Institute of Oriental Studies of the Russian Academy of Sciences

As a manuscript

Anna V. Kuritsyna

**SYNTAX OF TOCHARIAN LANGUAGES
IN COMPARATIVE ASPECT**

Dissertation summary

for the purpose of obtaining the academic degree

Doctor of Philosophy in Philology and Linguistics

Academic Supervisor:

Ilya B. Itkin,

Doctor of Philosophy in

Philology and Linguistics

Moscow

2023

This dissertation examines a set of previously unstudied (or poorly studied) phenomena of Tocharian syntax and compares their realization in two closely related languages – Tocharian A and Tocharian B; it also intends to discuss these phenomena from a typological and areal perspective.

Both Tocharian languages show features in nominal and verbal domain[1] which may have been induced by contact with some agglutinative language or languages, but the exact source is unknown. There are quite a few works exploring possible linguistic contacts of the Tocharians before and during their settling in the oases of the Tarim Basin. Just some of them are: Hansen 1940, Schwarz 1974, Schmidt 1985, Pinault 2002, Tremblay 2005, Blažek & Schwarz 2017 and Peyrot 2022: 84-87 on Tocharian-Iranian contacts or Winter 1963, Pinault 2001, Lubotsky & Starostin 2003, Kasai 2013 and Mair 2013 on Tocharian-Old Turkic contacts. Also, two important research projects at the Leiden University should be mentioned: "Tracking the Tocharians from Europe to China: a linguistic reconstruction" (2016-2021) dedicated to contacts between Tocharian and Old and Middle Iranian languages, and "The Tocharian Trek. A linguistic reconstruction of the migration of the Tocharians from Europe to China" (2018-2023) which examines Tocharian-Uralic and Tocharian-Turkic contacts. However, most of the listed studies concern lexical borrowings, while structural parallels remain understudied. This is partially caused by the relative youth of the scientific area itself (first manuscripts from the Tarim Basin were found in the end of the 19th – the beginning of the 20th century) as well as by the fragmentary character of the linguistic corpora.

To the best of my knowledge, separate structural similarities in the languages of the Tocharian linguistic area have been addressed only in few works. For instance, syntactic and semantic parallelism between Tocharian adjectives in A -ṣi / B -ṣṣe and the Old Uighur adjectives in -lXg is partially discussed in [Erdal 1991: 141-142]. G.-J. Pinault notices in [2002: 244] that Tocharian shares "with Bactrian and other Middle Iranian languages morphological and syntactic characteristics', illustrating this statement with the special marking of animate direct object and with the reanalysis of the optative as a narrative past tense. In [Peyrot 2017: 333-339] one can find quite a detailed comparison of how the Tocharian middle participle and the Old Uighur 'vowel

---

[1] Some of the Tocharian innovations in **nominal morphology and syntax**: «two-leveled» case system with primary and secondary cases, the latter added to the Accusative form; group inflection; non-cumulative expression of grammatical meaning by secondary cases, some Plural and Dual markers; generally loose differentiation between derivation and flection (derivational suffixes can link to flectional ones, e.g. to Plural markers etc.). Some innovations in **verbal system** can also indicate a substrate influence: a special role of the category of causativity; loss of flection by the present participle middle and its predominantly converbial use; numerous non-finite verbal forms and contextual orientation of some of them (gerund, past participle, present participle middle). More on this see in the Introduction of the dissertation.

converb' are used; it is proposed that the productivity of the Tocharian *m*-participle in converbial function may have emerged under Turkic influence. Also, the fact that the Tocharian absolutive is used for translating the Sanskrit forms in -*ya*, -*tya*, -*tvā* and can correspond to the Old Uighur converb in -*Xp* is discussed in [Pinault 2015: 95-97].

The most systemic attempt of a typological comparison between Tocharian and several agglutinative languages of Eurasia has been made in [Peyrot 2019] published as part of the mentioned above project "The Tocharian Trek". In this work the author proposes that similarities on several language levels speak in favor of a Uralic (more exactly of "an early form of Samoyedic" [ibid.: 73]) source for the Tocharian typological shift. These are similarities in both phonetic and case systems, group inflection and a broad use of converbs. However, the examined parallels appear the closest for the phonetic level of Tocharian and Samoyedic only, while similar syntactic features can be found in more than one language among those chosen for the comparison.

Some general remarks on a possible influence on Tocharian grammar from Altaic or Uralic languages can also be found in [Schmidt 1969; Zimmer 1982/83: 277-278, 285; Pinault 2007; Mallory 2015]. Structural innovations usually addressed in these works predominantly include group inflection, monofunctionality of secondary case markers, prevalence of postpositions above prepositions, but mostly without a detailed analysis of these phenomena.

Thus, typological peculiarities of Tocharian syntax and morphosyntax appear important for the discussion on the history of linguistic contacts of the Tocharians, but at least some of these peculiarities remain poorly described from both inner Tocharian and areal perspective.

This dissertation aims to research and to describe a set of previously understudied topics on nominal as well as on verbal syntax of Tocharian concerning group marking of coordinated nouns on the one hand and functioning of Tocharian participles and converbs on the other hand, specifically:

- marking of coordinated nouns by means of prepositions and adjectival suffixes;
- use of "middle" and "active" present participles and some agent nouns in coinciding syntactic positions;
- syntax of absolutive constructions and of the sentences including them.

The **objective of the dissertation** is to reveal possible factors which may determine:

- repeated vs. single use of prepositions governing coordinated nouns;
- single vs. repeated suffixation in adjectivization of coordinated nouns;
- the choice between middle and active present participles on the one hand and between active present participles and agent nouns on the other hand when the compared forms occur in similar syntactic contexts;
- the word order of sentences which contain absolutive constructions.

For each of the topics I am comparing data of Tocharian A and Tocharian B because the realization of syntactical characteristics may differ in them despite the close genetic relationship. The results are also discussed in areal and typological context. The areal comparison is limited to Old Uighur and Middle Iranian data only. Besides technical restrictions of my study, there are two main reasons behind this limitation: 1) speakers of these languages were direct neighbours of the Tocharians in the oases of the Tarim Basin, 2) all these languages share agglutinative features, even if the latter are expressed to quite a different extend[2].

