



Universiteit
Leiden
The Netherlands

Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian

Warries, A.R.

Citation

Warries, A. R. (2025, June 18). *Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian*. Retrieved from <https://hdl.handle.net/1887/4250485>

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/1887/4250485>

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

Tocharian and Samoyed

On the question of
Uralic substrate influence in Tocharian

Tocharian and Samoyed
On the question of Uralic substrate influence in Tocharian

Abel Warries
PhD dissertation
Leiden University, The Netherlands

Copyright © 2025 Abel Warries. All rights reserved.

Cover image: pastel painting by Helen van Hoorn.

This research was carried out as part of the research project “The Tocharian Trek: A linguistic reconstruction of the migration of the Tocharians from Europe to China”, funded by the European Research Council under the European Union’s Horizon 2020 Excellent Science programme (Grant agreement ID: 758855).

Tocharian and Samoyed

On the question of Uralic substrate influence in Tocharian

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 18 juni 2025
klokke 13:00 uur

door
Abel Radu Warries
geboren te Amsterdam
in 1997

Promotores: Prof. dr. A.M. Lubotsky
Prof. dr. M. Peyrot

Promotiecommissie: Dr. O. Khanina (University of Helsinki)
Prof. dr. A. Kloekhorst
Prof. dr. G.J. Kroonen
Dr. M. Zhivlov (National Research University HSE, Moscow)

Acknowledgements

This dissertation is the product of ten years at Leiden University, in the company of many inspiring friends and colleagues for whom I have the utmost appreciation. First and foremost, I would like to thank my supervisors, Michaël Peyrot and Sasha Lubotsky, for their encouragement and for sharing their insights throughout these years. They not only guided me throughout these arduous years of PhD research, but also helped form me as a linguist and a researcher. I am also grateful to the rest of the teaching staff at the Leiden Linguistics programme, in particular the CIEL section: Tijmen Pronk, who guided me when I wrote my BA thesis, Guus Kroonen, Alwin Kloekhorst, and Lucien van Beek.

I would not have stayed in Leiden for this long if I had not met so many wonderful people here. With Anouk Mudde and Lotte Meester I shared a love for Celtic, with Virginia Oogjen a budding interest in Uralic, and with Femke Montagne the world of birds (may we see many more together). When I embarked on my PhD, I also came to the world of natural history at Naturalis, which ended up being a continuous part of my life over the last five (coming on six) years. I would like to thank everyone at the museum for their kindness, and especially my PB colleagues for lifting my spirits in the difficult final stages of working on this dissertation. I sincerely thank Martijn Knapen and Thom van Hugte for their continued friendship over the years and lament the many kilometers of treacherous train tracks that separate us these days. Many thanks also to my longest Leiden friend and now paranymp Cid Swanenvleugel, without whom this day would have been unimaginable.

I have been fortunate to share the Tocharian part of this journey with Alessandro del Tomba, Chams Bernard, Federico Dragoni, Olga Lundysheva, Xiaoqiang Meng, and Yezi Mu, and I especially thank my fellow project members and friends Louise Friis, Niels Schoubben and Stefan Norbruis, and my office mates Ahmed Sosal and Saskia Dunn. Over the years, it has furthermore been a great pleasure and a privilege to work (and have lunch and borrels) together with such excellent people as Andrew Wigman, Anthony Jakob, Axel Palmér, Bente de Graeve, Carmen Spiers, Hester Groot, Jesse Wichers Schreur, Jiang Wu, Lis Kerr, Natasja Delbar, Nora Julmi, Olga Nozdracheva, Paulus van Sluis, Rasmus Torsø, Simon Poulsen, Sophia Nauta, Tobias Søborg, and Xander Vertegaal. As the last of the pandemic-time Indo-Europeanist PhD group to now leave the nest, I am eternally grateful for everyone's continued encouragement and support.

