

# **Parametric Syntactic Reconstruction**

Noun Phrases in Iranian, Proto-Indo-Iranian  
and Proto-Indo-European

INAUGURALDISSERTATION

zur

Erlangung des Doktorgrades der Philosophischen Fakultät der  
Universität zu Köln

im Fach

Historisch-Vergleichende Sprachwissenschaft

vorgelegt von

**SVENJA BONMANN**

aus Oberhausen.

Köln, 16. August 2023

Bei der vorliegenden Arbeit handelt es sich um eine leicht überarbeitete Fassung meiner Dissertation im Fach Historisch-Vergleichende Sprachwissenschaft, die am 13.04.2022 von der Philosophischen Fakultät der Universität zu Köln angenommen wurde. Mit dieser Veröffentlichung komme ich den Pflichten nach, die sich aus § 17 der Promotionsordnung der Philosophischen Fakultät der Universität zu Köln vom 12.03.2022 (Amtliche Mitteilungen 09/2020) ergeben. Die Prüfungskommission setzte sich aus den folgenden Mitgliedern zusammen:

1. Prof. Dr. Eugen Hill (Erster Referent)
2. Prof. Dr. Daniel Kölligan (Zweiter Referent)
3. Prof. Dr. Nikolaus P. Himmelmann
4. Prof. Dr. Birgit Hellwig
5. Prof. Dr. Peter Orth
6. Prof. Dr. Martin Becker

Datum der Disputatio: 02. Mai 2022

Die Arbeit an dieser Dissertation wurde durch den Schweizerischen Nationalfonds finanziell gefördert (SNF-Projektnummer: 170241, “The Evolution of Noun Phrases in Indo-Iranian: empirical foundations and theoretical modeling”; 01.09.2017–31.08.2020).



# Abstract

This dissertation is a defense of the hypothesis that syntax can be reconstructed using the time-tested Comparative Method and structural analogues to phonological features, so-called parameters.

Any application of the Comparative Method depends on a systematic comparison of basic units which are finite in number and have discrete values. This dissertation argues that syntactic parameters might equal distinctive phonological features and that different combinations or rather bundles of parameters might constitute syntactic ‘phonemes’. Building upon the noun phrase (NP) properties and behavior of nine Old and Middle Iranian languages, namely Old and Young Avestan, Old and Middle Persian, Parthian, Bactrian, Chorasmian, Sogdian and Khotanese Saka as well as the corresponding NP characteristics of three ancient Indo-European relatives, namely Vedic, Archaic (i.e., Mycenaean and Epic) Greek and Old Latin, this study tries to reconstruct, based on a parametric feature matrix, (a) the nominal syntactic behavior (formally representable in terms of parametric settings) of the last common ancestor of all Iranian languages, Proto-Iranian, (b) the parametric profile of the still earlier Proto-Indo-Iranian (PIIr.) language, and it makes (c) an educated first guess on the NP parameters of Proto-Indo-European (PIE).

Key findings of this dissertation are (I) the insight that parametric syntactic reconstruction parallels phonological reconstruction(s); (II) the observation that PIIr. and PIE had a right-branching nominal configurational syntax with a high amount of noun-raising (no rigid head-finality); (III) the inference that the non-IE substrate language(s) of the early towns of Bronze Age Central Asia (BMAC and related cultures) may have had articles.

**Keywords:** Linguistic Reconstruction, Syntactic Reconstruction, Noun Phrases, Iranian, Proto-Indo-Iranian, Proto-Indo-European



# Acknowledgements

I am very grateful to the Swiss National Science Foundation for the generous financial support of my PhD project. In my view, the academic freedom of research and teaching implies a strong obligation to advance human knowledge and to push the limits of a scientific branch like historical-comparative linguistics even further, and I hope that the money invested in my work has paid off.

It goes without saying that this project would not have been possible without the help of many people contributing to it in one way or the other. I would first like to express my gratitude to those people that helped me during my time in Zurich, most notably **Balthasar Bickel** for his constant push towards more robust and scientific linguistic research and **Andreas Hölzl** for the countless times we took a stroll along Seilergraben or through Niederdorf and spoke about everything and anything, and at times even about my dissertation.

Among the many people with whom I interacted in Zurich (many of which I shamefully must exclude from these Acknowledgements, because they *sensu stricto* did not directly contribute to my PhD project) I would like to highlight **Florian Sommer**, **Florian Wandl** and **Tim Felix Aufderheide** for discussions, exchange of literature and for giving me many valuable insights and ideas (either on Iranian or on Balto-Slavic, Indo-Aryan or other topics) that helped shape the text into its final form, even though this may not be obvious. Thank you very much, **Stefan Dedio**, for many helpful hints and tricks concerning L<sup>A</sup>T<sub>E</sub>X, and thank you, **Chundra Cathcart**, for honestly trying (yet ultimately failing) to convince me of the advantages of so-called “quantitative comparative linguistics”.

Proceeding to Cologne, a special thanks goes to **Eugen Hill** who not only taught me the working procedure of historical-comparative linguistics years ago, but who is also the best supervisor you could ever wish for. Thank you, Eugen (or should I say: *\*h<sub>1</sub>suǵenh<sub>1</sub>és*), for being an inspiring mentor and an invaluable source of information on all matters Indo-European, Uralic and “Palaeo-Siberian” and, of course, for being a wonderful colleague. I’m afraid I cannot share your fascination for Insular Celtic (even though you did your best to inspire passion

in me), but you succeeded in making me a reconstruction enthusiast.

Thank you, **Daniel Kölligan**, for having taught me several ancient Indo-European languages, some of which made it into this study (e.g., Mycenaean Greek), and for your constant optimism and feedback on my project and its theses and topics. Thank you very much, **Agnes Korn**, for helping me in so many ways and for always lending an ear for the smaller and larger problems occurring during the last months of my project. Your extremely detailed corrections of and comments on earlier versions of the manuscript, your Iranianist expertise and invaluable moral support when coping with the hardships of writing a dissertation helped tremendously in bringing this project to a successful end.

Thank you, **Nikolaus Himmelmann**, for being the best typological guide and teacher I could wish for and for motivating me to opt for a *disputatio* instead of a mere *defensio*. I will forever be grateful for introducing me (special thanks at this point also to **Sonja Riesberg** and **Maria Bardají i Farré**) to the fascinating Austronesian language family and for convincing me to submit my dissertation despite its many shortcomings and flaws.

Thank you very much, **Jakob Halfmann** and **Natalie Korobzow**, for sharing my fascination for undeciphered writing systems. I can only guess how many hundred hours each of us three invested in the decipherment of the “unknown Kushan script”, but every minute of this endeavor was worth the effort, because it, or rather you two, helped me forget the troublesome work on my dissertation for some time. Thanks also for the many discussions on typological and Indo-Iranianist topics, for your accurate proofreading of my text and for sharing literature with me.

Thank you very much, **Simon Fries**, for your inspiring insistence on Neogrammarian ideas and achievements and for being always helpful and selfless. Our discussions on the methodology of reconstructions significantly improved the scientific rigor of my manuscript, and I would like to say a heartfelt thanks to you. Thanks as well, **Michael Frotscher**, for your feedback on my theses.

Of course, the greatest thank you goes to my family. My parents, **Gregor** and **Klaudia Bonmann** (*\*ph<sub>2</sub>ter meh<sub>2</sub>tēr-k<sup>u</sup>e*), as well as my sister **Caroline Scholtysik, née Bonmann** and my brother-in-law **Raphael Scholtysik** always believed in me and supported me since day one of this project. They, as well as my nephews, **Henri** and **Klaas**, provided welcome diversion in the darkest hours of my PhD project, as did **Filou**, the loyal *\*k<sub>u</sub>on-* of my family. I dare say that I would not have been able to complete my doctorate without your support.

Thank you very much!

# Contents

<b>List of Figures</b>	<b>IX</b>
<b>List of Tables</b>	<b>XII</b>
<b>Linguistic Gloss Abbreviations</b>	<b>XV</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The challenge of syntactic reconstruction . . . . .	1
1.2 Linguistic reconstruction and the Comparative Method . . . . .	9
1.3 Sketching parametric syntactic reconstruction . . . . .	18
1.4 Hypotheses . . . . .	21
<b>2 Theoretical Background</b>	<b>25</b>
2.1 A short outline of the research history . . . . .	25
2.2 Parametric approaches to diachronic syntax . . . . .	39
2.3 Noun Phrase parameters . . . . .	45
2.3.1 The basic NP/DP structure . . . . .	50
2.3.2 NPs: core and periphery, arguments and adjuncts . . . . .	54
2.4 Parametric hierarchies . . . . .	57
<b>3 Noun Phrases in Iranian and beyond</b>	<b>85</b>
3.1 General remarks . . . . .	85
3.2 Notes on the classification of Iranian and on the choice of languages	87
3.3 Old Iranian languages . . . . .	95
3.3.1 Old Avestan . . . . .	95
3.3.2 Young Avestan . . . . .	105
3.3.3 Old Persian . . . . .	112
3.4 Middle Iranian languages . . . . .	119



3.4.1	Khotanese . . . . .	119
3.4.2	Sogdian . . . . .	126
3.4.3	Chorasmian . . . . .	137
3.4.4	Bactrian . . . . .	147
3.4.5	Parthian . . . . .	160
3.4.6	Middle Persian . . . . .	167
3.5	Selected ancient Indo-European languages . . . . .	174
3.5.1	Vedic . . . . .	174
3.5.2	(Archaic) Greek . . . . .	185
3.5.3	(Old) Latin . . . . .	193
<b>4</b>	<b>Parametric Syntactic Reconstruction</b>	<b>203</b>
4.1	A note on Indo-Iranian ethnogenesis and language contact . . .	204
4.2	Step 1: The syntaxeme inventory . . . . .	212
4.2.1	Proto-(Indo-)Iranian . . . . .	212
4.2.1.1	Unambiguous reconstructions . . . . .	213
4.2.1.2	Ambiguous evidence . . . . .	217
4.2.1.3	The (Indo-)Iranian parametric features . . . . .	230
4.2.2	Proto-Indo-Iranian = Proto-Indo-European? . . . . .	236
4.2.3	Proto-Indo-European: adding Greek and Latin . . . . .	237
4.3	Step 2: Is there a syntactic equivalent to sound laws? . . . . .	242
4.3.1	Syntactic Cognates . . . . .	243
4.3.2	Syntaxemes: syntactic phonemes . . . . .	250
4.3.3	Syntactic Minimal Pairs . . . . .	264
<b>5</b>	<b>Conclusion</b>	<b>269</b>
	<b>Appendix</b>	<b>277</b>
5.1	Parametric Feature Matrix . . . . .	277
5.2	Phonological Feature Matrix . . . . .	279
5.3	Farsi . . . . .	281
5.4	Pashto . . . . .	286
	<b>Bibliography</b>	<b>289</b>