To achieve the goals formulated above, it is planned to accomplish the following **tasks**:

● To examine the use of prepositions with coordinated nouns; to explain cases of prepositional reiteration against single preposition; to examine differences in conjunctional use between the comitative preposition and related words: A *śla*, *ślak*, B *śle*, *śale*, *ślek*; to search for possible cases of preposition repetition in other languages (Chapter I);

● To examine adjectivization of coordinated nouns by means of the Tocharian suffixes A *-ṣi*, B *-ṣṣe*, *-tstse* and *-ññe*; to explain differences in productivity of this derivational model in Tocharian B; to search for typological and areal parallels for this phenomenon (Chapter II);

● To examine the use of the Tocharian "active" and "middle" present participles in coinciding syntactic positions (attributive and predicative) and to reveal a possible distribution between both forms; to compare with the use of the "middle" participle in other ancient Indo-European and Eastern Middle Iranian languages; to typologize Tocharian system of present participles; (Chapter III);

● To examine the use of the Tocharian "active" present participle and some Tocharian B agent nouns in similar contexts and to find out possible factors influencing the choice between them (Chapter IV);

● To examine the use of absolutive constructions and to describe structural peculiarities of sentences containing such constructions; to compare the Tocharian absolutive constructions with functional parallels in areally close languages (Chapter V).

The research corpus includes all relevant manuscripts of Tocharian A and Tocharian B, available on the electronic sources CEToM ("A Comprehensive Edition of Tocharian Manuscripts"), TITUS ("Thesaurus Indogermanischer Text- und Sprachmaterialien") and IDP ("International Dunhuang Project"). Data from other languages of the Tocharian linguistic area

---

[2] For instance, in Tocharian group inflexion affects mostly secondary case markers, in Eastern Iranian languages it affects secondary case and plural markers [Skjaervo 2007: 65; Yoshida 2009: 312], and in Turkic languages including Old Uighur not only case and plural markers but also possessive suffix and ezafet in nominal domain and suffixes of causative and passive in verbal domain can undergo group inflexion [Lewis 1967; Gabain 1941: 158-160], see more detailed the Introduction to the Chapter 2 of the dissertation. Moreover, at least in Sogdian some verbal suffixes, even if not attested in group inflexion, are used monofunctionally like in Turkic [Vinogradova 1999: 67, 72]

has been taken from grammatical descriptions and dictionaries; for Old Uighur also the texts *Maitrisimitnombitig*[3] and *Daśakarmapathāvadānamālā*[4] have been used as a source.

The study uses philological **methods** including description, comparison, distributive method and contextual analysis.

The scientific relevance is based on an incomplete studiedness of several aspects of Tocharian grammar on the one hand, and on the quick development of this scientific area in the last decades on the other: the online projects CEToM ("A Comprehensive Edition of Tocharian Manuscripts"), TITUS ("Thesaurus Indogermanischer Text- und Sprachmaterialien") or IDP ("International Dunhuang Project") make more Tocharian manuscripts accessible which offers new possibilities for research.

The scientific novelty of the chosen theme lies in the fact that the research topics have not been addressed earlier and remain undescribed in the previous works.

The results achieved in this dissertation complement our knowledge about several aspects of Tocharian nominal and verbal syntax; they also may be important for further studies in areal linguistics of East Turkestan and in typology of syntactic borrowings. Therein I see the **theoretical significance** of the study. Its **practical significance** lies in the fact that the results allow us to specify existing descriptive grammars of Tocharian, they can also be used in university courses and may be useful for interpretation and translation of some already known manuscripts as well as of potential new findings in Tocharian languages.

The main propositions of the study:

1. If governing coordinated nouns, Tocharian prepositions must repeat before each conjunct. This rule is obligatory in both Tocharian languages and can be broken only due to metrical reasons (if the number of syllables in the metrical line is insufficient for a repeated preposition).

2. The rule of preposition repetition is opposed to group inflexion and must be caused by the inability of the preposition as the head of a right-branching construction to govern several coordinated noun phrases in Tocharian as a predominantly left-branching language.

3. Tocharian adjectival suffixes A *-ṣi*, *-eṃ*, B *-ṣṣe*, *-tstse* and *-ññe* can take scope over coordinated noun phrases. Analogically to group inflexion of case suffixes, the use of a single derivational suffix with conjuncts can be treated as group derivation and as a type of suspended affixation. A similar adjectivization model exists in Turkic languages including Old Uighur.

---

[3] Tekin 1980 for the Siŋgim version, Geng & Klimkeit 1988, Geng, Klimkeit & Laut 1988, 1998 for the Hami version.

[4] In Wilkens 2016.

4. Although the Tocharian middle present participle (PPM) mostly functions as a converb of simultaneity, in attributive use it is complementary distributed with the active present participle (PPA): PPM relativizes the subject of an intransitive verb and the object of a transitive verb (S = P) while PPA relativizes the subject of a transitive verb (A). The ergative model of the Tocharian (attributive) PPM is unique for a present participle from Indo-European but also from areal perspective. In Eastern Middle Iranian languages as well as in Old Turkic all present participles / imperfective participles are active (A = S). However, the ergative pattern is widespread in formation of deverbal nouns and adjectives of Old Turkic and in past participles of Middle Iranian. A remarkable typological parallel is the participle in *-gæ* in modern Ossetic.

5. Tocharian B active present participles in *-ñca* (PPA) and agent nouns in *-auca* have equal syntactic compatibility and turn out to be complementary distributed depending on the present stem class of the verb: verbs with nasal present classes (6 and 7) obtain a form in *-auca* instead of the usual active participle in *-ñca*. Thus, the form in *-auca* should be treated as one more type of present participle active (PPA II).

6. Tocharian absolutive constructions have two basic positions in the sentence: preposition to subject and predicate (Abs-S(O)V) if absolutive construction contains a sentential pronoun, and interposition (S-Abs-(O)V) in all other cases. This rule is strong in Tocharian A but has some deviations in Tocharian B. Moreover, absolutive can have a subject different from that of the main clause; this is a unique property of the absolutive compared to other Tocharian predicates. A similar syntactic behavior is typical for Old Uighur secondary converbs in *-mIš+tA* and *-dOk+dA*.

