



Universiteit  
Leiden  
The Netherlands

## The Tocharian Gender System: A Diachronic Study

Tomba, A. del

### Citation

Tomba, A. del. (2020, March 24). *The Tocharian Gender System: A Diachronic Study*. Retrieved from <https://hdl.handle.net/1887/87130>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/87130>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/87130> holds various files of this Leiden University dissertation.

**Author:** Tomba, A. del

**Title:** The Tocharian Gender System: A Diachronic Study

**Issue Date:** 2020-03-24

## CHAPTER ONE

### INTRODUCTION

This study aims to describe the evolution of the Tocharian gender system, investigating the inflectional morphology of grammatical gender in Tocharian nouns, adjectives, and pronouns.

#### 1.1. TOCHARIAN

At the end of the 19<sup>th</sup> century, ancient manuscripts were found in the Tarim Basin in Chinese Turkestan (present-day Xinjiang Uyghur Autonomous Region). In these manuscripts, different languages were documented, many of which were already known. Two of them, however, were completely unknown to the scientific community of the time. In 1908, the Indologists Emil Sieg and Wilhelm Siegling announced that the new languages were related to each other as an independent branch of the Indo-European language family (cf. already Müller 1907). As is clear from the name of the article, Sieg and Siegling named these languages “Tocharisch. Die Sprache der Indoskythen”.

Thus, Tocharian is the conventional name of two extinct Indo-European languages: Tocharian A (also named East Tocharian or Agnean) and Tocharian B (also named West Tocharian or Kuchean).

Despite the differences between the two idioms, the comparison of Tocharian A and Tocharian B allows to reconstruct a coherent picture of an immediate antecedent, which is usually named Proto-Tocharian. The differences between the two Tocharian languages are significant and testify an independent evolution during at least some centuries. Tocharian B texts are dated from the 5<sup>th</sup> to 10<sup>th</sup> centuries CE, while Tocharian A is attested over a shorter period, from the 7<sup>th</sup> to 10<sup>th</sup> centuries CE (Pinault 1989: 7-12; Tamai 2011: 370-5). From a linguistic point of view, the largest differences between Tocharian A and B are found in the phonology (the vowel system in particular), the formation of some case and verbal endings, the verbal stems, and other elements of the morphology. On the lexical level, they are similar, but some differences can equally be found both in technical Buddhist terms and in some words belonging to the basic lexicon (Lane 1966). Another significant difference between Tocharian A and B is their sociolinguistic status: while Tocharian B has dialectal and chronological layers, Tocharian A appears to be linguistically uniform. Furthermore, apart from sporadic exceptions, texts drafted in Tocharian A are of religious content, while secular documents are more common in Tocharian B. This led some scholars to think that Tocharian A was not a language of everyday use (cf. e.g. Winter 1963; Peyrot 2010a). However, recent investigations on the Tocharian A materials have cast doubt on this analysis (see Ogihara 2014 with references), showing that Tocharian A was a spoken vernacular too. Following a classification

definitively standardised by Peyrot (2008), Tocharian B is usually divided into three main linguistic periods: archaic, classical, and late. Colloquial forms can be found across all periods (Ching & Ogihara 2013).

Tocharian A documents were mostly found in monastery complexes in the vicinity of Šorčuq (near Qarašähär) and in Turfan region. Tocharian B documents were found in the same areas as Tocharian A and in the region of Kuča (particularly in the Miñ-Öy grottoes near Qizil). Most of the Tocharian leaves are damaged and sometimes abraded, and many manuscripts are mutilated and fragmentary. With the exception of fragments from two manuscripts drafted in Manichean script, Tocharian documents are written in a modified variety of the Indian Brāhmī script, commonly named North-Turkestan Brāhmī (*Nordturkistanische Brāhmī*). A peculiarity of Tocharian Brāhmī is the use of the so-called “Fremdzeichen”, signs with inherent vowel *ä*, which are not found in the Indian Brāhmī. Furthermore, the anusvāra, a diacritic used to indicate nasalisation and usually transcribed with <ṃ>, appears often in word-final position, where it always corresponds to /n/.