# List of Figures

1.1	Schematic representation of linguistic reconstruction I. . . . .	12
1.2	Schematic representation of linguistic reconstruction II. . . . .	13
1.3	Schematic representation of linguistic reconstruction III. . . . .	13
1.4	Schematic representation of parametric reconstruction (following Roberts 2021, pp. 505–506.) . . . . .	19
2.1	The three most basic parameters FGP, FGN and FGG. . . . .	49
2.2	Base-generated DP/NP structure. . . . .	51
2.3	Example for functional and free genitives. . . . .	55
2.4	Free genitives, ‘base-generated’ order. . . . .	56
2.5	Free genitives in a language with ‘raised’ N. . . . .	56
2.6	Relative clauses. . . . .	56
2.7	Noun-raising parameters. . . . .	59
2.8	Number-related parameters. . . . .	65
2.9	Definiteness-related parameters. . . . .	67
2.10	‘Boundedness’-related parameters. . . . .	70
2.11	Further definiteness-related parameters. . . . .	73
2.12	Adjectival parameters. . . . .	73
2.13	Further adjectival parameters. . . . .	74
2.14	Genitival parameters. . . . .	75
2.15	Further genitival parameters. . . . .	77
2.16	Relative clause parameters. . . . .	77
2.17	Parameters with regard to possessives. . . . .	78
2.18	Linker/ <i>eṣāfe</i> parameters. . . . .	81
3.1	Traditional phylogeny of Iranian. . . . .	89
4.1	Premodern Iranian languages: relative chronological layers. . . . .	212

4.2	Parametric syntactic reconstruction: flat hierarchy (X = some parameter). . . . .	213
4.3	Parameter NOO: Noun raising over ordinals. . . . .	218
4.4	Parameter NOC: Noun raising over cardinals. . . . .	218
4.5	Parameter NDE: Noun raising over demonstratives. . . . .	220
4.6	Parameter DGP: Grammaticalized partial definiteness. . . . .	221
4.7	Parameter CGB: Grammaticalized boundedness. . . . .	225
4.8	Parameter GUN: Uniform genitive. . . . .	226
4.9	Parameter GPR: DP/NP over free genitive. . . . .	226
4.10	Parameter APO: Adjectival possessives. . . . .	227
4.11	Parameter AGE: Adjectival genitives. . . . .	228
4.12	Parameter TPL: Strong partial locality. . . . .	228
4.13	Parameter VRC: Verbless relative clauses. . . . .	229
4.14	Parameter FLI: Facultative linker. . . . .	229
4.15	Proto-Indo-Iranian, basic genitival NPs, syntaxemic representation (e.g., Y. 50.10 <i>raocā xʷāng</i> ‘the rays of the sun’ or RV 3.22.3 <i>rocané (...) sūryasya</i> ‘in the sun’s realm of light’). . . . .	255
4.16	Pashto, basic genitival NPs, syntaxemic representation (e.g., <i>de asád (de) plār</i> ‘Asad’s father’, following Tegey and Robson 1996, p. 185). . . . .	256
4.17	Proto-Indo-Iranian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (e.g., RV 6.56.3 <i>cakrām hiranyāyam</i> ‘the golden wheel’). . . . .	258
4.18	Proto-Indo-Iranian, basic NP with demonstrative, syntaxemic representation (e.g., Y. 36.6: <i>imā raocā</i> ‘these lights’ or Y. 32.13; 45.3,4 <i>aṇhāuš ahiiā</i> ‘of this world’). . . . .	259
4.19	Proto-Indo-Iranian, basic NP with cardinal numeral, syntaxemic representation (e.g., Khotanese, Z. 22.162 <i>ratanyau haudyau</i> ‘(full of) seven jewels’). . . . .	259
4.20	Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 31.6 (West 2011, p. 147) <i>mąθrām yim hauruuatātō</i> ‘a spell [which is one of] integrity ..’). . . . .	260
4.21	Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 35.4 <i>tāiš šiiāoθanāiš yāiš vahištāiš</i> ‘with those deeds that (are) the best’). . . . .	261
4.22	Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 35.4 <i>tāiš šiiāoθanāiš yāiš vahištāiš</i> ‘with those deeds that (are) the best’); base-generated order. . . . .	261

- 4.23 Young Avestan, complex NP with a verbless relative clause being in a genitival relation to a head noun, syntaxemic representation (e.g., Young Avestan [Yt. 13.91] *staota ašahe yaṭ mazištaheca vahištaheca sraēštaheca* ‘The prayer of Aša, the greatest and best and most beautiful’). . . . . 262
- 4.24 Vedic, adjectival NPs with genitival reading, syntaxemic representation (e.g., RV 10.96.6 *dāsā vrtrāṇi (āryā)* ‘the Dasic (and Aryan) enemies’). . . . . 264
- 4.25 Archaic Greek, basic adjectival NPs (attributive, not predicative), syntaxemic representation (e.g., Il. 2.658 (etc.) βίη Ἡρακλείη ‘for Herakles’ force’ or ‘for the Herakleian force’). . . . . 265
- 4.26 Old Persian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (DZc 6 *bumiyā vazrakāyā* ‘great earth’). 265
- 4.27 Middle Persian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (*zamīg ēw wuzurg (ud istabr)* ‘a great [and solid] earth’, following Durkin-Meisterernst 2014, p. 361). . . 266



# List of Tables

1.1	Polynesian sound correspondences (adopted from Campbell 2020, pp. 180–181, but slightly corrected). . . . .	12
1.2	Proposed solutions to the ‘correspondence problem’ in syntactic reconstruction. . . . .	17
2.1	Parametric Feature Matrix: NP parameters in Old and Middle Iranian languages and selected other early attested IE languages; the parameters, their abbreviations, and the logical relationships are based on Longobardi et al. (2013). . . . .	48
3.1	Premodern Iranian languages and their main grammatical descriptions. . . . .	87
4.1	Language contact scenarios following Hajnal (2018). . . . .	210
4.2	Parametric Grid: NP parameters in Vedic, Old and Middle Iranian languages and Proto-Iranian. . . . .	232
4.3	Lexically identical syntactic cognates of the oldest Iranian and Indo-Aryan representatives, based on Schlerath (1968, pp. 148–164). . . . .	245
4.4	Lexically identical cognate NPs of the oldest Iranian and Indo-Aryan representatives with irrelevant surface differences in terms of word order, based on Schlerath (1968, pp. 148–164). . . . .	247
4.5	Basic genitival NPs (Old Avestan, Young Avestan, Vedic and Proto-Indo-Iranian), parametric configuration. . . . .	248
4.6	Basic genitival NPs (Proto-Indo-Iranian and Pashto), parametric archetype/configuration. . . . .	248

5.1	Parametric Feature Matrix: nominal macro- and mesoparameters of Old and Middle Iranian languages compared with those of other early attested IE languages. . . . .	278
-----	--	-----

# Linguistic Gloss Abbreviations

1	first person	IPRF	imperfect
2	second person	LINK	linker
3	third person	LOC	locative
A	agent	M	masculine
ABL	ablative	MED	medium
ACC	accusative	N	neuter
ACT	active	NOM	nominative
AOR	aorist	OBL	oblique
ART	article	OPT	optative
DAT	dative	PFV	perfective
DEM	demonstrative	PL	plural
DES	desiderative	PN	personal name
DIR	direct case	PRET	preterite
DO	direct object marker	PRF	perfect
DU	dual	PRS	present
EZ	<i>eṣāfe</i>	PST	past
F	feminine	PTCL	particle
GEN	genitive	PTCP	participle
IMP	imperative	PTRN	patronymic
IMPF	imperfective	REFL	reflexive
IND	indicative	REL	relative
INDF	indefinite	SBJV	subjunctive
INF	infinitive	SG	singular
INJ	injunctive	VN	verbal noun
INS	instrumental	VOC	vocative





Wer kann was Dummes, wer was Kluges denken,  
Das nicht die Vorwelt schon gedacht?

---

MEPHISTOPHELES  
*Faust II, Vers 6809 f.*



# Chapter 1

## Introduction

### 1.1 The challenge of syntactic reconstruction

Well, *\* $\hat{k}léuos \acute{n}d^h g^{uh} itom$* . If someone asks me why I wrote a dissertation about the reconstruction of syntax, this may be the answer. This small syntagma is the classic handbook example of a reliably reconstructible Proto-Indo-European (PIE) noun phrase (NP),<sup>1</sup> and everlasting or imperishable fame surely is a worthy goal for a young researcher. As regards a verb phrase (VP), or rather a whole (albeit very short) sentence, one could mention the famous *\*(h<sub>1</sub>é)g<sup>uh</sup> ent h<sub>1</sub>óg<sup>uh</sup> im* ‘he slew the serpent’ (Watkins 1995, pp. 302, 365, 370).<sup>2</sup> Such phrases, together with the statement that PIE was an SOV (Subject–Object–Verb) language with little configurationality, correlative relative clauses, eight cases in the nominal domain and a ‘Wackernagel position’ for clitics, is basically what every student of Indo-European (IE) or comparative linguistics learns as the state of research on PIE syntax.

---

<sup>1</sup>The reconstruction given above must be the predecessor of Homeric κλέος ἄφθιτον (Il. 9.413) and Vedic *śrávas ... akṣitam* (RV 1.9.7bc) ‘imperishable fame’, cf. Kuhn (1853, p. 467), Schmitt (1967, p. 1) or Watkins (1995, pp. 173–178).

<sup>2</sup>Ved. *áhann áhim* (RV 1.32.1 etc.) ‘he slew the serpent’, Gk. κτεῖνε ... ὄφιν (Pindar *Pythian* 4.249) ‘he killed the serpent’, Av. *yōjanaṭ ažiṃ dahākəm* (Y. 9.8) ‘who slew Aži Dahāka’. Note that this specific ‘formula’ reflects a marked word order (VO) according to Watkins (1995, pp. 301–302).