Structure and volume of the dissertation

The study consists of an introduction, five chapters with subdivisions and conclusions in each chapter, a general conclusion, a list of abbreviations and a list of references. The dissertation volume comprises 231 pages, the list of references includes 185 works. All used examples have a through numeration within each chapter and are glossed in a simplified way.

CONTENT OF THE DISSERTATION

The introduction gives a brief information on Tocharian languages and previous works concerning their areal and typological study; it also contains formal characteristics of the dissertation such as objectives and tasks of the research, its relevance, theoretical and practical significance, research corpus and methodology, novelty of the research and main propositions submitted for the dissertation defense.

Chapter I examines the use of the Tocharian prepositions A *śla*, B *śle* "with" and A *sne*, B *snai* "without" with coordinated noun phrases in prepositional (1.1.) and conjunctional function (1.2.). In section 1.1. (§ 1.1.2.) all found examples of prepositions with coordinated nouns are discussed. The preposition can either be single before the first of the coordinated elements or repeat before each of them, like in the examples 1a-b for TA and 2a-b for TB below:

1 a) *śla ṣulas ṣṭāmäntu*
        "with mountains (and) trees" (A 320 a7).

   b) *sne rse sne yäslurñe*
        "without hatred (and) without enmity" (YQ N.5 a8).

2 a) *snai laiwo ṣpane*
        "without rest (and) sleep" (B 31 a5).

   b) *[ono]lmi snai yräm snai keś snai ṣa(ṃ)ṣäl*
        "[bein]gs without measure, without number, without count" (B 169 a4).

As shown in §1.1.2., iteration of the preposition does not depend on whether the coordinated nouns build a collocation or not: among the found examples there are undoubtful hendiadyses as well as quite rarely co-occurring pairs of words; also the former (hendiadys) can be used both with single and with iterated preposition.

It could be expected that the cases with preposition repetition are a calque from Sanskrit but the Tocharian data disproves this. Among all the examples only three have a known Sanskrit parallel (examples 3-5 below). One of them (number 3) shows some kind of similarity: two Tocharian conjuncts with repeated preposition "without" render two Sanskrit compounds with the negation prefix *a(n)*-. In the two other examples Tocharian prepositional phrases present a more free and redundant translation of single Sanskrit adjectives *arupi* "formless" and *atandritam* "tireless" respectively, cf.:

3   *asaṅgam*   *anapagraham •*   *sn[ai]*   *treṅ[käl]*   *snai*   *krämpālyñetse •*
     unattached   undistracted   without   attachment   without   distraction

     "Without attachment, without distraction" (B 251 b2).

4   *ākāśam*   *arupi •*   *ākāś*   *sne*   *arämpātum*   *sne*   *wrātal •*
     air   formless   air   without shape   without   form

7

"Air without form (and) without shape" (A 387 a3).

5   *aiśaumyi*   *ceu*   *pällāntär*   *krento*   *āstreṃ śaul*

aiśaumyi	*ceu*	*pällāntär*	*krento*	*āstreṃ*	*śaul*
sage:PL	DEM.OBL	praise:PR.3PL	good:OBL	pure:OBL	life

śayeñcai	*wnolme:*	*snai*	*laiwo*	*ṣpane*
live:PPA.OBL	being	without	tiredness	sleep

"The sages praise him as a being living a good, pure life, without lassitude or sleep." (B 31 a5; translation: CEToM).

Cf. the Sanskrit parallel (Uv 13.12):

tam	*praśaṃsanti*	*śuddhājīvam*	*atandritam* ‖[5]
DEM.ACC	praise:PR.3PL	purely_living:ACC	tireless:ACC

At the same time preposition repetition correlates with the text type: repeated preposition (§1.1.2.2.) is possible in both metrical and prosaic texts while single preposition (§1.1.2.1.) occurs only in verse. Since the Tocharian verse is syllabic so that every meter has a fixed number of syllables for each pāda and for each subdivision inside the pāda, it can be assumed that single use of the preposition is due to lack of an "additional" syllable for one more preposition in a specific metrical scheme. This appears to be proved with the next example:

6

ṣäptäñcäṃ	*koṃ*	*śla*	*klop*	*wraṣäl*	|	*ṣpät*	*pā(k*	*ats*
7:ORD.OBL	day	with	pain	suffering		7:CARD	part	EMPH

la)p	*wākña-ci*
head	split:KO.1SG-PN.2SG

"On the seventh day, with sorrow and pain, I will split your head into seven parts." (A 215 b1; translation: CEToM).

This line has the metrical scheme 8¦7¦7 with the first caesura (marked as "/") directly following the prepositional phrase; here a repeated preposition would require an extra ninth syllable.

The influence of metrics is more evident in the example 7 below where the second preposition can be kept due to its devocalization before the vocal in the beginning of the following word (sandhi is marked with "=") so that no excess syllable is required:

7

puttiśparṣās	*sambhāräntu*	*sne*	*cärk*
Buddha_rank:ADJ.OBL.PL	collection:PL	without	stop

sñ=	*āñu*	*kropnämāṃ*
without	rest	gather:PPM

"Without pause and without end collecting preparations for the rank of a Buddha" (A 405 a1; translation: CEToM). Here the Meter must be 4×25 (5¦5¦4¦4¦4¦3) or 20/22/10/15 (5¦5¦5¦5 / 4¦4¦4¦3¦4¦3 / 5¦5 / 4¦4¦4¦3).

Some of the examples found can hardly be attributed to verse or prose because of highly fragmentary character of the manuscripts but well-preserved examples show following proportion of single vs repeated preposition:

---

[5] [Bernhard 1965: 204].