Tocharian shows significant effects of contact with neighbouring languages. While the contacts with Indian (Sanskrit and Middle Indian languages), Chinese, and Old Uygur are from a relatively recent date and mostly surface in loanwords, contacts with Iranian languages have taken place over a much longer period. We find loanwords from Khotanese, Bactrian, and Sogdian, but an archaic layer must derive from an otherwise unattested Old Iranian variety.

The phylogenetic position of Tocharian within the Indo-European domain is debated. Affinity has been claimed with practically every other branch of the Indo-European family (Malzahn 2016: 281), from Germanic (Adams 1984) to Celtic (Pedersen 1913), and from Slavic and Armenian (Meillet 1914; Hamp 1998) to Greek and Phrygian (Benveniste 1936; cf. Klingenschmitt 1994). In recent years, a broad, though not universal, consensus seems to be reached, according to which Tocharian was the second branch that split off from Proto-Indo-European, after the earlier departure of Anatolian (Carling 2005; Ringe 2017; Jasanoff 2017; Weiss 2018; Lundquist & Yates 2018; Kim 2018a; Peyrot 2019). The supporters of the so-called “Indo-Tocharian hypothesis” tested the evidence either with the traditional comparative method (Schmidt 1992; Ringe 1991; Winter 1997; Jasanoff 2003; Kim 2018a), or with the computational cladistic method (Ringe et al. 2002; Bouckaert et al. 2012; Chang et al. 2015).

In order to bring the discussion forward and to possibly settle the debate, two points remain crucial: (1) the identification of possible common innovations of the remaining non-Anatolian and non-Tocharian branches; (2) more generally, the relative chronology of unitary and differential structural isoglosses between Tocharian, Anatolian, and the other Indo-European languages.

Indeed, most compelling evidence said to militate in favour of the early split-off of Tocharian comes from the lexical level. This has been recently reviewed by Malzahn (2016), who concluded that these lexical isoglosses could neither prove the early split of Tocharian nor deny it.

At any rate, the lexicon is of less value for establishing cognacy and tree-branching (Ciancaglini 2009; Ringe et al. 2002: 99). Indeed, lexical and semantic arguments may only serve as additional evidence, while the method must be based on structural evidence, i.e. phonological and morphological correspondences. In this respect, verbal morphology has been the privileged area of research, where, according to some scholars, the archaic status of Anatolian and Tocharian could be suggested by peculiar linguistic traits that these two Indo-European branches share, but that are not (or are only partially) attested in other Indo-European languages (Jasanoff 2003, 2017; Kim 2007a; Ringe 2000). On the other hand, nominal morphology has been less used to substantiate the early split-off of Tocharian. Indeed, in contrast to the verbal system, it is generally assumed that Tocharian nominal morphology is less conservative and archaic with respect to what is reconstructed for Proto-Indo-European. At first glance, this assumption seems reasonable and founded. A good example is the case system, where on top of only four nominal cases inherited from the proto-language, an additional agglutinative case layer has been formed.

In recent years, however, the gender system of Tocharian has been called into question, since it has been claimed to preserve a more archaic status with respect to the other non-Anatolian languages. This thesis is also concerned with this topic.

## 1.2. TOCHARIAN GRAMMATICAL GENDER

The Tocharian gender system is not complex. It distinguishes three category values, the masculine, the feminine, and the so-called alternating gender, which mostly evolved from the Proto-Indo-European neuter. A three-gender contrast is found in most of the ancient Indo-European languages. Still, the Tocharian gender system is different from that of the ancient Indo-European languages. These differences are many and can be addressed from both a diachronic and a synchronic point of view.