As a rule, current handbooks or introductions to Indo-European studies or, in more general terms, historical-comparative linguistics only cursorily describe how this *communis opinio* on PIE syntax came into being, what the methodology for reconstructing syntax looks like and how it works.<sup>3</sup> Instead, sentences, clauses or smaller phrases of older IE languages are juxtaposed, and the reader is then encouraged to accept that these carefully selected examples point towards a certain pattern or feature inherited from PIE. This is syntactic reconstruction as students learn it.

To illustrate this procedure, consider the following verb-final clauses or sentences from ancient or medieval representatives of the Anatolian, Indo-Iranian, Greek, Italic, Celtic, Germanic and Tocharian branches of Indo-European:<sup>4</sup>

(1) Examples for verb final order in ancient or medieval IE languages

a. Hittite (KBo 3.7 i 11 [Illuyanka myth, §3])

*nu=za*      <sup>MUŠ</sup>*illuyankaš* <sup>D</sup>*IMan*      *tarahta*  
and=REFL PN.NOM.SG      Stormgod.ACC.SG overcome.PRET.3SG

“And Illuyankaš himself overcame the Stormgod.”

b. Runic Norse (Gallehus horn)

*ek hlewagastiz holtijaz*      *horna*      *tawido*  
I PN.NOM.SG Holtian.NOM.SG horn.ACC.SG make.PRET.ACT.1SG

“I, Hleugast, Holtian, made (this) horn.”

<sup>3</sup>See the introductory chapters on PIE syntax and syntactic change in Clackson (2007), Fortson (2010) or Meier-Brügger (2010). Szemerényi (1990) and Tichy (2009) do not discuss syntax at all, Beekes (2011) has two (!) pages on the topic (pp. 96–97, of 415 pages all in all). Similarly, introductions written for a more general (i.e., not specifically Indo-Europeanist) readership usually describe mechanisms (and case studies) of syntactic change, but not the methodology for reconstructing syntax (e.g., see Hock and Joseph 2009, pp. 183–202, or Crowley and Bowerman 2010, pp. 217–245); Campbell (2020, pp. 297–305) has at least a few pages devoted to the issue.

<sup>4</sup>Examples 1a, 1b, and 1c are taken from Fortson (2010, p. 157), the others were added by me. In general, and unless noted otherwise, translations are my own.

- c. Tocharian A (Pun̄yavantajātaka 1)

*kāsu*                      *ñom-klyu*                      *tsraššī*                      *śāk*  
 good.NOM.SG.M name-fame.NOM.SG vigorous.GEN.PL ten  
*kälymentwaṃ*    *sātkatār*  
 direction.LOC.PL spread\_out.MED.PRS.3SG

“Good fame of the vigorous (ones) spreads out in ten directions.”

- d. Vedic (RV 4.27.1)

*ádha śyenáh*                      *javāsā*                      *nís adīyam*  
 but falcon.NOM.SG speed.INS.SG out fly.IPRF.ACT.1SG

“but I flew out with rapid speed (as/like) a falcon.”

- e. Greek (Iliad 17.80)

Τρώων                      τὸν                      ἄριστον                      ἔπεφνε  
 Trojan.GEN.PL ART.ACC.SG best.ACC.SG slay.AOR.ACT.3SG

“He slew the best of the Troians.”

- f. Venetic (\*Es 122, normalized transliteration)

*vinetikaris vivoi*                      *oliiale=kve murtuvoi*                      *atisteit*  
 PN.NOM.SG alive.DAT.SG ?=CONN dead.DAT.SG set\_up.PRS.3SG

“Vinetikaris<sup>5</sup> (oliiale, ADV?) sets up [this stone] for (the) living and (the) dead.”

- g. Gaulish (RIG II.1, L-3; Eska 2007)

*ratin*                      *briuatiom*                      *frontu*                      *tarbetisonios*  
 fort.ACC.SG bridge\_dweller.ACC.SG PN.NOM.SG PTRN.NOM.SG  
*ieuru*  
 dedicate.PRET.3SG

“Frontu, son of Tarbetisu, dedicated the fort of the bridge-dwellers.”

The conclusion that PIE had an unmarked verb-final and/or OV order seems almost inevitable, given the fact that so many and such diverse daughters attest

<sup>5</sup>Cf. the Gaul. PN *ueni-carus*.

verb-finality, and consequently this is usually claimed in introductions.<sup>6</sup> Alternatively, I could have collected examples in which a subordinate clause has a verb-final order and an immediately following main clause a verb-initial one. This pattern is found both in Latin (e.g., Terence, *Eunuchus* 629–631: *dum rus eo, coepi egomet mecum ... cogitare*; Cicero, *Epistulae ad Fam.* 1,7: *quod mihi de filia et de Crassipede gratularis, agnosco ...*) and Ancient Greek (e.g., Xenophon, *Anabasis* 1.8.8: ἡνίκα δὲ δειλὴ ἐγίγνετο, ἐφάνη κονιορτός ...; Plato, *Phaido* 59 d: ἐπειδὴ δὲ ἀνοιχθείη, εἰσῆμεν ...). Should it be reconstructed, then, also for PIE?

The problem with this approach is that many ancient and medieval IE languages evince a bewildering variety of apparently legit surface realizations of syntactical constituents such as predicates or NPs. This makes it difficult to determine which surface realization is or was the unmarked or default one, as we cannot simply ask speakers of these languages to give us judgements on topicalization, fronting, stylistic differences vel similia. Even though verb-final word order is attested for many of the ancient IE languages, the claim that this is the unmarked or default order in all these IE languages could be doubted for at least some of them. For instance, the just cited Venetic and Gaulish sentences are practically the only ones with verb-final position. The question arises whether precisely verb-finality or SOV order is (or was) a marked construction in these languages and whether SVO or OVS was the default pattern instead. Most other attested sentences do not have a basic verb-final order, cf., e.g.,

---

<sup>6</sup>See e.g., Fortson (2010, p. 157): “It is almost universally asserted that most of the ancient IE languages were verb-final and that PIE was as well; more specifically, that they were SOV (Subject-Object-Verb).” Beekes (2011, p. 97) states (bold type as per l.c.) that “[t]he oldest IE languages indicate the presence of a **word order** wherein the object was placed before the verb (OV, object-verb).” Similarly Clackson (2007, p. 166): “If PIE were a ‘non-configurational’ language, with completely free word order, we would still have to explain why the unmarked place of the verb is sentence-final in Hittite, Sanskrit and Latin, and why word comparisons which reflect original juxtapositions of noun and dependent genitive agree in showing the order genitive – head noun (...) Agreements such as these could lead to the conclusion that the unmarked order was SOV for the PIE sentence, and head-final for the PIE noun phrase.” Note also Hock (2013).

- (2) Venetic (Es 53, normalized transliteration)

*me go    donasto                      reitiiai                      nerrka*  
 I.ACC.SG give.IND.AOR.MED.3SG theonym.DAT.SG PN.NOM.SG  
*lemetorna*  
 daughter.of.Lemeter(?).NOM.SG  
 “Nerka Lemetorna gave me to Reitia.”

- (3) Gaulish (CIL XIII 2880; RIG L-13)

*martialis    dannotali    ieuru                      ucuete                      sosin*  
 PN.NOM.SG PTRN.GEN.SG dedicate.PRET.3SG theonym.DAT.SG DEM.ACC.SG  
*celicnon*  
 Celicnon.ACC.SG  
 “Martialis, son of Dannotalos, dedicated this Celicnon to Ucuētis.”

Sentences like 1f and 1g are frequently brought forward as being particularly significant for a reconstruction of syntactic patterns of the proto-language, because their assumed non-regularity from a synchronic perspective can be claimed to be a syntactic archaism. This may be the case. But it could also be the other way round, and languages with strong verb-finality like Latin and Hittite could be innovative, and Venetic and Gaulish conservative. Who can tell? We should likewise not ignore the fact that examples for verb-initial sentences of more ‘canonical’ IE languages (compared with Venetic and Gaulish) can be presented without great difficulty:

- (4) Sanskrit (Mahābhārata 3.50.1)

*āsīd                      rājā                      nalo                      nāma*  
 be.IPRF.3SG king.NOM.SG PN.NOM.SG by\_name  
 “There was a king named Nala.”

- (5) Greek (Sappho, 49.1 L-P; following Watkins 1995, p. 107)

*ἠράμην                      μὲν ἔγω σέθεν,                      Ἄτθι,                      πάλαι πότε*  
 love.IPRF.MED.1SG PTCL I    you.GEN.SG PN.VOC.SG earlier once  
 “I loved you once, Atthis, long ago.”



As regards NPs, a noun phrase PIE *\*dems potis* ‘master of the house’ (with GEN–NOM), as reflected in Ved. *dám pati-*, Av. *dāng paitiš* ‘lord’ and Gk. δεσπότης, is the standard example presented in handbooks to demonstrate PIE head finality in NPs (see e.g., Clackson 2007, p. 166). It is, however, in no way difficult to find noun phrases with a following instead of preceding genitival attribute in the ancient IE languages; returning to the ‘master of the house’, Ved. also has the order NOM–GEN, *pátir dán*, which is usually not mentioned in handbooks (Keydana 2018, p. 2205 is a notable exception).<sup>7</sup> Note also the match – and hence probable inheritance from common late PIE times – between the religious/mythological collocation ‘daughter of the day’: we find Ved. *duhitā diváh* and Gk. θυγάτηρ Διός besides *divó duhitā* and Διὸς θυγάτηρ. Many noteworthy examples for NOM–GEN order from other languages and branches could be brought forward, importantly also for rather rigorous head-final languages in the verbal domain (e.g., Lat. *tribunus plebis*, *magister equitum*, *orbis terrarum*, *praefectus urbis* etc.). Clearly, there was variation, and it is not as straightforward as usually claimed to determine which one was the default or unmarked NP order. Syntactic reconstruction, at least in the way it is presented in current handbooks, is not as unambiguous and objective as suggested. This is acknowledged by Hale (1987b, p. 2):

“Discontinuous NPs, no fixed Adj/N-N/Adj or Gen/N-N/Gen order, no fixed verb position, and unusual relative clause structures (to name but a few characteristics) are constantly encountered in the texts in these languages. Indeed the same complications clearly con-

<sup>7</sup>Cf. also RV 10.61.20 *śísur dán* ‘child of the house’. The argument that the phrase PIE *\*dems potis* shall be particularly revealing with regard to PIE NP-internal word order due to its fossilized nature does not convince me, given the Greek theonym Ποσειδῶν (Homeric Ποσειδάων, Doric Ποτειδάων, Ποτιδᾶς etc.) in which the exact opposite order appears to be fossilized. Whatever the second part of the compound is or was – a ‘master of water(s)’ is tantalizing, with an otherwise lost cognate to Avestan *dānu-* ‘river’, Ossetic *don* ‘water’ and preserved Scythian river names in eastern Europe like *\*dān(u)-apara-* ‘river to the rear’, *\*dān(u)-nazdya-(ra-)* ‘river to the front’, not to mention the *Danube* –, it seems clear that the first part reflects PIE *\*poti-* ‘lord, master’ and the second some other noun in the genitive (*\*d(e)h<sub>2</sub>uo-*?), interestingly both in the GEN.PL and GEN.SG.

tinue to mystify typologists, who have claimed PIE to have been SVO, SOV and VSO in their efforts to fit that language into their neatly designed ‘ideal’ types.”