Table 1

	Preposition repetition		Single preposition	
	TA	TB	TA	TB
Verse	6	10	8	7
Prose	11	6	1	0
Total	33		16	

Thus, it is shown in §1.1. that preposition repetition with coordinated nouns is compulsory in both Tocharian languages and the lack of repetition can be caused only by metrical requirements of the respective verse line. It is proposed that the inability of prepositions to govern coordinated noun phrases is due to the branching direction: Tocharian as a predominantly left-branching language may have had a constraint for the preposition as the head of a right-branching construction to take scope over several coordinated nouns.

As discussed in §1.1.1., preposition repetition is described for quite a few other languages including some geographically close Eastern Middle Iranian languages (Sogdian and Bactrian, cf. [Sims-Williams 1973: 96; Steblin-Kamenskij 1981: 344; Yoshida 2009: 314]); however, in the latter this phenomenon seems to be limited to appositive noun phrases only. Moreover, Tocharian and Eastern Middle Iranian languages may share the use of circumpositions (preposition + postposition / secondary case suffix).

In section 1.2. I am also analyzing the use of the comitative preposition A *śla*, B *śle* in conjunctional function beside the cognate conjunctions A *ślak*, B *śale*, *ślek*. According to Thomas 1979 (also Hacktsein 1997, Adams 2013, Kim 2014, CEToM) A *śla*, B *śle* is treated as equally combining prepositional and conjunctional function, and TB *śale* is claimed to be an accented variant of *śle*. However, these statements are based only on separate examples. A comprehensive analysis of all relevant contexts (§1.2.1. for TB and §1.2.2. for TA) shows that all the forms can be attested in differing functions. Namely, *śale* cannot be found in prepositional use; in conjunctional use it is only attested in coordination of clauses while *śle* conjoins only nouns; *ślek* shows a mixed use; moreover, there is insufficient data for argclaiming conjunctional use of TA *śla*. Thus, it is concluded that the forms in question should be treated separately from lexicographical perspective despite their common origin.

In **Chapter 2** I am examining the ability of some Tocharian adjectival suffixes to be morphosyntactic heads of coordinated nouns. Tocharian suffixes A *-um*, *-ṣi*, *-eṃ*, B *-ṣṣe*, *-tstse* and *-ññe* can adjectivize subordinated phrases which has already been described in detail in [Zimmer 1982/83; Adams 2009; Hajnal 2004; Burlak & Itkin 2013: 438-439]. In this dissertation it is shown

that most of these suffixes can also link coordinated phrases. This is mainly true for A -*ṣi*, B -*ṣṣe* like in the examples 8 (TA) and 9 (TB) below.

8	*kus*	*ne*	*ārkiśoṣṣaṃ*	*k(ā)wältune*	*ciñcroneṣi*	*wākä(m)*	*ṣ[e](ṣ)*
	who	EMPH	world:LOC	beauty	tenderness:ADJ	paragon	be:IPF.3SG

"But who was the paragon of beauty and tenderness in the world" (A 175+178+THT 2968.g a5[6]).

9	*atiyai*	*pisäl*	*melte*	*oraṣṣe*	*puwar*
	grass:OBL	chaff	dung	wood:ADJ	fire

"A fire of grass, chaff, dung and wood" (B 194 b1).

Single examples with the suffixes A -*eṃ*, B -*tstse* and -*ññe* can also be found. For instance, the following Tocharian A text shows a case of -*eṃ* (plural -*eñi*) governing two nouns in coordination:

10	*(sne*	*yä)rm*	*sne*	*mem*	*kinnareñ*	*gandharveñi*	*lāñś*
	without	measure	without	number	kiṃnara:NOM.PL	gandharva:ADJ.NOM.PL	king:NOM.PL

"The innumerable and countless kings of the Kinnaras and Gandharvas (saw the venerable Metrak)" (YQ II.9 b3; translation: CEToM).

Tocharian A examples including this one (10) are discussed in detail in §2.3. of the dissertation.

Tocharian B examples with the suffixes -*ññe* (11) and -*tstse* (12):

11	*ṣaḍvarginta*	*yasa*	*ñkante*	*wrākaññeṃ*	*wmera*	*makci*	*priyeṃ*
	Ṣaḍvargika:PL	gold	silver	pearl:ADJ.OBL.PL	jewel:PL	self:NOM.PL	wear:IPF.3PL

"Ṣaḍvargikas[7] were by themselves wearing jewels of gold, silver (and) pearl" (PK AS 18A a2).

12	*snai*	*trenkäl*	*snai*	*krämpālyñetse*
	without	attachment	without	distraction:ADJ.NOM

"Without attachment (and) without distraction" (B 251 b2).

These and some more Tocharian B examples are discussed in §2.4.

Although paired nouns giving adjectives of this type are hardly distinguishable from compounds, a few examples may speak in favor of the syntactic independence of adjectivized elements (§2.2.). One of the most demonstrative examples of this kind (13) presents an adjective in -*ṣṣe* containing six coordinated noun phrases (names of medical plants), two of which are modified noun phrases (*wi praha(ti)nta* "two Solanum" and *klyotaiṣṣana witsaka-* "roots of Tribulus"):

13	*hirant* •	*pilamāti* •	*wi*	*praha(ti)nta*	• *mā(tu)l(uṅk*	*pāṣā)ṇa(bhit)*•
	Ricinus	quince	2	solanum:PL	citron	coleus

klyotaiṣṣana	*witsakaṣṣe*	*kaṣāy*
tribulus:ADJ.PL	root:OBL.PL.ADJ	decoction

---

[6] Joined according to [Itkin 2019: 21].

[7] "Member of a particular group of (possibly heretical) Buddhist monks" [Adams 2013: 710].

"A decoction consisting of ricinus, of the heart of the fruit of Aegle marmelos, of two Solanum indicum, of citron, of (Coleus) aromaticus [and] roots of Tribulus terrestris..." (PK AS 2A a3-5; translation: CEToM).