The Uralic world and the Samoyed part of this journey opened up to me more properly at a later stage while working on this PhD project. I am grateful for the teachings of Gerson Klumpp, Josefina Budzisch, Mikhail Zhivlov, Tapani Salminen, Valentin Gusev, and I especially thank Olesya Khanina for welcoming me at Helsinki University during the spring of 2023 and for the many thought-provoking discussions. My sincere thanks also to Juho Pystynen, Kaisla Kaheinen, Petri Kallio, and Sampsa Holopainen for

discussing with me all aspects of historical Uralic and Samoyed linguistics, and for their helpful feedback. My stay in Helsinki was further brightened by the friendship of Simone and Tuomo.

On this long and winding road I would further like to give mòran taing to the team of the Sabhal Mòr Ostaig summer courses, especially to my teachers Alec Bhaltois, Catriona Nic an t-Saoir, and Muriel Fisher, and also to the many other kind people I met during my visits to Eilean a' Cheò. I have the fondest memories of days and late nights spent in merry Gàidhlig conversation with Eilís, Emma, Jennie, Josie, Lewis, Liam, Marcel, Melvin, Natalie, Nathaniel, and many more.

In the one free week between the end of the MA and the start of the PhD, I took a short vacation to Vienna and met Eugenio Loi. I could never have imagined then that we would still be such close friends all these years later, and I am overjoyed that he agreed to be here by my side as one of my paranymphs. I am also grateful to Simon Rey for his continued friendship, and I look forward to continuing our walks, talks and museum visits for many years to come.

Last, but certainly not least, I come to the people who have been in my life the longest. Joosje, Koen, Marion, and René have been like a second family for many years, and I have the greatest appreciation for their continued friendship and hospitality. To my parents, Helen and Michael, I owe the biggest debt of gratitude. Without their help and advice and their belief in me I would be standing here. Thank you both.

Table of Contents

1	Introduction	1
1.1	A Uralic piece to the Tocharian puzzle	1
1.2	Samoyed and the other Uralic languages	3
1.3	Archaeological background.....	7
1.4	Methodological considerations	10
1.5	Research questions	13
2	Relative chronology of Tocharian sound changes	17
2.1	Introduction.....	17
2.2	Tocharian and Proto-Indo-European phonology	18
2.3	Early consonant developments.....	21
2.3.1	Laryngeals	22
2.3.2	Syllabic resonants	25
2.3.3	Palatovelars and labiovelars.....	27
2.3.4	The stop system.....	36
2.3.5	Summary	49
2.4	Palatalization	51
2.4.1	Palatalization as it affected the consonant system.....	51
2.4.2	Developments of PIE <i>*i</i> and <i>*e</i>	54
2.4.3	Developments of PIE <i>*u</i>	55
2.4.4	Palatalization as it affected the vowel system.....	57
2.4.5	Developments of PIE <i>*ō</i> and <i>*eh₂/ā</i>	59
2.4.6	Summary	63
2.5	Late vowel developments.....	65
2.5.1	<i>u</i> -umlaut of pre-PT <i>*e</i>	65
2.5.2	<i>o</i> -umlaut of pre-PT <i>*u</i>	66
2.5.3	<i>o</i> -umlaut of pre-PT <i>*e</i>	67
2.5.4	<i>o</i> -umlaut of pre-PT <i>*a</i>	69
2.5.5	<i>a</i> -umlaut of pre-PT <i>*e</i>	70
2.5.6	Effects of labial consonants on vowels.....	72
2.5.7	Summary	83
2.6	Information from language contact	83