The argumentation and the concrete reconstruction will crucially depend on the choice of examples, if the process of reconstruction entails and is understood just as a mere projection back in time of what one can find in – preferably early attested – daughter languages. This, in turn, means that there is an inevitably high degree of subjectivity in both the selection of examples supporting the argument and in the syntactic reconstruction itself. You reconstruct what you search for, in essence.

Here we should stop for a moment and reflect on the issue. What exactly do we mean by the term ‘syntactic reconstruction’? Is the procedure just sketched a ‘real’ linguistic reconstruction? Do we want to project back in time, as the handbooks suggest, superficial patterns found in the oldest attested daughter languages? How, then, can we avoid the pitfalls of selecting exactly those examples that fit our pre-existing expectations?<sup>8</sup> And how can we know that our presumed cognates are, in fact, cognates? Is searching for superficially similar utterances from a wide variety of languages sufficient? Does this procedure not resemble language comparison of amateurs, of the type Modern Persian (Farsi) *bad* : English *bad*; Korean *mani* : Engl. *many*, Nahuatl (Aztecan) *huel* : Engl. *well*? Didn’t our forerunners develop a sober methodology in order to precisely avoid such superficial comparisons? Isn’t linguistic reconstruction much more rigorous and, to be quite plain, scientific in other areas?

In fact, the way syntactic features are reconstructed for PIE does in no way resemble morphological or lexical reconstructions arrived at by using a rigorous methodology. Reading current handbooks – and ‘real’ studies on syntactic re-

---

<sup>8</sup>Following ‘Teeter’s law’ – C. Watkins, *Selected Writings I*, Innsbruck 1994, p. 247 (see also Tichy 2009, p. 8): “The language of the family you know best always turns out to be the most archaic” – one could state: “The syntactic feature you know best always turns out to be the most archaic.”

construction (more on which in Chapter 2) – creates the impression of language comparison and reconstruction in a pre-Neogrammarian style of the early 19th century. By just presenting carefully selected examples from the older languages, one can ‘prove’ and ‘reconstruct’ almost anything.

One of the few places in the literature in which this methodological pitfall is clearly articulated – though not condemned –, is Jamison (1993, p. 219):

“For, when all the techniques and methods have been laid out, in the end the best way to approach the writing of the synchronic syntax of a dead language is to try to cultivate a pseudo-Sprachgefühl, which we can only acquire by the diligent and careful reading of texts. This was, of course, the ‘secret’ of the great syntacticians of the 19th century, one that seems all too often forgotten now.”

I could not disagree more. This is exactly the reason why the “great syntacticians of the 19th century” failed, and why syntactic reconstruction is commonly ignored (just recall footnote 3). To follow Jamison in her approach to syntax implies an abandonment of objectivity, replicability and falsifiability of our reconstruction(s) and hence of our scientific methodology and approach. “Pseudo-Sprachgefühl” is but a euphemism for intuition and subjective judgement. I do not doubt that such distinguished scholars as Delbrück, Brugmann, Wackernagel, Jamison or Watkins may have had a good conscious or unconscious understanding of the underlying syntactic rules of Vedic, Avestan, Greek, Latin etc. After all, each of them was (or is, in the case of Jamison) an excellent philologist. I also do not doubt that they thought, while writing their treatises on syntax, they could project their intuitive understanding of syntactic rules and patterns to PIE itself. But what I doubt is that this approach is a scientific one. Syntactic reconstruction should not lead us to dismiss all our practices. We do not reconstruct morphemes or lexemes based on our “pseudo-Sprachgefühl”, but by means of a rigorous and replicable methodology. Why should syntax be any different?

It is true that many features encountered in the oldest attested languages can lead the linguist to a conscious or unconscious feeling of what sounds natural

or right. But even if a so-called “pseudo-Sprachgefühl” takes shape after reading relevant texts, it must in no case lead us astray to mistake our feelings for scientific evidence. This cannot be stressed enough.

## 1.2 Linguistic reconstruction and the Comparative Method

The only solution to this problem is a recourse to our standard practice of reconstruction, and not a recourse to completely new methods.<sup>9</sup> The procedure for the reconstruction of syntactic aspects should be virtually the same as the one used for the reconstruction of morphemes or lexemes. This, in turn, implies an application of the time-tested Comparative Method (CM), the “gold standard” (Kiparsky 2014, p. 65) of our discipline (on which see Harrison 2003, Hale 2014 or Clackson 2017, among others). The Comparative Method, despite being more a heuristic tool than a real diagnostic method in the narrower sense (Nichols 1996), fulfills at least the scientific criteria of replicability/reproducibility and falsifiability; it is as close to a scientific approach as one can get in historical linguistics. We would therefore do well to adopt the CM in our efforts to reconstruct syntax.

Many scholars rush to judgement here, as it is commonly believed that “syntactic reconstruction is a different type of enterprise from phonological reconstruction, and [that] it is not possible to compare the two directly” (Clackson

---

<sup>9</sup>This is not the right place to discuss the activities of linguistic amateurs (arachnologists, ornithologists, primatologists, mathematicians, statisticians, geographers etc.) trying to reinvent historical-comparative linguistics by using methods adopted from bioinformatics – the so-called ‘quantitative comparative linguistics’ program. Suffice to say that the results published so far have met severe criticism (see e.g., Pereltsvaig and Lewis 2015 for a book-length critique; cf. also Campbell 2020, p. 441: “The models adopted from biology do not fit or do not take into account many aspects of language change known to be significant, for example analogical change, chain shifts, directionality of many changes, how entities are constrained by and dependent upon other elements in particular language subsystems, sociolinguistic conditioning of change, impacts of language contact, reanalysis, grammaticalization, avoidance of homophony, aspects of semantic change and neologisms, taboo, and on and on.”).

2007, p. 158, see also Clackson 2017).<sup>10</sup> Consequently, even though in theory the CM would be the method of choice, many linguists refrain from using it due to this perceived difference. And did we not just see that there are indeed many problems with the syntax of ancient IE languages? How, then, can we still lay hope on an application of the CM?

We should first of all remind ourselves of the basic procedure, and we should consider how the CM works and what it entails and what not. This does not seem to be well understood by many people active in the field (e.g., by those engaged in so-called ‘Construction Grammar’, as we will see in Chapter 2). To avoid a common misconception, the Comparative Method comprises a search for systematic correspondences between basic discrete units – phonemes. Linguistic reconstruction is first and foremost an entirely mechanical, formal process. Words in any language are chains or strings of phonemes which only secondarily carry a certain semantic (or pragmatic) load. What is of interest to us is the chain of phonemes, the string of basic units, as this, and only this, can be logically analyzed; semantic identity is, at best, of secondary value.<sup>11</sup>

We always compare and reconstruct cognate sequences of sounds. It is true that in the end form-meaning pairs are proposed, but only after the heuristic process is – provisionally – completed. Not infrequently, the process gets stuck or is never really completed, namely when the formal part of the reconstruction is clear, but the meaning of the reconstructed string of phonemes is shady.<sup>12</sup> The

---

<sup>10</sup>Note also the comparisons between an application of the CM in phonology and syntax in Jeffers (1976), Lightfoot (1979, pp. 154–167), Harris and Campbell (1995, pp. 343–378), Roberts (2007, pp. 357–368), Roberts (2021, pp. 499–516), Willis (2011), Walkden (2013), Barðdal and Eythórsson (2012, pp. 261–262) or Barðdal (2013, 439 ff.).

<sup>11</sup>See Allen (1953, pp. 57–60) for a general critique on semantic matches in linguistic comparisons.

<sup>12</sup>Cf. cases like Lat. *nepōs*, Vedic *nāpāt-* ‘grandson’, OHG. *nefo* ‘grandson; nephew’, Welsh *nei* ‘nephew, sister’s son’, or their female equivalents Lat. *neptis*, Ved. *naptī-* ‘granddaughter’, OHG. *nift*, Welsh *nith* ‘niece’. We can formally reconstruct these lexemes for PIE (\**nepōt-* in the former case, \**neptih<sub>2</sub>-* in the latter), but we do not actually know what they meant back in the times the parent language was spoken (‘nephew’ or ‘grandson’?, and ‘niece’ or ‘granddaughter’?). Similarly, did \**syekruh<sub>2</sub>-* – formally without doubt reconstructible – originally denote the ‘mother-in-law’ (on both sides) or only ‘husband’s mother’?

formal part must be complete, the reconstruction of the string of sounds must be unambiguous, whereas the meaning allows for a certain flexibility.

While applying the CM, we routinely set up so-called correspondence sets, compare the items therein, search for systematic, i.e., recurring patterns between the languages compared (say, a phoneme /f/ in language A always corresponds to a /p/ in languages B, C and D) and explain individual pathways as due to certain changes of the basic discrete units (e.g., \*/p/ > /f/ in language A of our just mentioned example). Such developments are without exceptions according to the traditional Neogrammarian tenet, and exceptions which seem to contradict the pattern can be either explained by more sophisticated ‘conditioned sound changes’ or ‘sound laws’ (e.g., \*/p/ > /f/ only after unaccented syllables, otherwise it remained /p/) or by secondary processes (analogy, language contact including interdialectal borrowing, neologisms etc.). If we assemble a correspondence set like that of Table 1.1 (Polynesian languages, adopted from Campbell 2020, pp. 180–181; see the next page), it won’t be difficult for us to discern systematic matches between the languages involved and the apparent outcomes of (conditioned) sound changes from their last common ancestor to each of them.

If we are interested in reconstructing a certain phoneme X of Proto-Polynesian, we will compare the corresponding values or features of its daughter languages. We will set up a list of words which correspond to each other and we will compare the values of the phoneme(s) in question, as in Table 1.1 (only with more examples). To give but one example, let us assume that we are interested in the reconstruction of a phoneme which has a binary opposition between /t/ and /k/ in the daughter languages (no phonemic split or conditioned sound change etc. has occurred); see Figure 1.1.