Such a use of adjectival suffixes can be compared with group inflexion / phrase marking by means of Tocharian secondary case markers which can link both subordinated and coordinated groups, cf. Tocharian B pairs of words "hunger (and) thirst" (14) and "doubt (and) fear" (15) with the adjectival suffix *-ṣṣe* and the perlative suffix *-sa*:

14	*kest*	*yokaiṣṣe*	*läkle*[8]	vs.	*kest*	*yokaisa*[9]
	hunger	thirst:OBL.ADJ	suffering		hunger	thirst:OBL.PERL
	"The suffering of hunger (and) thurst"				"because of hunger (and) thurst"	

15	*sklok*	*pr(o)sk(aiṣṣe*	*or)k(a)mñe*[10]	vs.	*sklok*	*proskaisa*[11]
	doubt	fear:OBL.ADJ	darkness		doubt	fear:OBL.PERL
	"The darkness of doubt and fear"				"because of doubt (and) fear".	

It is concluded that the set of suffixes which can take scope over coordinated noun phrases in Tocharian is wider than generally considered and includes not only inflectional but also some derivational suffixes. Therefore, it is proposed to treat this Tocharian phenomenon under the terms *suspended affixation*[12] or *group derivation* (analogically to group inflection). However, it must be pointed out that the suffix A *-ṣi,* B *-ṣṣe* has an exclusive productivity in the described derivational model while there are only single examples with other suffixes: in Tocharian A I could find seven adjectives in *-ṣi* against only one adjective in *-eṃ*; in Tocharian B near 20 examples with the suffix *-ṣṣe* are attested while *-tstse* and *-ññē* occur only once each. Possible reasons of such specific productivity of the suffix A *-ṣi,* B *-ṣṣe* in this adjectivization model are discussed in §2.4.

Although the typology of suspended affixation remains relatively understudied (previous works on it are discussed in §2.5.), some parallels can be found in the modern Turkic [Akkuş 2016] but also in the geographically closer Old Uighur. In the latter such parallels are comprised mainly of the adjectives in *-lXg*. An exceeding functional similarity of the Old Uighur suffix *-lXg* and the Tocharian suffix A *-ṣi*, B *-ṣṣe*, including the ability to adjectivize subordinated constructions, was earlier noticed in [Erdal 1991: 139-140]. However, there is no detailed discussion on whether *-lXg* can also govern coordinated constructions, except for the mention that *-lXg* can be added to "binomes" (ibid.: 147). Nevertheless, some additional examples can be found which prove parallelism between *-lXg* and its Tocharian counterpart in coordinative sequences as well, like in the following example where *-lXg* refers to four conjuncts:

---

[8] B 284 a2.

[9] B 286 b3.

[10] PK AS 17K a1.

[11] B 409 a1.

[12] The term was first proposed in [Lewis 1967] for phrase marking of coordinated constructions.

16	*ačmak*	*[suvsa]mak*	*busuš*	*kadgulug*	*otın*
	hunger	thirst	sorrow	solicitude:ADJ	fire:INSTR

"Wir [brennen(?)] im Feuer des Hungerns, Dürstens, der Sorge und des Leids"
(MaitrHami XX 14 verso 10-12; Geng, Klimkeit & Laut 1998: 82).

Thus, the ability to undergo suspended affixation may be an areally induced characteristic for the Tocharian adjectival suffix A -*ṣi*, B -*ṣṣe* and the Old Uighur suffix -*lXg*.

The **Chapter 3** is dedicated to the problem of potential functional differences between the Tocharian present participles in A -*nt*, B -*ñca* and A -*māṃ*, B -*mane*. These forms are traditionally denominated as "active" (further PPA) and "middle" (further PPM) because they continue Indo-European participles in *-*nt*- and *-*mh₁no* respectively, but in Tocharian languages their use does not depend on the voice of the finite verb, and both participles have mostly different syntactic functions: forms in A -*nt*, B -*ñca* are used substantively while forms in A -*māṃ*, B -*mane* represent a converb of simultaneity (previous works on the Tocharian present participles are surveyed in §3.1.). At the same time, both forms can occur in attributive and predicative functions. This dissertation is the first research on possible differences between PPA and PPM in such overlapping syntactic positions. In §3.2. all found Tocharian A and Tocharian B contexts with attributive PPA (§3.2.1.) and PPM (§3.2.2.) are analyzed. It is revealed that if used as adjectives, PPA and PPM do have differences in voice semantics based on transitivity / intransitivity of the deriving verb, namely: transitive verbs build active participles in A -*nt*, B -*ñca* and passive forms in A -*māṃ*, B -*mane*, while intransitive verbs build PPM with active meaning. This can be illustrated with the examples 17 (transitive PPA and intransitive PPM in one and the same clause) and 18a-18b (transitive active PPA vs. transitive passive PPM) below.

17	*yaläṃ*	*wram*	*ypant*	*wrasom*	*nu*	*pälkäṣ*
	do:GER	thing	do:PPA	being	EMPH	shine:PR.3SG
	mäṃt ne	*sälpmāṃ*	*por :*			
	like	glow:PPM	fire			

"A being which does a thing that must be done is shining like a glowing fire" (A 3 b4).

18a	*lāñci*	*waṣt*	*pāṣäntās*	*śāwes*	*empeles*	*(n)ā(gā)s*
	royal	house	guard:PPA.OBL.PL	big:OBL.PL	horrible:OBL.PL	snake:OBL.PL

"… big and horrible Nāgas who are guarding the royal palace" (A 1 b4-5).

18b	*pāsmāṃ*	*niṣpal*	*lo*	*näkṣäl*
	guard:PPM	property	away	perish:GER

"The protected property will be lost away… (A 2 b3-b4)".

This distribution appears quite strong and has only single deviating examples which are however fragmentary and syntactically ambiguous.

In §3.3. the use of PPA and PPM in predicative function is additionally studied. It is concluded that predicative participles don't show the same strong distribution which can be observed in attributive cases: PPM derived from transitive verbs can have active meaning.

The distribution of voice semantics between Tocharian PPA and PPM in attributive position can be presented in the following table:

Table 2

Present participle in attributive function	TA		TB	
	PPA	PPM	PPA	PPM
Total cases	21	31	20	13
● vt	20	6 (passive: 4)	19	3 (passive: 3)
● vi	1	25	1	10

In this chapter participles derived from the verb *mäsk-* "be, become" are treated in a separated section (§3.4.) because at least in Tocharian B both forms, PPA *mäskeñca* and PPM *mäskemane*, seem to have different lexical meaning: *mäskeñca* occurs only in contexts with a depending locative and has the meaning "staying in X" while the form *mäskemane* is copular / existential. Based on this fact and some deviating cases it is proposed that not only verbal valency but also verbal agentivity may play a role in the participle choice in case of an overlapping syntactic position.