In the last few years, historical linguistics oriented towards Tocharian morphology has mostly focussed on the verbal system, which resulted in the publication of two monumental monographs, Melanie Malzahn's *The Tocharian Verbal System* (2010) and Michaël Peyrot's *The Tocharian Subjunctive* (2013). A new monograph by Markus Hartmann (2013) is an exception.<sup>1</sup> This publication is a slightly revised version of his *Habilitationsschrift*, submitted in 2011. The published version is not much different from the thesis, since recent bibliography has not been included (with the exception of Hackstein 2012 and an article by Melchert later published in 2014). As suggested by the title *Das Genusssystem des Tocharischen* (2013), Hartmann's aim is to describe the Tocharian gender system from both a synchronic and a diachronic perspective, thus clarifying the status of Tocharian with respect to the other Indo-European languages. Is it therefore necessary to have another study on the grammatical gender of Tocharian?

Hartmann's book has been reviewed by two of the major experts from the Tocharian field: Georges-Jean Pinault (2015a, 20 pages) and Douglas Q. Adams (2016, 7 pages). These

---

<sup>1</sup> Hartmann's book has been recently joined by Kim's monograph on the Tocharian dual (2018).

reviews evaluate the book very differently: Pinault strongly criticises both the method and the results, while Adams considers it a rich contribution from the point of view of the analysis provided and the data collected.

In my opinion, the truth may lie somewhere in between. Hartmann's book can be divided into two sections: the first is synchronically orientated, while the second is diachronically orientated. I think these sections differ from each other not only in the final goals and in the perspectives, but also in the achieved results.

In the first section, Hartmann deals with the "gender resolution", i.e. the recognition of formal and functional patterns to predict the gender of a given noun. The data are detailed and analysed systematically, in order to provide an in-depth discussion about the gender assignment system of Tocharian from an inflectional, derivational, and semantic perspective.

On the other hand, Hartmann's diachronic analysis mainly concerns etymological problems related to the gender of some Tocharian nouns, which can be neither sorted into Tocharian inflectional classes nor traced back to common PIE stems. As a consequence, Hartmann did not deal with some other and, in my view, more important problems related to e.g. the origin of the third Tocharian gender, the so-called *genus alternans*, and the evolution of the feminine markers from Proto-Indo-European to Tocharian.

Another point that Hartmann did not consider in his book is the morphology of adjectives and pronouns. The Tocharian gender system must have arisen from formal and functional mergers of the genders inherited from Proto-Indo-European. These mergers have first taken place in the adjectival system, and they have reshaped not only the gender system, but also the nominal morphology of Tocharian as a whole. But even on a more general level, I think that grammatical gender must be considered, first of all, in light of the morphosyntactic relations between a noun and a modifier. This is just as true in the synchrony of the language, as it is in its diachrony. Indeed, in the Indo-European domain, nouns can be divided in genders only when they start to agree with targets, i.e. adjectives, pronouns, demonstratives, numerals, etc. This means that an important task for the diachronic analysis is to consider the marking of gender in the inflection of the modifiers.

Hartmann's book is part of a flourishing line of research that has the Tocharian category of gender as the central subject of investigation. The issues that this type of study face pertain to various aspects of the language from a phonological, morphological, and typological point of view. In some recent works it has been argued that Tocharian has inherited a gender system different from that of the other Indo-European languages (cf. Kim 2009, 2014, 2018b; Hackstein 2012; Kortlandt 2017). The most debated issue concerns the evolution of the feminine gender: on the one hand, it has been argued that, when Tocharian separated from Proto-Indo-European, the feminine had not yet risen as a grammatical category value; on the other hand, it has been hypothesised that Tocharian has maintained an older stage, where the suffix *\*-ih<sub>2</sub>* was generalised in both the thematic and the athematic declensions. These theories have received some support, but they have also been criticised (cf. e.g. Pinault 2008: 516f., 2012; Malzahn 2011; Fellner 2014, 2014a).

As one can see, most of these recent articles have been published even after Hartmann (2013). This is indicative of the fact that the problems linked to this topic have not been solved yet. I therefore believe that there is ample reason to look at the subject again, starting, of course, from what Hartmann has either already clarified, or left unsolved (Adams 2016: 253).

### 1.3. AIM

The aim of this thesis is to illustrate what type of gender system Tocharian has and particularly to investigate how it developed. Consequently, the approach is mainly diachronic.