We would do this for dozens of lexemes in order to get a systematic picture of the patterning of the languages. If we take into account both economic considerations (the less changes to the daughter languages a model assumes, the better) and our knowledge of a cross-linguistically/typologically common direction of change for this feature/phoneme (if such is indeed known) – say, a change from

Māori	Tongan	Samoa	Rarotongan	Hawaiian	gloss
tapu	tapu	tapu	tapu	kapu	‘forbidden, taboo’
pito	pito	—	pito	piko	‘navel’
puhi	puhi	—	pu’i	puhi	‘blow’
taha	tafa (‘edge’)	tafa	ta’a	kaha	‘side’
tae (‘trash’)	ta’e	tae	tae	kae	‘excrement’
tanata	tanata	tanata	tanata	kanaka	‘man, person’
tai	tahi	tai	tai	kai	‘sea’
kaha	kafa	’afa	ka’a	’aha	‘strong’
ma:rohi-	ma:lohi	ma:losi	ma:ro’i	—	‘strong’
karo	kalo	’alo	karo	’alo	‘dodge’
aka	aka	a’a	aka	a’a	‘root’
au	’ahu	au	au	au	‘gall’
uru (‘tip of weapon’)	’ulu	ulu	uru	ulu (‘centre’)	‘head’
uhi	’ufi	ufi	u’i	uhi	‘yam’
ahi	afi	afi	a’i	ahi	‘fire’
fa:	fa:	fa:	’a:	ha:	‘four’
peke	feke	fe’e	’eke	he’e	‘octopus’
ika	ika	i’a	ika	i’a	‘fish’
ihu	ihu	isu	puta-i’u (‘nostril’)	ihu	‘nose’
hau (‘wind’)	hau	sau	’au	hau	‘dew’
hika	—	si’a	’ika	hi’a	‘firemaking’
hiku (‘fishtale’)	hiku	si’u	’iku	hi’u	‘tail’
ake	hake	a’e	ake	a’e	‘up’
uru	huu	ulu	uru	ulu	‘enter’

Table 1.1: Polynesian sound correspondences (adopted from Campbell 2020, pp. 180–181, but slightly corrected).

the value /t/ to /k/ is common, the opposite not –, we can easily explain daughter 5 (Hawaiian here, with daughters 1 to 4 being Māori, Tongan, Samoan and Rarotongan) as being deviant and we may then reconstruct a phoneme /t/ for our proto-language.

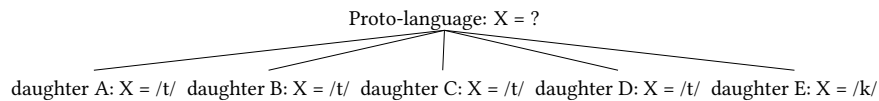


Figure 1.1: Schematic representation of linguistic reconstruction I.

Now, compare this with a more elaborate version, in which X does not stand for one particular phoneme generally in the languages, but for a whole word.

This word is analyzed as a string of four different slots A, B, C and D, and these slots, in turn, are filled by different phonemes (left empty in Figure 1.2).

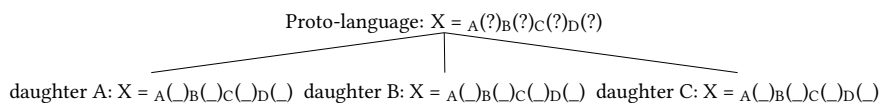


Figure 1.2: Schematic representation of linguistic reconstruction II.

To illustrate this, consider the concrete example of ‘ten’ in four different Indo-European languages, viz. Greek, Avestan, Latin and Tocharian A.

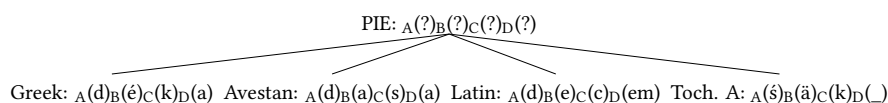


Figure 1.3: Schematic representation of linguistic reconstruction III.

As can be seen, Latin has two phonemes /e/ and /m/ in slot D, whereas Tocharian A has no phoneme at all. Three of the compared languages have a velar plosive in slot C, whereas one has a sibilant. Also three languages, though not the same ones, have a dental plosive in slot A, whereas one has a sibilant, etc. Evidently, a slot can be filled by one single phoneme or by more than one, if phonemic split has occurred in the prehistory of one of the daughter languages. Likewise a phonemic merger could have happened or the phoneme in one particular position (e.g., at the end of the word) could have been lost due to certain developments (loss of all word-final vowels etc.).

Based on these and other cognates from other branches (Germanic, Armenian, Balto-Slavic etc.), and of course our accumulated knowledge about individual sound changes (including general and conditioned ones) and cross-linguistic directions of change for various sounds, we reconstruct a voiced (perhaps pre-glottalized/emphatic) dental plosive /d/ for slot A, a vowel /e/ for slot B, a ‘palatal’ velar /k̑/ for C and a syllabic nasal /m̥/ for D, and hence *\*dek̑m̥* ‘ten’ for PIE.



What is meant by stating that the Comparative Method – and linguistic reconstruction in general – is a heuristic process is that after a while, after general sound laws and smaller, conditioned sound changes have been worked out, semantics becomes nearly irrelevant. Admittedly, the process starts with compiling a correspondence set based on semantically ideally congruent lexemes in the languages compared, but after some time the reconstructive process is decoupled from semantics. As soon as sound laws are known, one can compile other sets of cognate lexemes; cognates one only recognizes on the basis of their phonological, purely formal ‘skeleton’ and the sounds found therein. For example, it is telling that a trained Indo-Europeanist can easily discern the cognacy of Armenian *erkn* ‘birth-pangs’, Greek ὀδύνη ‘pain, harm, sorrow’ (Aeol. ἐδύν-), Luvian *attuwal-* and Hittite *idālu-* ‘evil’, despite the diverging semantics of the individual lexemes. Evidently, the non-Anatolian words are based on earlier PIE *\*h<sub>1</sub>dun-/h<sub>1</sub>(e)d<sub>u</sub>on-* and the Anatolian ones on a differently suffixed *\*h<sub>1</sub>(e)d<sub>u</sub>ol-*.<sup>13</sup> It surely helps to know that Armenian has a fully regular sound change of PIE *\*d<sub>u</sub> > rk* when judging the cognacy of its stem with the Greek and Anatolian ones; based on semantics alone, we would, perhaps, have detected the cognacy with Greek, but probably not with Anatolian.

Linguistic reconstruction thus is a formal, mechanical process. It entails a comparison of discrete units which are organized in cognate strings, and semantics only comes at the end of the formal reconstruction. Ideally the slots in a cognate string are filled with exactly one phoneme each, but a slot can also consist of two phonemes or none at all. Reconstruction, then, comes in two ways:

1. it can be very abstract, systematic, aiming at an elaboration of the phonemic inventory of the proto-language and diachronic pathways (sound laws, conditioned sound changes) leading to each daughter (= phonological reconstruction);

---

<sup>13</sup>A further connection with the verbal root PIE *\*h<sub>1</sub>ed-* ‘to eat (< to bite?)’ is far from obvious from a purely semantic point of view; from a formal, sound change based one, on the contrary, it is perfectly possible.

2. it can be very concrete, if one chooses to focus on particular cognate words and then on their cognate slots (= morphological and lexical reconstruction).

The laws of logic require that concrete morphological or lexical items can only be reconstructed in a second step, after diachronic pathways and (conditioned) sound changes of each branch have been worked out. Only through the knowledge of systematic diachronic pathways and conditioned sound changes is it possible to differentiate between real and false cognates.

Without this methodology, without our accumulated knowledge about systematic correspondences and language-specific sound laws, we would not be able to recognize that French *huit*, Italian *otto* and Spanish *ocho* are, in fact, exact correspondences, all meaning ‘eight’ and going back, phoneme-by-phoneme, to earlier (Vulgar) Latin *\*okto*. Conversely, without our methodology, we could be tempted to compare terms like Farsi *bad* with Engl. *bad*, or Ossetic D. *ævellon* ‘extraordinary, astonishing’ (< ‘unbelievable’, see Cheung 2002, p. 166) with the famous *Insula Avallonis* of the medieval Arthurian epics.<sup>14</sup> We can easily demonstrate that the two latter terms are surely not related by explaining the individual phonemes in terms of their Ossetic and Insular Celtic context;<sup>15</sup> while without the knowledge of the sound laws, this is far from obvious. Only by means of our rigorous method are we able to discern false cognates and to predict expected outcomes of lexemes and morphemes assumed to be inherited from a common parent language.

<sup>14</sup>To my knowledge first attested in Geoffrey of Monmouth’s *Historia Regum Britanniae*; cf. also the Middle Welsh name [*Ynys*] *Afallach*.

<sup>15</sup>The Oss. term (probably) is a derivative – with privative prefix – of the specific East Iranian noun *\*paṛa-* ‘belief’ (thus *\*a-paṛ-jo-n-*) and shows the typical Alanic sound change of *\*r<sub>i</sub> > l* – or, in other words: if an Oss. term has /l/, one will usually explain this as going back to earlier *\*r<sub>i</sub>*. We know of this Alanic sound change because correspondence sets with other Iranian and IE languages have been set up in the past. The Arthurian name *Avalon*, on the other hand, is usually explained as being a Celtic word ‘(island) of apples’ vel sim., which may go back to something similar to Gaulish nom. *\*aballū*, oblique stem *\*aballon-*, as reflected in the French commune of Avallon (originally ‘apple-tree place’ vel sim.).

Syntax, being “the set of rules and principles for the combination of words into larger units – phrases, clauses, and sentences” (Fortson 2010, p. 152), is at first glance something different than morphology or phonology. And it is true that nobody has ever set up, at least to my knowledge, a genuine correspondence set for syntactic cognates. In fact, this is the very reason why syntactic reconstruction has failed so far and why the issue is so often tacitly ignored.<sup>16</sup> The statement that “(...) the discussion about syntactic reconstruction has resurfaced at regular intervals, but in the end one cannot claim the debate to have yielded any new results” (Ferraresi and Goldbach 2008, p. 5) still holds true.