In §3.5. the Tocharian system of present participles is discussed from Indo-European and areal perspective. In ancient Indo-European languages which preserve the opposition of PPA and PPM, both participles are distributed between the verbs with active and middle paradigm and syntactically oriented towards the subject of a transitive and an intransitive verb (A = S) which with some reservations is also mainly true for PPM, while the object of a transitive verb (P) is relativized separately (like in Balto-Slavic languages or in Sanskrit). In a general case (i.e., in non-attributive use) Tocharian present participles seem to have the same orientation: PPA is active (A = S), PPM is also active or probably contextually oriented (A = S = P)[13]. But in shared attributive function both forms comprise an agentive-absolutive[14] correlation model: the attributive PPA relativizes the subject of a transitive verb (A) only while the attributive PPM qualifies the object of a transitive and the subject of an intransitive (and usually non-agentive) verb (P = Sp). In Middle Eastern Iranian languages which belong to the Tocharian linguistic area both forms are active participles (A = S). With rare exceptions, the same is true for Old Uighur imperfective participles as well. Nevertheless, absolutivity (ergativity) is widely present in both Middle Iranian and Old Turkic, although it affects mainly forms with resultative meaning: Middle Iranian past participles

---

[13] In fact, examples of converbial passive PPM are very rare: I could find only four cases in TA and no cases in TB.

[14] In terms of the typology of participles used in [Shagal 2019: 75-78].

(giving there rise to ergative alignment in the past)[15] and most Old Uighur deverbal nouns[16]. Thus, the absolute pattern (S = P) may appear unique for present participles, but not generally for deverbal formations in the Tocharian linguistic area. Moreover, an interesting typological parallel is the absolutive present participle in -*gæ* in modern Ossetic [Abayev 1970: 611-612].

In non-attributive use the Tocharian active present participle is often treated as an agent noun [Dietz 1981: 3-4, 9-10, 144; Pinault 2008: 616-617; Peyrot 2017: 328-329]. At the same time, in Tocharian B there are up to eight "true" agent nouns (with the suffixes -*nta*, -*ntsa*, -*uki*, -*uca*, -*auca*, -*i*, -*a* and -*tse*). **Chapter 4** examines whether PPA and agent nouns (NA) in TB show any functional differences.

In the section §4.1. the situation in Tocharian A is discussed where there is only a small number of agent nouns (with the only suffix -*nt*) and the latter differ from the PPA only in vowel length of the deriving stem. In section §4.2. the more complex situation of Tocharian B is researched where participles and some of the agent nouns (in order to avoid confusion I am using traditional terms for the forms in question) are attested in similar contexts. For instance, in the following two examples the PPA (*tanmaṣṣeñca* and *keṣṣeñca*) and the NA (forms in -*uki* and in -*auca* ~ -*oca*) co-occur within one and the same manuscript in a similar syntactic environment:

19 Text B 197:

a) | *tarya* | *vedantats* | *putkau* | ***tanmaṣuki*** | *viṣai* |
|---|---|---|---|---|
| 3:CARD | sensation:PL.GEN | distinction | engender:NA | sphere_of_senses |

mā	*nesäṃ*
NEG	be:PR.3SG

"The external object (*viṣaya*) does not engender the distinction of three sensations (*vedanā*)" (lines b3-b4; translation: Schaefer 1997: 165).

b) | *sak* | ***tanmaṣṣeñca*** | *mäsketrä* |
|---|---|---|
| happiness | engender:PPA | become:PR.3SG |

"(*kakse*-object in the body) engenders pleasure" (lines b4-b5).

20 Text B 295:

a) | *po* | *lläklenta* | ***keṣṣeñcai*** |
|---|---|---|
| all | suffering:PL | extinguish:PPA.OBL |

"Destroying all sufferings…" (line a9).

b) | *pontäṃntso* | *akalkänta* | ***kärstoca*** |
|---|---|---|
| all:GEN.PL | wish:PL | cut_off:NA |

"(The monster of death) cutting off the wishes of everybody" (lines b2-3).

---

[15] [Haig 2017; Livshits 1999: 45; Vinogradova 1999: 71, 92].

[16] [Erdal 2004: 151-152].

Further the PPA and the most frequent Tocharian B NA (those in *-nta*, *-ntsa*, *-uki*, *-uca*, *-auca*) are compared with respect to their morphological characteristics (§4.2.1.) and syntactic compatibility (§4.2.2.). As a result, it can be revealed that forms in *-auca* in fact comprise a PPA II which is complementary distributed with the PPA I (participle in *-ñca*) depending on the present class of the deriving verb. Namely, verbs with nasal *Auslaut* of the present stem (classes VI and VII) build their present participle from the subjunctive stem (such verbs almost always have subjunctive class V) and by means of the suffix *-auca* while all other verbs build their present participle from the present stem and by means of the suffix *-ñca*. Cf. the choice between the forms *keṣṣeñcai* vs. *kärstauca ~ kärstoca* in the examples 18a and 18b given above:

- PPA I *keṣṣeñca < käs- / kes-* "extinguish", present class 2, 3sg pr *keṣäṃ*
- PPA II *kärstoca < kärst-* "cut off", present class 6, 3sg pr *kars**n**aṃ*

Only three verbs show a deviation from this distribution: they have both participle forms (in *-ñca* and in *-auca*). However, two of them are suppletive and have subjunctive class V:

- *āk- ~ wāy-* "lead" (Pr II *āk-*, Ko V *wāy-*, Pt Ib *wāy-*) > PPA *āśeñca* & NA *wayauca*
- *läk- ~ pälk-* "see" (Pr IX *läk-*, Ko V *läk-*, *pälk-*, Pt Ib *läk-*, *pälk-*) > PPA *lkāṣṣeñca* & NA *pälkauca*

The third of these verbs, *kälp-*, is not suppletive and has subjunctive class 6 but the latter must be secondary [Malzahn 2010: 408]. It is moreover probably that the present class 9 is an innovation as well (since in TA *kälp-* has the stems Pr VI, Ko V, Pt I):

- *kälp-* "receive" (Pr IX, Ko VI, Pt Ia) > PPA *kälwāṣṣeñca** & NA *kälpauca*

Thus, the form *kälwāṣṣeñca* must have occurred after the substitution Pr VI > Pr IX so that this verb is no real exception to the observed distribution between PPA and the form in *-auca*.