2.6.1	Early Iranian languages	83
2.6.2	“BMAC language”	86
2.6.3	Summary	86
2.7	Overview of Tocharian relative chronology.....	87
3	Relative chronology of Samoyed sound changes	95
3.1	Introduction.....	95
3.2	Samoyed and Proto-Uralic phonology	95
3.3	Loss of final PU <i>*-i</i>	98
3.4	Changes affecting single consonants.....	99
3.4.1	PU <i>*δ</i> , <i>*s</i> and <i>*š</i>	99
3.4.2	PU <i>*l</i> , <i>*δ'</i> and <i>*j</i>	101
3.4.3	PU <i>*w</i> and <i>*γ</i>	104
3.4.4	Summary	106
3.5	Changes affecting consonant clusters.....	107
3.5.1	PU <i>*rk</i> and <i>*lk</i>	107
3.5.2	PU <i>*kC</i> , <i>*Ck</i> and geminates.....	108
3.5.3	PU pre-consonantal glides.....	110
3.5.4	Summary	110
3.6	Changes affecting non-low vowels	111
3.6.1	PU <i>*u</i> , <i>*i</i> and <i>*ü</i>	111
3.6.2	PU <i>*e</i> and PS <i>*e</i> and <i>*i</i>	116
3.6.3	PU <i>*o</i>	118
3.6.4	Summary	121
3.7	Changes affecting PU <i>*a</i>	121
3.7.1	PU second-syllable <i>*-a</i> to <i>*-â</i> and <i>â</i>	122
3.7.2	PU first-syllable <i>*a</i> to <i>*â</i> and <i>*a</i>	124
3.7.3	PU second-syllable <i>*-a</i> to <i>*-ä</i>	126
3.8	Information from language contact: Yukaghir.....	126
3.9	Overview of Samoyed relative chronology.....	128
4	Phonological comparison.....	131
4.1	Introduction.....	131
4.2	Comparison of the consonant systems	132
4.2.1	The pre-Proto-Tocharian and pre-Proto-Samoyed stop systems.....	132

4.2.2	Stop systems of other languages in the region.....	134
4.3	Comparison of the vowel systems	136
4.3.1	The pre-Proto-Tocharian and pre-Proto-Samoyed vowel systems.....	136
4.3.2	Vowel systems of other languages in the region.....	139
4.4	Overview and conclusions	141
5	The Tocharian accent	143
5.1	Accentuation in Tocharian	143
5.1.1	Basics of the Tocharian B accent	144
5.1.2	Problems with the Tocharian B accent	145
5.2	Epenthesis and syncope	147
5.2.1	General remarks on epenthesis.....	147
5.2.2	Epenthesis in <i>na</i> -presents.....	149
5.2.3	Epenthesis in <i>s</i> -preterites	162
5.2.4	Epenthesis in <i>Cst</i> and <i>Csk</i>	163
5.2.5	Syncope.....	164
5.2.6	The accent through the lens of syncope and epenthesis	165
5.2.7	Intermediary summary.....	167
5.3	The accent in finite verbs	168
5.3.1	Mobile and fixed accent in the Tocharian B verb	168
5.3.2	The accent in the thematic present [2] and <i>s</i> -present [8]	172
5.3.3	The accent in the <i>s</i> -preterite [3]	174
5.3.4	The accent in the subjunctives [1] and [5].....	179
5.3.5	The accent in the causative present-subjunctive [9] and preterite [2]..	196
5.3.6	The phonologization and morphologization of the accent	199
5.4	The accent in other categories.....	200
5.4.1	Vowel contractions and the accent in the preterite participle.....	200
5.4.2	Nasal prefixes and the accent	208
5.4.3	The accent in kinship terms.....	214
5.4.4	The accent in the secondary cases	215
5.4.5	The accentuation of compounds	217
5.4.6	The accentuation of loanwords.....	218
5.5	Tocharian A accentuation.....	219
5.5.1	Apocope, vowel weakening and syncope in Tocharian A	219
5.5.2	Vowel contractions in Tocharian A.....	221
5.5.3	Indirect indications of pre-Tocharian A accent	222

5.6	Conclusions on the accent	223
6	Agglutinative case systems.....	225
6.1	Introduction.....	225
6.2	The Tocharian case system	228
6.2.1	The functions of the Tocharian secondary cases	229
6.2.2	The forms of the Tocharian cases	234
6.3	The Samoyed case system	241
6.3.1	The Samoyed cases and their functions.....	241
6.3.2	The Proto-Samoyed and Proto-Uralic case systems	246
6.4	Comparison of the Tocharian and Samoyed case systems	250
6.4.1	Case inventory	250
6.4.2	Functional comparison	251
6.4.3	Comparison of development	252
6.5	Conclusions on the agglutinative case systems.....	253
7	Participial systems	255
7.1	Introduction.....	255
7.1.1	Participles and their orientation.....	256
7.2	Tocharian	260
7.2.1	The <i>nt</i> -participle.....	260
7.2.2	The <i>m</i> -participle	262
7.2.3	The preterite participle	268
7.2.4	The gerundives: necessitative and potential participles.....	281
7.2.5	The privative: a negative participle.....	288
7.2.6	Summary	290
7.3	Samoyed.....	291
7.3.1	Uralic and areal context	297
7.4	Conclusions on the participial systems	298
8	Object marking on the verb	301
8.1	Introduction.....	301
8.2	The Tocharian pronoun suffixes.....	303
8.2.1	Functions of the pronoun suffixes.....	304
8.2.2	Origins and status of the pronoun suffixes	307
8.2.3	Similar pronominal elements in other Indo-European languages.....	309