As long as we do not come up with correspondence sets and are able to make predictions on expected outcomes of syntactic cognates in a purely mechanical fashion, without even the slightest hint at a “pseudo-Sprachgefühl” or the like, it will not be possible to differentiate between real cognates and false ones, between superficially perhaps differing, but actually identical cognates (of the type French *huit* : Spanish *ocho*), and vice versa to recognize that superficially similar patterns are in fact mere pseudo-cognates like Farsi *bad* : English *bad*. One can hardly avoid the conclusion that we need syntactic analogues to phonemes in order to tackle this problem.

The fact that syntactic reconstruction suffers from a lack of clear comparanda has been repeatedly remarked (already in the beginnings of modern research on syntactic reconstruction, see Jeffers 1976 or Lightfoot 1979); it is nowadays known under the heading “correspondence problem” or “correspondence set problem” (see Roberts 2021, pp. 505–514, Willis 2011, p. 411 or Walkden 2013). The most important larger ‘schools’ or collaborative teams as well as individual scholars and their ideas with regard to the correspondence problem are given in Table 1.2.

Most approaches suffer from two problems: one regards their scope; instead of actually trying to set up a correspondence set, studies like Walkden (2014),

---

<sup>16</sup>The widespread skepticism is also due to influential critics like Jeffers (1976), Winter (1984), Harrison (2003), Holland (2003), Pires and Thomason (2008), von Mengden (2008); particularly influential also Watkins (1976) and especially Lightfoot (1979, 1980, 1991, 1998, 2006).

School/Scholar	Proposed syntactic comparanda/cognates
Traditional ‘IE school’ (e.g., M. Fritz in Meier-Brügger 2010, p. 374–412)	morphology as ‘paleo-syntax’ (Balles 2008); morphemes
Watkins (1976, 1995) etc.: comparative poetics	formulaic utterances (phrases, fixed collocations)
(Radical) Construction Grammar (Barðdal and Eythórsson 2012, Barðdal 2013, Barðdal and Gildea 2015, Barðdal and Eythórsson 2020 etc.)	‘cognate argument-structure constructions’ in the sense of (Radical) Construction Grammar
Kikusawa (2003, 2020)	‘cognate structures’
Harris (1985, 2008) and Harris and Campbell (1995)	‘syntactic patterns’ which constitute “a set of corresponding sentences” (Harris and Campbell 1995, p. 348)
Lightfoot (2002, p. 135), but also Lightfoot (1979) etc.	whole formulaic sentences (if at all)
Willis (2011)	“obvious answer” would be <i>cognate sentences</i> , but language grammar (l.c., p. 411) makes more sense
Walkden (2013, 2014)	no real correspondence <i>set</i> , just language/lineage-specific single correspondences in grammatical structure
Roberts (2021, pp. 505–514)	parameters (for an application of the CM)
Longobardi (2003), Longobardi and Guardiano (2009, 2017), and Longobardi et al. (2013) and Celin et al. (2020)	parameters (used for phylogenetic studies)

Table 1.2: Proposed solutions to the ‘correspondence problem’ in syntactic reconstruction.

for example, highlight one specific pattern or construction and try to shed some light on its genesis (e.g., the middle affixes *-mk* and *-sk* going back to reinterpreted first- and third-person pronouns in Old Norse). Studies like Walkden’s focus on very concrete syntactic phenomena instead of abstracting and adopting a systematic or rather bird’s-eye view on a whole syntactic module. By doing this, by limiting the focus on sporadic instead of systematic matches between languages – or even focusing on just one single language –, an application of the CM is practically impossible (at best, internal reconstruction would remain as a possibility). This is not syntactic reconstruction in a sense similar to reconstructing morphemes or lexemes; it is rather reminiscent of detailed treatises on the etymology of single lexical items (‘syntactic etymology’, so to speak).

The other problem is a formal one, as Lightfoot (1979, p. 165) half-incorrectly, half-correctly stated: “Reconstruction will be possible via the comparative method only where the daughter languages show identical constructions either

in attested forms or in internally reconstructed abstractions.” This holds true in relation to the CM’s dependence on cross-linguistically comparable basic discrete units, though this does not mean that “identical constructions” must be the syntactic analogue – one just needs something comparable over branch- and lineage-boundaries, something formal like phonemes. And recall that phonemes are by no means identical in the daughters of some parent language; in fact, this is the starting point of the CM.

### 1.3 Sketching parametric syntactic reconstruction

Only the last two of these approaches concentrated on something that could, with a few adaptations, result in correspondence sets as known from phonology or morphology. Both approaches used or use so-called *parameters* as comparanda and are firmly rooted within modern generative grammar (as is, however, also true for Walkden’s approach). Parameters are conceptualized as abstract syntactic choices, as logical possibilities, as classification characteristics with binary values set in each language according to yes/no-questions. Parameters are a tool for us to enable comparability among the languages of the world. I will give a detailed account of parametric frameworks and concepts in Chapter 2; here I will only sketch the aspects relevant for syntactic reconstruction.

The so-called polysynthesis parameter (Baker 1996, Baker 2001, p. 111) is a good example: “Verbs must include some expression of each of the main participants in the event described by the verb (the subject, object, and indirect object).” Parameters like the polysynthesis parameter are discrete abstractions with binary values; there is not much place for a “pseudo-Sprachgefühl” or subjectivity here. The parametric world is black and white, without gray scale. A language like English has the value *no* or *negative* or – (whatever the notation), a language like Mohawk *yes* or *positive* or +.

The members of a research group around Giuseppe Longobardi and Cristina Guardiano (see Longobardi 2003, Longobardi et al. 2013, Guardiano and Longo-

bardi 2017; Longobardi and Guardiano 2009, Longobardi and Guardiano 2017, Ceolin et al. 2020, 2021) have set up grids with parametric values, but they are limiting their efforts to phylogenetic studies.<sup>17</sup> Ian Roberts, on the other hand, once described – very briefly – how syntactic reconstruction within a parametric approach could be possible. He devoted a few pages to the idea in his monograph about a parametric approach to diachronic syntax (Roberts 2007, pp. 364–367). In the second, revised edition (see Roberts 2021, pp. 505–514), this section is somewhat extended, but without substantially changing the basic idea.

If one adopts a parametric approach to syntax, it may be possible to set up correspondence sets and to carry out a classical reconstruction, similar to a comparison of phonemes. Remember that traditional linguistic reconstruction by means of the CM is a heuristic process during which a comparison between cognate features (so far phonemes) enables us to make predictions on their common ancestor, and that the reconstruction gets refined the more items one compares and the more one unveils about diachronic pathways in the languages in focus.

In Figure 1.1, further up in the text, I sketched the basic procedure of abstract, systematic, phonological linguistic reconstruction (with a phoneme as compared feature). Roberts suggests that the same procedure may also be carried out with other abstract entities, namely syntactic parameters. The process can be abstractly visualized as in Figure 1.4 (p = any parameter):

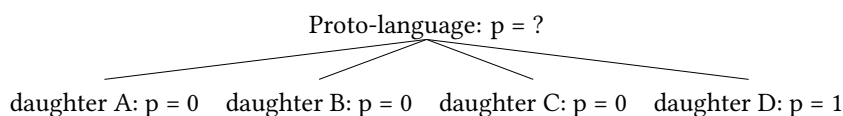


Figure 1.4: Schematic representation of parametric reconstruction (following Roberts 2021, pp. 505–506.)

<sup>17</sup>Curiously, Longobardi et al. take the term ‘reconstruction’ to entail the uncovering of phylogenetic ‘signals’ and not of recovering a proto-language’s grammatical system.

By means of the traditional procedure, we can state that according to economic principles (the less changes one assumes from the parent language to the observed facts of the daughters, the better) daughter D appears to be deviant and that daughters A, B and C are more conservative with regard to this particular parametric value. If we, then, happen to know that a change of the value  $0 > 1$  is a common one for this parameter, this may further strengthen our assumption that daughter 4 is deviant. Consequently, we may reconstruct a parametric value of 0 for the parent language.

This parallelism to a classical reconstruction of phonemes is intriguing. If parameters are used as analogues of phonemes, if reconstruction is indeed a mechanical process dependent upon classifiable and comparable units, it must be possible to recover something like the phonology of a proto-language, i.e., the set of parametric values. If one can make an educated guess on whether or not a proto-language like PIE had three series of ‘velars’ or three laryngeals by means of a thorough comparison of its daughters, why, then, shall it not be possible to determine whether the very same proto-language had other discrete units – ontologically differing, but structurally similar to phonemes?

Phonemes are, from an abstract point of view, a finite set of discrete basic units. In their entirety, phonemes are universally present or rather common to all human languages, but in the concrete case, a synchronically seemingly arbitrary subset of them makes up the phonological system of the individual language, thus allowing for cross-linguistic variation and scientific comparison.

The same applies to parameters, as we will see in Chapter 2. Suffice to say that the process of linguistic reconstruction via the CM depends upon structural analogues to phonemes, and that parameters fulfill this criterion. This makes it reasonable to use parameters as syntactic comparanda.

## 1.4 Hypotheses

A test of this idea is at the heart of this study. The following pages are devoted to Robert's somewhat ignored or at least not immediately followed up suggestion of reconstructing syntax by means of a systematic, i.e., correspondence-set-based comparison of parameters. My hypotheses are:

- Syntactic reconstruction is possible by means of a parametric approach, as parameters can overcome the “correspondence problem” in historical syntax.
- Once cognate sets of parameters have been worked out, the Comparative Method can be applied in order to carry out a reconstruction similar to phonological reconstruction.
- Structural analogues to sound change may exist – syntactic/parametric change, so to say –, and perhaps even parametric analogues to conditioned sound changes (in daughter language A parameter Y changes its value from 0 > 1, if parameter X is set to 0 and parameter Z to 1, vel sim.).
- Parametric syntactic reconstruction is language-unspecific (= universally applicable), falsifiable and reproducible (= independent of the researcher) and thus meets all requirements of being considered a scientific approach.

A parametric approach to syntactic reconstruction, as just sketched, brings with it a rebuttal of Clackson's statement that

“[s]yntactic reconstruction is a ‘second-order’ operation, as is morphological reconstruction and the reconstruction of lexical fields; all rely upon both the reconstructed phonology and the knowledge that the comparanda come from genetically related languages.” (Clackson 2007, p. 158).<sup>18</sup>

---

<sup>18</sup>See also Ferraresi and Goldbach (2008, p. 4) and Harris and Campbell (1995, ch. 12) for similar views.