Further it is revealed that some of the agent nouns in *-uca* can be used attributively like the NA in *-auca* but it remains unclear whether both must be treated as one and the same NA type or not because the group of nouns in *-uca* is too small and at the same time heterogeneous with respect to the syntactic characteristics of its constituent forms.

Moreover, a study of the contexts containing nouns in *-uki* in §4.2.2. allows to assume that these agent nouns may also be a special type of PPA used in sentences with condition or negation. However, the number of relevant contexts is too small for a sure conclusion.

If forms in *-auca* are PPA II, they must undergo the distribution rule of attributive PPA and PPM proposed in Chapter 3, i.e., it is expected that in adjectival use forms in *-auca* are built only to transitive verbs. Such a review is carried out in §4.2.3. and shows that the interpretation of NA in *-auca* as PPA II does not contradict to the conclusions made in Chapter 3.

Section 4.3. summarizes the results of the research conducted in Chapter 4. It is proposed that the connection of the forms in *-auca* with the nasal present classes VI and VII on the one hand and

with the subjunctive class V on the other has a single origin: the absolute majority of the verbs with present stems belonging to the classes VI and VII have subjunctive class V so that the use of the suffix -*auca* instead of -*ñca* in verbs with nasal present could have been reanalyzed as the rule to use -*auca* instead of -*ñca* in verbs with the subjunctive class V. It is concluded that such a rule exists due to the intention to avoid a chain of several nasals between the stem and the suffix. The dissimilation may also explain those forms in -*auca*/-*uca* which don't belong to present classes VI-VII but in fact have a nasal stem *Auslaut* such as *pännauca* instead of **pänneñca* from *pänn-* "tear" (Pr II) and the prevalence of the form *ynūca* above the form *yneñca* from the stem *yn-* of the verb *y-* "go" (Pr I).

In Section 4.4. a full list of Tocharian B PPA and NA forms discussed in this chapter is given.

Chapter 5 examines the word order in sentences containing absolute construction (AC)[17]. The Tocharian absolute construction functions as a gerund of antecedence [Thomas 1960; Pinault 2008: 612-613; Burlak & Itkin 2013: 429, 473]. Although AC are highly frequent in both Tocharian languages, the syntax of sentences containing this construction has not been studied separately. It is believed (e.g., in [Pinault 2015: 95, 97; Peyrot 2017: 334]) that the AC usually takes the initial position in the sentence. Indeed, initial position is typologically most common for a converb of antecedence [Tikkannen 2001: 1118]. However, according to my preliminary estimate, deviations from this word order may account for up to half of all cases in Tocharian.

In this research, 94 sentences with AC in Tocharian A and 63 sentences with AC in Tocharian B have been gathered; they are analyzed in §5.1. and §5.2. respectively. As a result, it could be found that in both Tocharian languages absolute construction distinguishes two basic positions in the sentence:

✓ **Interposition (S-Abs-(O)V)**: in a "normal" case the AC is placed between the subject and the predicate of the main clause:

21	*tane*	*plaktu(kä)ñña*	***brāhmaṇeṃ***	***lyelyakormeṃ***
	then	female_doorkeeper	brahman:OBL.PL	see:ABS

	kercīyenn(e)	*yopsa*
	palace:LOC	enter:PT.3SG

"Then the (female) doorkeeper, having seen the Brahmins, entered the palace" (B 81 b3-4).

✓ **Preposition (Abs-S(O)V)**: if the AC contains a sentential pronoun which functions as an anaphoric reference to the previous text, such AC regularly takes initial position in the sentence (п. 5.2.2.):

---

[17] Absolutive is the ablative form of the deverbal noun in -*r* built to the past participle (TA -*r-äṣ*, TB -*r-meṃ*):
- TA *klyos-* "hear" > PPt *kaklyuṣu* > *kaklyuṣur** > Abs *kaklyuṣuräṣ*;
- TB *klyaus-* "hear" > PPt Nom. *keklyauṣu*, Obl. *keklyauṣo-* > *keklyauṣor* > Abs *keklyauṣormeṃ*.
In this chapter only the Abs I (ablative absolute) is discussed. The rarer Abs II (perlative absolute, TA -*r-ā*, TB -*r-sa*) is not concerned in the dissertation.

22   **täm**   **kärsoräṣ**   *ṣuliñi*   *ñäktañ weyeṃ*
     SENT   know:ABS   mountain:ADJ.NOM.PL god:NOM.PL   astonished

     nāṃtsuṣ   *tränkiñc*
     be:PPT.NOM.PL   say:PR.3PL

"The mountain gods, who took note of this, speak in wonder" (A 75 a1; translation: CEToM).

In Tocharian A almost no deviations from this rule were attested. In Tocharian B some irregularities can be observed which are discussed in §5.2.3. It turns out that in all cases of an apparently deviating word order (5 to 8 cases in total) we deal with an initial AC without a sentential pronoun but never with an interpositioned AC containing such a pronoun. It is also remarkable that AC with a demonstrative pronoun tends to be in the initial position in Tocharian B but always adheres to the basic pattern formulated above in Tocharian A, i.e. it takes interposition.

As further discussed in §5.3., absolutive construction in both Tocharian languages can have a subject different from that of the main clause. This is a unique property of the absolutive compared to other Tocharian predicates. In such examples the absolutive construction always has initial position, for instance:

23   **te**   **keklyauṣormeṃ**   *araṇemiñ*   *lānte*   *pit*   *maiwāte-ne*
     SENT   hear:ABS   Araṇemi:GEN king:GEN gall   tremble:PT.3SG-PN.3SG

   "When he had heard this (lit. "Having heard this"), the gallbladder of king Araṇemi trembled." (B 85 b4-5).