8.3	The Uralic objective conjugation	312
8.3.1	General overview of the objective conjugation.....	312
8.3.2	Formal reconstruction of the objective conjugation	314
8.3.3	Functional reconstruction of the objective conjugation	350
8.4	Comparison between Tocharian and Uralic	362
9	Potential loanwords.....	367
9.1	Introduction.....	367
9.2	Earlier comparisons.....	368
9.2.1	pre-PT <i>*septə(m)</i> ‘seven’ ⇒ pre-PS <i>*säjt³wə</i> ‘id.’	368
9.2.2	pre-PT <i>*wesa</i> ‘gold’ and PS <i>*wäsa</i> ‘metal, iron’.....	369
9.2.3	pre-PT <i>*mēnē-</i> ‘moon’ ⇒ PS ± <i>*mänä/iwǰə</i> ‘full moon’	370
9.2.4	pre-PT <i>*monko-</i> ‘lack, deficit’ ⇒ PS <i>*māṅkâ</i> ‘poverty, need’	370
9.2.5	Pre-PT <i>*kwənə(m)</i> acc. of ‘dog’ ⇒ pre-PS <i>*wēn</i> ‘dog’	371
9.3	New comparisons.....	372
9.3.1	pre-PT <i>*lonko-</i> ‘steep slope; cliff’ ⇒ PS <i>*lāṅkâ</i> ‘steep bank; steep’	372
9.3.2	pre-PT <i>*lenk-</i> ‘dangle, sway, hang’ ⇒ pre-PS <i>*lāṅkâ-</i> ‘hang, dangle’	372
9.3.3	pre-PT <i>*der-</i> ‘split, separate’ ⇒ pre-PS <i>*tar-</i> ‘divide, separate’	373
9.3.4	pre-PT <i>*newo-</i> ‘new’ ⇒ Pre-PS <i>*ñuwV</i> ‘child, young boy’	374
9.3.5	pre-PT <i>*mēywyā</i> ‘power, strength’ ⇒ pre-PS <i>*mejâ</i> ‘sturdy, solid’	374
9.4	Discussion.....	376
10	Conclusions.....	379
10.1	Overview	379
10.2	Phonology.....	380
10.3	Accent.....	382
10.4	Morphology	383
10.5	Lexicon	385
10.6	Evaluation.....	386
	Etymological reference list 1: Tocharian.....	389
	Etymological reference list 2: Samoyed	413
	Bibliography.....	457

Abbreviations of languages

Av.	Avestan	Md.	Mordvin
E	Erzya (Mordvin)	Ms.	Mansi
EM	Erzya and Moksha	Mt.	Mator
En.	Enets	Ne.	Nenets
EnF	Forest Enets	NeF	Forest Nenets
EnT	Tundra Enets	NeT	Tundra Nenets
Goth.	Gothic	Ng.	Nganasan
Gr.	Greek	OSIr.	Old Steppe Iranian
Fi.	Finnish	PIE	Proto-Indo-European
FU	Finno-Ugric	PNe.	Proto-Nenets
Hitt.	Hittite	PS	Proto-Samoyed
Hu.	Hungarian	PSaa.	Proto-Saami
Iir.	Indo-Iranian	PT	Proto-Tocharian
Kh.	Khanty	PTrk.	Proto-Turkic
KhKaz.	Kazym Khanty	PU	Proto-Uralic
KhO	Obdorsk Khanty	PYuk.	Proto-Yukaghir
KhSur.	Surgut Khanty	PYen.	Proto-Yeniseian
KhTrj.	Tremjugan Khanty	SaaN	North Saami
KhVj.	Vasjugan Khanty	SaaS	South Saami
KhVVj.	Vakh-Vasjugan Khanty	Sk.	Selkup
Khot.	Khotanese	Skt.	Sanskrit
KoZ	Komi Zyrian	TA	Tocharian A
Lat.	Latin	TB	Tocharian B
M	Moksha (Mordvin)	Udm.	Udmurt