If parameters can be used for reconstructing something like the phonology of a proto-language (the set of parametric values, as well as individual pathways leading to the daughters), syntactic reconstruction will be a ‘first-order’ operation. A demonstration of a successful parametric reconstruction will open up important possibilities for our discipline, as the last major linguistic subfield – viz. syntactic reconstruction – will then be no unknown territory anymore. For this prospect, it was crucial to find a language family with (a) a comparatively high horizontal breadth or synchronic diversity and (b) a considerable attested history or vertical depth as a test object. The more members the family in focus has, the more contact languages its members had and have, the better the historical attestation of former states and varieties is, the better it serves for a test of a ‘new’ approach like this.

This left Indo-Iranian and Semitic as candidates. There are, of course, other languages with an early written attestation – Sumerian, Hurrian, Urartian, Elamite or Greek immediately come to mind. But they are all either too small to serve as a good test object (as is the case with Greek, as here only a few closely related dialects contrast with each other) or they have died out millennia ago (not to mention the problem of being isolates). A traceability in time over several millennia (up into the Bronze Age) combined with a good inner-familial variation is found only in Semitic and Indo-Iranian. Considering my training, Indo-Iranian was the obvious test object.<sup>19</sup> It is true that a test of the proposed parametric approach to syntactic reconstruction would be more reliable, if the database consisted of *all* Indo-Iranian languages. However, for practical purposes I decided to concentrate on Iranian only.

As regards the overall structure of this study, Chapter 2 will recapitulate the history of syntactic reconstruction and it will detail the intricacies and theoretical implications of parametric approaches to syntactic variation. Chapter 3

---

<sup>19</sup>Besides, the focus on Indo-Iranian has another advantage, as “[t]he reconstruction of the syntax of the Indo-European protolanguage simply cannot make meaningful progress without the development of a firm understanding of the Proto-Indo-Iranian situation.” (Hale 2018, p. 1940). By studying IIr. languages, one can make an important contribution to PIE syntax.

will provide a typologically useful collection of material on noun phrase properties in premodern Iranian languages. It will give a detailed description of each language's parametric values, including examples and references for potential falsification of the values listed there. This is the empirical part of this study, and the parameters will be presented in form of a parametric feature matrix in the Appendix. Chapter 4 will then provide the attempt of a parametric syntactic reconstruction. The data discussed in Chapter 3 will be used to test the idea outlined above and to discuss the results.



# Chapter 2

## Theoretical Background

### 2.1 A short outline of the research history

In this Chapter, I will explain in more detail the parametric approach I have chosen. I would like to begin by summarizing the main approaches to syntactic reconstruction which have dominated the field during the last 150 years, following the overviews presented in Ferraresi and Goldbach (2008), Viti (2015), Eythórs-son and Barðdal (2016) and Gildea et al. (2020), but interspersed with my personal thoughts. As I see it, four main directions of research and approaches to the issue can be distinguished, viz.

1. the traditional Indo-Europeanist one, whose representatives regard morphology as ‘paleo-syntax’ and reconstruct either the general morphosyntax of PIE or any other proto-language (how many cases, genders, numbers, tenses, moods etc. can be reconstructed) or morphologically cognate phrases and collocations;
2. a typological approach inspired by the works of Joseph Greenberg which dominated the research of the 1970s (quickly given up afterwards by the majority of interested people, but partly revived by Harris and Campbell 1995);

3. attempts to reconstruct syntax within the framework of ‘Radical Construction Grammar’ (J. Barðdal et al.); and
4. an approach rooted in Generative Grammar, aiming at an elaboration of the “configurational syntax” (Hale 2018) of languages.

The first of these approaches is simultaneously the oldest one, going back to the 19th century, whereas typologically inspired attempts to reconstruct syntax had their heyday in the 1970s. In contrast, both the Construction Grammar and generative approaches to syntactic reconstruction are rather young phenomena, having gained prominence only in the last 20 years. All four still have proponents; it would not be accurate to state that one approach replaced all others.

1. Serious studies on syntax and its reconstruction began with the Neogrammarians in the late 19th and early 20th centuries. Delbrück (1907) or Havers (1911) were concerned with (P)IE cases, while Delbrück (1878) and Thurneysen (1892) were with verb placement in the older IE languages and PIE. Wackernagel (1892) was a highly influential publication for research on early IE clitics, whereas Hermann (1895) raised the question of whether or not PIE had subordinate clauses. Wackernagel also addressed several (para-)syntactic aspects in his lectures (see Wackernagel 2009 for an English version), and Behaghel (1909) described what is now known as ‘Law of Increasing Members’. Modal categories were discussed by Jolly (1872) and Delbrück (1871). The issue of whether or not early or Pre-PIE had ergative-absolutive alignment instead of the attested nominative-accusative one of the older (non-Anatolian) IE languages was first brought up by Uhlenbeck (1901), later on taken up by Vaillant (1936), rejected by Rumsey (1987) and has still adherents today (e.g., de Vaan 2019, p. 178 or Willi 2018, pp. 504–540).

And yet, even though large-scale, monumental works like the Brugmann-Delbrück *Grundriß* with its three-volume section on syntax (see Delbrück 1893–1900)<sup>1</sup> or Hirt (1934, 1937) and, for that matter, also syntactic descriptions of

<sup>1</sup>Totaling a little less than 2000 pages. But note also volume II, written by Brugmann, which

single languages like Delbrück (1879, 1888), Speyer (1896) or Sommer (1931) are interesting to read, they are not comparable to contemporaneous monographs or articles on sound changes and their intricacies.

The Neogrammarians were neither able to discover syntactic analogues to sound laws nor to tackle the problem of directionality in syntactical change. They did not know how to analyze and categorize syntax – at least not in a way similar to phonology and morphology – and instead relied on a “pseudo-Sprachgefühl”, as Jamison (1993, p. 219) properly termed it. Nothing was known about possible ‘syntaxemes’, i.e., about the smallest syntactic constituents or units that distinguish one phrase, clause or sentence from another, analogous to morphemes and phonemes. Too much was vague, too much was based on intuitive insights rather than on objective criteria; criteria as were already known and widely used in the analysis, comparison and reconstruction of phonological or morphological elements.

Even the monumental Delbrück (1893–1900) reads for the most part like the earliest works on comparative IE linguistics – works the Neogrammarians, as their *geuzennaam* suggests, wanted to surpass in terms of scientific rigor. Yet with syntax, they remained on the level of their own predecessors. As M. Fritz (in Meier-Brügger 2010, p. 374) puts it,

“Anders als es seit A. Schleicher (...) bei Phonologie und Morphologie üblich ist, wird eine Syntax der uridg. Grundsprache von B. Delbrück nicht rekonstruiert, sondern es werden die einzelsprachlichen Erscheinungen nur vergleichend nebeneinandergestellt. Ob dieser ‘vorschleichersche’ Zustand in der Erforschung der Syntax bei manchen syntaktischen Fragestellungen, die größere Satzeinheiten betreffen, an sich unüberwindlich ist, wird die Zukunft weisen. Rekonstruierbar sind nur solche Satzeinheiten, die in den Einzelsprachen auf der Ausdrucks- und Inhaltsseite Vergleichbares aufweisen.”

---

covers other syntactic aspects (Brugmann 1909, 1911, 1913, 1916).

The last sentence is of pivotal importance: only morphologically and semantically cognate collocations are assumed to be reconstructible. The reason for this is the lack of syntactic theory, then and now. Since no ‘syntaxemes’ were and are known, Indo-Europeanists routinely relied and rely on compensatory morphological reconstruction. Building on morphological material, it is possible to reconstruct so-called formulae consisting of two or three words. The ‘imperishable fame’ of the introductory Chapter, PIE *\*kléuos ḡd<sup>h</sup>g<sup>uh</sup>itom*, is an excellent example, and poetic expressions, idioms, proverbs and the like may indirectly preserve earlier syntactic traits (word order vel sim.) due to an assumed conservative nature. Watkins (1995) is the most prominent work representing this morphosyntactically-based tradition within IE linguistics, but – and this point cannot be stressed enough – his approach has not much in common with actual, or rather “configurational syntax” in the sense of Hale (2018, pp. 1924–1925). It is justified quoting Clackson (2017, pp. 194–195):

“even if Watkins is right in thinking that these sentences or phrases do reflect inherited formulae, it is not clear how much of the syntactic structure of the parent language is conveyed in these reconstructed formulae, since in their transmission syntactic features, such as word order and co-ordination, have more frequently been disrupted rather than preserved (...) If the assortment of phrases in Proto-Indo-European reconstructed from this rich material is too patchy to approach anything like a correspondence set, it is unlikely that this enterprise will be any more successful in other language families.”

Similarly, morphology per se is sometimes analyzed as “paleosyntax” (Balles 2008). The reasoning here is that earlier syntactic features can be inferred by analyzing and understanding morphology as grammaticalized and hence fossilized earlier syntax. For instance, the fact that a morphological nominative and an accusative case can be reconstructed for PIE means that PIE must have been a nominative-accusative language, and not one with an ergative alignment system (but note the ongoing discussion on this particular topic, as indicated above).

Larger research objects, such as clauses or whole sentences, are in any case not analyzable by means of a morphology-based comparison, as admitted by Fritz (l. c.). Scholars like Delbrück, Brugmann, Wackernagel, Sommer etc. all knew this and consequently kept working on mere juxtapositions of syntactic phenomena of ancient Indo-European languages, aiming at an exhaustive description of the morphosyntax of ancient IE languages. Many Indo-Europeanists of the 20th and 21st century followed and still follow this largely descriptive agenda,<sup>2</sup> and there is, as I should clearly state, nothing wrong with this. A description of very general morphosyntactic features is, however, a different task from reconstructing syntax in the narrower sense. Hale (2018, pp. 1924–1925) consequently calls this longstanding Indo-Europeanist agenda “traditional syntax” and contrasts it with more modern, generative approaches designated as “configurational syntax”.

One consequence of the failure to postulate or work out formal units suitable for a proper analysis of syntax – structural analogues to phonemes – is a highly critical attitude towards syntactic reconstruction among Indo-Europeanists, as expressed by Viti (2014, p. 75):

“it is perhaps appropriate to revise these traditional assumptions about linguistic reconstruction and language change, at least in regard to certain linguistic domains such as syntax. Our claim is that syntactic anomalies are a manifestation of syntactic variation, which is the natural condition for a language. Heterogeneous syntactic structures may spontaneously emerge and be recreated anew, and they also have a lively function synchronically.”

