary started under the leadership of Gerd Carling in 2014. Initially, Rob Verhoeven (engineer, Lund University) in collaboration with Edin Kuckovic (engineer, Lund University), built a database and an editor interface (2014-2015). Filip Larsson (research assistant, Lund University) introduced and checked all occurrences from Poucha (1955) (POU), controlling against the database CEToM. In addition, he introduced the review by Couvreur (Couvreur 1955/56 (COU), 1966), Ji – Winter – Pinault (1998) (JWP), the Nachlass by Rudolf Dietz (VTW, see description below), and the remaining dictionary material compiled by Gerd Carling (between 2003-2007) but not included in DTA (2009). This work by Rob Verhoeven, Edin Kuckovic,

and Filip Larsson was done between 2014-2015 under the supervision of Gerd Carling. For each lemma, Filip Larsson performed a search in CEToM, checking for additional occurrences in THT texts. Fellow Tocharologists Melanie Malzahn, Hannes Fellner, Michaël Peyrot, Olav Hackstein, Georges-Jean Pinault and Fanny Meunier improved a handful of lemmas during a workshop in Lund, February 16-20, 2015. Thereupon, from 2015-2017, Gerd Carling worked over all lemmas, checking and improving the entire material, including sources, bibliographic references, and new occurrences. The first part (*a-j*) was improved by reviews of the earlier volume (see the references below in the full bibliography) and additional references later than 2009, and all verbs were checked against TVS. Gerd Carling and Georges-Jean Pinault worked together in Paris in February 2018 on problematic lemmas. Thereupon, Olof Lundgren (research assistant, Lund University) proof-read the entire dictionary for consistency, added and checked translations, and introduced glosses when missing (2018-2019). Again, Gerd Carling proof-read and controlled the entire dictionary (2019), checking Sanskrit and Middle Indic sources, verbs, and references. Gerd Carling and Georges-Jean Pinault worked together with various lemmas through correspondence and during a work session in Paris, November-December 2019, solving many problematic issues in translating lemmas and occurrences. From January 2020 until April 2022 a team in Paris, under the supervision of Georges-Jean Pinault, worked over and controlled all lemmas of the dictionary. This team included PhD students of the École Pratique des Hautes Études (EPHE, Paris Sciences et Lettres), who were preparing their dissertation on various aspects of Tocharian philology and linguistics: in alphabetical order, Timothée Chamot-Rooke, Athanaric Huard, Véronique M. Kremmer, and Boris Petipas. Since several lemmas contains lots of occurrences, the survey of the materials was distributed between the members of this team: *ñ* to *p* (Petipas), *b* to *ly* (Kremmer), and *v*, *w*, sibilants, *ts*, *h* (Pinault). The whole revision was slowed down since March 2020 and until October 2021 by the consequences of the Covid-19 pandemic on academic and research activities in France and abroad. Permanent access to the Lund database (CeDICT) has been granted since 2019 to Chamot-Rooke, Huard and Pinault, in order to introduce all changes which were written before. The work done by the Paris team was elaborated during several sessions at Pinault's office in Paris, together with Chamot-Rooke and Huard. Both assisted Pinault in checking many points, which resulted in reassessing—or establishing for the first time—many interpretations and translations of passages. The revision required to improve many items of Poucha's dictionary, beyond the corrections made by Couvreur (1955/56, 1956, 1959). The work done by Chamot-Rooke was substantial for a lot of improvements of the entire database, such as, to name only a few, the thorough revision of the demonstratives, the identification of the occurrences of the indefinite pronoun (see Chamot-Rooke 2021), the link of the imperative prefixed stem forms to their respective verbs, the identification of gender of nouns, and the identification and translation of several fragments. Kremmer revised the identification of compounds with first members *śla* and *sne*. Huard has controlled the meaning of Buddhist terms, including comparison with other languages. He has identified and translated several fragments, which had been poorly understood so far. In the meantime, Chamot-Rooke, Huard and Kremmer wrote papers (several still under the press at the beginning of 2022) about data from texts in Tocharian A, which bring significant improvements to a number of lexical, syntactic, and morphological problems. This material, as well as those from Huard's PhD thesis (Huard 2022), have been entered into the database. All data were inserted into the CeDICT database by Chamot-Rooke (80%) and Huard (20%). Specific contributions of Pinault's collaborators have been highlighted by the mention 'p.c.' following their names or by reference to their works. This mention applies also to Bai Yu, PhD student at the EPHE, PSL, under the joint supervision of Sylvie Hureau. In addition, Gerd Carling did a final and complete proof reading for

completion and consistency with some assistance of Victor Bogren-Svensson (PhD, Lund University).

A substantial part of the lemma has the source VTW in the Literature field. This refers to an unpublished manuscript by the name “Vorarbeiten zu einem tocharischen Wörterbuch von Werner Thomas and Rudolf Dietz”. In 2012, Olav Hackstein was officially asked by the Deutsche Morgenländische Gesellschaft (DMG) on behalf of Rudolf Dietz to secure this document for further scientific work. The material was made available to the current project in 2014. It is not clear how this material was produced. The material covers only the alphabet up to *ma-*, and the remaining part is either lost or not existing. The document, which is an outline of a dictionary, is written by means of an old-time typewriter, with diacritics added by hand. Besides, there are hand-written additions and comments. The parts by typewriter are probably from the 1980s or earlier and have no references to the Yanqi manuscripts or the publication by Ji – Winter – Pinault (1998). The comments added by hand contain revisions which insert Uighur parallels, referring to Tekin (1980) and Ji – Winter – Pinault (1998), including the review of this edition (JWP) by Thomas (2003). However, it should be said that the meritorious work done by Dietz under the supervision of Thomas is by now outdated, due to the advances of Tocharian philology in the two past decades.

We are fully aware that there might be limitations and shortcomings in the present work. However, we hope that the present volume will stimulate philological research on Tocharian, little more than a century after the decipherment (1908) of this language as belonging to the Indo-European family. The expected advances in Tocharian A and B lexicography will result from the better understanding of fragmentary texts, based on the joining of fragments and on the discovery of further parallel texts in other languages of Buddhism. We welcome all suggestions for additions, corrections and improvements proposed by fellow researchers.