Abbreviations of grammatical terms and glosses

1	first person	agn.	agent noun
2	second person	all.	allative
3	third person	aor.	aorist
A	agent	caus.	causative
abl.	ablative	cneg.	connegative
abs.	absolutive	com.	comitative
acc.	accusative	conj.	conjunction
act.	active	conn.	sentence-connector
adj.	adjective	dat.	dative
adjz.	adjectivized	def.	definite

dem.	demonstrative	opt.	optative
dist.	distal (demonstrative)	ord.	ordinal
dst.	(pre)destinative	P	patient
ela.	elative	pass.	passive
emph.	emphatic	per.	perlative
erg.	ergative	pers.	personal (pronoun)
ext.	extension	pf.	perfect
f.	feminine	pfv.	perfective
freq.	frequentative	pl.	plural
fut.	future	PN	proper noun
gen.	genitive	poss.	possessive
ger.	gerundive	priv.	privative
ill.	illative	pro.	prolative
imp.	imperative	prog.	progressive
indef.	indefinite	prohib.	prohibitive
inf.	infinitive	pron.	pronoun
ins.	instrumental	prs.	present
ipf.	imperfect	prt.	preterite
ipfv.	imperfective	pst.	past
itr.	intransitive	ptc.	participle
iter.	iterative	ptcl.	particle
lat.	lative	refl.	reflexive
loc.	locative	rel.	relative (pronoun)
m.	masculine	rprt.	reportative
mid.	middle	S	subject
n.	neuter	sbc.	subjective conjugation
necess.	necessitative	sbj.	subjunctive
neg.	negative	sg.	singular
nom.	nominative	vn.	verbal noun
obc.	objective conjugation	voc.	vocative
obl.	oblique		

Other abbreviations

arch.	archaic	id.	same meaning as previous
der.	derivation	lit.	literally
dial.	dialectal	p.c.	personal communication
fn.	footnote	trsl.	translation
ibid.	same reference as previous		

General symbols

[_]	phonetic representation
/_/	phonological representation
_	morphological representation
<_>	infix in morphological representation
<_>	orthographic representation
*_	reconstructed form
**_	non-existent or incorrect form
*_	expected form, e.g., in a partially attested paradigm
A > B	A developed into B by sound change
A >> B	A developed into B by analogy
A → B	B is derived from A by derivation or semantic shift
A ⇒ B	B is derived from A by borrowing

Special symbols

<i>ć</i>	provisional notation for an intermediary stage between PIE <i>*t_i</i> and PT <i>*t^s</i> , see 2.3.4.2
<i>đ</i>	provisional notation for an unclear intermediary development of PIE <i>*d</i> in early pre-Proto-Tocharian, see 2.3.4.4
<i>ž^h</i>	provisional notation for an intermediary stage between PIE <i>*d^hǵ</i> and PT <i>*t^s</i> , see 2.3.4.2

Cover symbols

<i>C</i>	any consonant
<i>N</i>	any nasal
<i>n¹</i>	Proto-Samoyed <i>*n</i> or <i>*ń</i>
<i>P</i>	any labial or labialized consonant
<i>R</i>	any resonant (<i>l, m, n, r</i>)
<i>t¹</i>	Proto-Samoyed <i>*c</i> or <i>*t</i>
<i>t²</i>	Proto-Samoyed <i>*c</i> or <i>*s</i> or <i>*t</i>
<i>t³</i>	Proto-Samoyed <i>*c</i> or <i>*k</i> or <i>*s</i> or <i>*t</i>
<i>V</i>	any vowel