Nonetheless, a synchronic part of the thesis is aimed at clarifying an important and still unsolved problem of Tocharian grammar, namely how many genders Tocharian has. In order to answer this question, I will refer to the problematic status of the third Tocharian gender, i.e. the so-called *genus alternans*, putting forward typological arguments and invoking cross-linguistic comparisons with languages that show a similar gender system.

The diachronic part has a threefold aim. It aims at:

- (1) describing the evolution of the gender markers in the inflection of the gender agreement controllers, i.e. nouns;
- (2) describing the evolution of the gender markers in the inflection of the gender agreement targets, i.e. pronouns and adjectives;
- (3) understanding what type of morpho-phonological mergers between the three inherited genders can be reconstructed.

It follows that the main subject of investigation concerns the question how grammatical gender developed over time and what type of gender system Tocharian inherited from Proto-Indo-European. The principal problems are related to the historical evolution of the feminine and the neuter.

As to how the feminine gender evolved in Tocharian nouns, I will consider inflectional classes that may be traced back to either the  $*(e)h_2$ -inflection (of both the ablauting and non-ablauting type) or the  $*ih_2$ -inflection (of both the *deví*-type and *vṛkí*-type). As far as the adjectival declension is concerned, I will focus on the relation between the “thematic”  $\bar{a}$ -inflection and the “athematic” ablauting  $*ih_2$ -inflection, in order to verify whether Tocharian has inherited a different gender marking with respect to the other Indo-European languages, and how this system evolved in the inflection of the modifiers. Furthermore, I will consider the evolution of the demonstratives and the pronominal adjectives, because they show clear formal and functional differences between the masculine, the feminine, and the Tocharian “neuter”. Indeed, the pronominal paradigm is in part different from that attested in the adjectival inflection. In addition, phonological problems related to the outcome of PIE  $*-eh_2(-) > *-\bar{a}(-)$  will be addressed, since it still represents a debated issue of Tocharian historical phonology.

As to how the neuter gender evolved in Tocharian nouns, I will investigate the outcome of selected classes of nouns, continuing both athematic and thematic paradigms. The ultimate goal is to understand how the PIE neuter lost its functions as a category of target gender and how it evolved into the Tocharian *genus alternans*.

In order to achieve these aims, I will also investigate the morpho-phonological mergers between case forms and gender markers that affected (Pre-)Proto-Tocharian in an unattested stage. These mergers have been caused by a general consonant apocope that led to confusions between the inherited case endings. In addition, I will consider typological comparisons in the evolution of languages which synchronically show a similar gender system to that of Tocharian. In particular, in the evolution from Latin to the Romance languages, the neuter gender suffered of gradual depletion. In a transitional stage, however, the neuter displayed an alternating agreement pattern, which is fully parallel to the Tocharian *genus alternans*. The same system was attested in some Old Italian dialects and survives even today in e.g. Romanian. The advantage that this type of comparison in the diachronic typology may give to the analysis of the Tocharian gender system is that the attestation of Latin as the antecedent of the Romance languages could, in a way, give insight into the Pre-Proto-Tocharian developments leading to the system attested by Tocharian A and B.

#### 1.4. STRUCTURE AND OUTLINE OF THE THESIS

The thesis is organised as follows. Chapter 2 offers some remarks on the typology of grammatical gender and gives a general introduction to the gender system of Tocharian. In addition, it deals with the analysis of the Tocharian *genus alternans* from the perspective of general linguistics and offers an introduction to the gender assignment system of Tocharian. Chapter 3 addresses issues related to the inflection of nouns, clarifying for each of the selected inflectional classes the origin and the evolution. Chapter 4 provides an overview of the pronominal and adjectival system of Tocharian and treats the development of the three inherited genders across the inflection of both pronouns and adjectives. Chapter 5 recapitulates the findings and provides conclusive remarks. The main body of the thesis consists of the second, the third, and the fourth chapters. As each of them needs independent clarifications on the aim and the structure, such matters are discussed in the introductions of the relative chapters.

Since the category of gender requires the investigation of a wide variety of inflectional and derivational patterns as well as its intersection with the other nominal categories of case and number, this may perhaps be seen as a study of Tocharian nominal morphology as a whole.