It is possible that non-coinciding subjects in sentences with AC are admissible because morphologically the absolutive is a noun.

In section 5.4. converbial constructions with the meaning of antecedence in Old Uighur and partially in Sogdian are discussed. Beside the OU corresponding constructions with the converb in *-Xp* (noticed in Pinault 2015: 95-97) another parallel which might be interesting from the structural point of view is comprised of 'secondary' converbs in *-dOk-dA* and *-mIš-tA*, i.e. locative constructions with absolute past participles in *-dOk* and *-mIš*. Exactly the sentences with the constructions in *-dOk-dA* and *-mIš-tA* (but never those in *-Xp*) allow non-coinciding subjects:

24   *kim* *ymä* **tängri tängrisi**   **burxan**   **nirwanqa**   **b(a)rmïšta**
     REL   CNJ   god   god:IZ   Buddha   Nirvana:DAT go:PTCP.LOC

     kin *šazanta*   *uluɣ* *asaɣ* *tusu*   *qïltïlar*
     then disciple:LOC   big   use   benefit create:PT.3PL

"(Wiederum verneige ich mich vor der Majestät derjenigen,) die ebenfalls, nachdem der Göttergott Buddha ins Nirvana eingegangen ist, in der Disziplin großen Nutzen und Vorteil gebracht haben". (MaitrHami 12b 1-3; Geng 1988: 53).

As far as can be judged from the examples found, the 'synthetic' converb in *-(X)p* is only used if its subject coincides with the subject of the main clause. Instead, forms in *-mIš+tA* and *-dOk+dA* often (but not always) occur in "heterosubjective" sentences. This property of 'secondary' converbs is pointed out in [Erdal 2004: 320]. According to Erdal, in such a case the subject of a 'secondary' converb stands usually in nominative but can also be in genitive. This is comparable with the situation in Tocharian where the non-canonical subject of an AC can also take nominative (examples 74-76 in Chapter 5 of the dissertation) as well as genitive case (examples 83 and 87).

Concerning the word order, it seems that Old Uighur converbial constructions with the meaning of antecedence always stand in preposition to the main clause, although this observation is made on a limited number of examples. If this is true, the Tocharian syntax with the varying position of the absolutive construction is unique.

In Section 5.5. the results got in Chapter 5 are summarized. Absolutive construction has clause-initial position only in a half of cases in Tocharian A and in two thirds of cases in Tocharian B, for which mostly an anaphoric element (a sentential pronoun) is required, see table 3 below. It is pointed out that the rule formulated above is stronger in Tocharian A than in Tocharian B. This fact may be caused by a certain "artificiality" of Tocharian A because most of the known Tocharian A texts are religious compositions, and it is moreover not excluded that Tocharian A was a dead language by the time it was written down.

Table 3

AC position	AC type	TA	TB
Interposition	without anaphoric element	46	18
	with sentential pronoun	0	1
	with demonstrative pronoun	5	2
Initial position	without anaphoric element	2	8
	with sentential pronoun	37	14
	with demonstrative pronoun	1	7
	with non-canonical subject	6	9
Postposition		2	4
Contexts in total		94	63

The **Conclusion** lists and summarizes the main results of each Chapter of the dissertation. It is noted that this study can be one of the steps towards compiling comparative grammars for the linguistic area of the Tarim Basin; it also can be of interest for research on the typology of syntactic borrowings.

PUBLICATIONS REFLECTING MAIN RESULTS OF THE STUDY

Publications in journals included in the international Scopus database and in the lists of high-profiled journals recommended by NRU HSE:

1.   Kuricyna A.V. Ešče raz o sojuznoj funkcii toxarskix slov A śla, ślak, B śle, śale, ślek // Indoevropejskoe jazykoznanie i klassičeskaja filologija XX (1) – SPb: Nauka, 2016 – ss. 559-572.

2.   Kuritsyna Anna V. Tocharian B manuscripts 498 and PK AS 3A: Two parallel medical texts // Manuscripta Orientalia 23 (1) – St. Petersburg, 2017 – pp. 54-60.

3.   Kuritsyna Anna V. Suspended affixation with Tocharian adjectival suffix A -ṣi / B -ṣṣe and its possible parallel in Old Uighur. // Journal of Language Relationship: International Scientific Periodical / Russian State University for the Humanities, Russian Academy of Sciences. Institute of Linguistics; ed. By V.A. Dybo – Moscow, 2018 – № 3-4 (16) – pp. 265-276.

4.   Kuritsyna Anna V. Preposition Repetition in Tocharian // Indo-European Linguistics 4 – Brill, 2016 – pp. 190–208 – URL:

http://booksandjournals.brillonline.com/content/journals/22125892/4/1

Other publications:

5.   Kuricyna A.V. Lingvističeskaja zadača «V stixax i proze» na materiale toxarskogo A i toxarskogo V jazykov // XLVI Moskovskaja tradicionnaja Olimpiada po lingvistike: Zadači. II tur – M., 2016.

6.   Itkin Ilya B., Anna V. Kuritsyna & Sergey V. Malyshev. Tocharian A text THT 1331 and the "Höllenkapitel" of the "Maitrisimit nom bitig": some more remarks // Tocharian and Indo-European Studies 18 – Museum Tusculanum Press, 2017 – pp. 71-82.

7.   Itkin Ilya B. & Anna V. Kuritsyna. Chapter XX of "Maitreyasamiti-Nāṭaka" and its hellish sufferings: the fragment THT 1308.a1 // Tocharian and Indo-European Studies 18 – Museum Tusculanum Press, 2017 – pp. 63-70.

8.   Itkin Ilya B. & Anna V. Kuritsyna. Again on nomina agentis in Tocharian B: some new observations. // Tocharian and Indo-European Studies 20 – Museum Tusculanum Press, 2020 – pp. 83-106.