Syntactic reconstruction in Viti’s view consists of a determination of the limits of syntactic variation of a proto-language, because linguistic reconstruction as we know it from other domains is allegedly not possible with syntax. However, elucidating variation does not resemble the clear-cut morphological or phonological reconstructions we arrive at by analyzing phonemes and their

---

<sup>2</sup>Hettrich (1988), besides Hettrich (1984, 2007), immediately comes to mind, but West (2011), Bauer (2014), Devine and Stephens (2000) can be mentioned as well.



precise ordering and appearance in specific lexemes and morphemes. Irrespective of information structure and processes like fronting etc., synchronic syntactic variation, or even seemingly chaos, may be perfectly explicable as soon as the historical dimension is taken into account. What we perceive as fronting or left-dislocation may turn out to be something completely different – syntactic variation can, but need not be related to pragmatic issues. The task for us is to focus on the underlying building blocks creating the linguistic surface variation, no matter whether morphological, lexical or syntactical in nature. Not the multitude of ‘alloys’ is interesting, but the elements they consist of and their exact constitution, in chemistry as in historical-comparative linguistics.

I cannot see any progress beyond ‘pre-Schleicherian language comparison’ (as per Fritz, quoted above), if we focus on surface syntactic variation instead of trying to find structural syntactic analogues to phonemes. Unless syntax is phonemicized, broken down into its smallest units, syntactic reconstruction will not be possible. And that is the reason why the few Indo-Europeanists who have ventured into the issue of syntactic reconstruction always failed – judged on our discipline’s standard – and why, up till now, no syntactic laws have been discovered or postulated (apart from Wackernagel’s and Behaghel’s so-called ones). Without the notion of phonemes, we could likewise only delimit the amount of surface variation in terms of pitch, timbre, tone, or in general the physical properties of speech found in utterances of related languages, but no actual reconstruction could be attempted. Linguistic reconstruction rests on a previous theoretical atomization and abstraction of the objects to be reconstructed.

2. The second approach is associated with the name of Joseph Greenberg. In the 1960s and 1970s, inspired by a renewed interest in typology, syntactic reconstruction appeared to be possible in an objective, neutral manner. Higher-order typological universals and cross-linguistic tendencies, based on statistical trends, were now the main focus of research (cf. Greenberg 1966a,b,c). As is well known, Greenberg was convinced that there were correlations between different

patterns concerning the order of elements (e.g., genitives and nouns or adpositions and nouns). Greenberg regarded this type of correlation an implicational universal (with ‘implicational’ here meaning: ‘if A, then B’), and assumed that such universals were unexceptional and unilateral.

Though his research agenda centered on synchronic and typologically utilizable generalizations, he was also aware of the fact that his approach could be used to search for other correlations between synchronic patterns and diachronic pathways (cf. Greenberg 1966a, p. 10). A few scholars adopted his idea to search for correlations between superficial word order patterns, others regarded this as too simplistic to explain syntactic differences between different languages. J. N. Adams (1976), Dressler (1971), Friedrich (1975), Lehmann (1974, 1976), Miller (1975), and Watkins (1964, 1976) all were stimulated (affirmatively or critically) by this line of research, to name but a few (see also Li 1973).

Lehmann (1973, 1974), for instance, focused on word-order changes, assuming that Greenberg’s universals might reflect possible constraints on syntactic changes. Following Lehmann, all languages tend to evolve into fully regular or consistent types; during transitions, however, or maybe at the beginning, languages may also be inconsistent. This is exactly what Lehmann (1974) reconstructed for PIE, as he labeled it an inconsistent SOV language, licensing also other orders like SVO in some contexts. This inconsistency, together with an assumed permanent impulse to regularize patterns, is supposed to have led into different directions, and the Indo-European daughter branches and languages not being SOV languages anymore must be the result of diachronic regularizations. In a similar (though not identical) vein, even Watkins once reconstructed the orders #V(E)...# and #.(E)...V# for Proto-Indo-European, with the elements E (enclitic pronouns) and V (finite verb forms),<sup>3</sup> cf. Watkins (1964).

But there was also early criticism on this focus on implicational universals and word order patterns. For instance, Friedrich (1975, p. 3) noted that “the network of variables (...) cannot be mapped onto a two-dimensional space”. Trivial

---

<sup>3</sup>Note that in the latter case the clitic is attached to some other element, marked by a dot.

Object–Verb or Verb–Object orders were not detailed enough to cover all differences between the languages of the world, and syntax evidently (or not?) implies a non-horizontal understanding and analysis. Watkins, too, distanced himself from the typological agenda he earlier endorsed, stating that too many linguists were focusing on “the magic letters S, V and O” (Watkins 1976, p. 305). His 1976 article was highly influential; it is no exaggeration to state that Watkins’ critical verdict stopped any attempt to reconstruct syntax – not only concerning word order patterns, but in general. For roughly two decades research nearly completely stopped.<sup>4</sup> Scholars diverged into different directions, and Watkins developed his own morphology-driven focus on formulae, culminating in his opus magnum of 1995.

The Greenbergian impact on syntactic reconstruction vanished as quickly as it had come.<sup>5</sup> A more detailed theoretical underpinning was necessary to analyze cross-linguistic statistical tendencies (what was then called ‘universals’),<sup>6</sup> and to speculate on previous word order patterns without accompanying (or rather preliminary) syntactic theory was, at best, premature. Nevertheless, the idea to focus on cross-linguistic recurring patterns was a step in the right direction, as these empirical facts helped theoretical linguists to develop a better understanding of syntax.

Typological attempts to syntactic reconstruction were partly reinvigorated by Harris and Campbell (1995), later taken up and complemented by Harris (2008). However, most linguists interested in the topic rather endorse one of the two approaches now to be presented. Nonetheless, Harris and Campbell (1995) inspired a younger generation of linguists to rethink the reconstructibility of syntax per se, and that was worth a lot after Watkins’ disservice to the field of diachronic syntax.

---

<sup>4</sup>One of the few exceptions from the time frame mid-70s–mid-90s is Harris (1985).

<sup>5</sup>Though typological work focusing on word order patterns has never stopped, cf. e.g., the *World Atlas of Language Structures* (Dryer and Haspelmath 2013), Dryer (1992, 2013, 2019) or Claudi (1994).

<sup>6</sup>Meanwhile, the notion of typological universals in a rigid sense has been replaced by the much more neutral idea of statistical tendencies.

3. The first is rooted in ‘Construction Grammar’ and has gained some prominence within the last decade (Barðdal 2013, 2014; Barðdal et al. 2013; Barðdal and Eythórsson 2012, 2020; Barðdal and Gildea 2015; Barðdal and Smitherman 2013; Gildea 2000; Gildea et al. 2020 etc.). Construction Grammar assumes, or rather posits, that vast quantities of language-specific so-called ‘constructions’, i.e., form-meaning pairings, are the fundamental building blocks every natural human language is built around. These constructions are culturally transmitted and can consist of almost everything, from single words to idioms and proverbs to extremely specific constructions or whole sentences.

Linguists engaged in syntactic reconstruction within a broader Construction Grammar frame – i.e., the research group around J. Barðdal and S. Gildea – endorse the so-called “Radical Construction Grammar” developed by William Croft (see Croft 2001). Since Radical Construction Grammar completely rejects the notion of syntactic relations and replaces them with semantic ones, syntactic reconstruction (or linguistic reconstruction in general) and Radical Construction Grammar seem to exclude each other. Meaning in reconstructed items is always open to a certain degree of subjectivity, as discussed in Chapter 1, and thus sharply contrasts with the objectively reconstructible formal ‘skeleton’ of lexemes and morphemes. To make a conscious decision for an approach specifically focusing on the inherently vague semantic part of variation among natural human languages is difficult to understand, at least for me, as the target of the reconstruction is syntax, not semantics, and thus a purely formal set of rules. Any application of the Comparative Method depends on formal units, not semantics. Furthermore,

“[m]orphemes and words are memorized and entered into the mental lexicon during language acquisition, but phrases, clauses, and sentences are not (except for fixed idioms). In the course of ordinary conversation, sentences are produced that have never been uttered before; clearly, a memorized body of sentences, no matter how extensive, would be of little use on its own in producing novel ones. (...) The rules constituting the syntax are deduced from the sentences

that are heard during language acquisition in childhood.” (Fortson 2010, p. 152).

The eponymous ‘constructions’ of Radical Construction Grammar do not describe rules, but lexicalized and fixed expressions. A set of lexicalized constructions is, however, not the same as a set of syntactic rules, and hence any attempt to reconstruct ‘constructions’ is not congruent with reconstructing syntax. At best, one can reconstruct a very concrete, lexicalized case of an application of not further defined actual syntactic rules, but not the complete set of such rules. In this regard, the efforts of J. Barðdal and colleagues rather resemble the formula-centered approach of Watkins.

The other reason why constructionalist approaches to syntactic reconstruction – or specifically those within Radical Construction Grammar – should be assessed critically is the framework’s denial of a finite set of cross-linguistically comparable units. If one assumes a basically infinite number of ‘constructions’, if one lays the focus of attention on the meaning of highly language- and culture-specific constructions, idioms, proverbs and the like (e.g., ‘woe’ + ‘to be’ + noun in the dative case), the Comparative Method cannot be applied, *contra* Barðdal and Eythórsson (2012). The CM needs analogues to phonemes which are finite in number, have discrete values and can (in theory) be arbitrarily set in every language, thus being purely formal, comparable units. I cannot see how ‘constructions’ could possibly fulfill these criteria, as they are the external expression of many (in theory even unlimited?) cognitive concepts or memes. Whereas classical reconstruction relies on small building blocks whose synchronically arbitrary combinations result in the observable variation of the world’s languages, I cannot see how ‘constructions’ could contain anything structurally similar to phonemes. It is difficult to avoid the impression that Radical Construction Grammar and an application of the CM mutually exclude each other – at least, if one is interested in syntax, and not in morphologically cognate formulae.

4. The last approach is the generative movement with its assumptions about an innate linguistic faculty, a so-called *Universal Grammar (UG)*, and its deeply formal view on language. Works like, e.g., Adger (2003), Batllori et al. (2005), Bowerman (2008), Crisma and Longobardi (2009), Faarlund (1990), Hewson and Bubenik (2006), Longobardi (2003), and Roberts (1998, 2007, 2012, 2021); Anderson (1993), Biberauer and Roberts (2005, 2015, 2016, 2017), Biberauer and Walkden (2015), Ceolin et al. (2020, 2021), Garrett (1990), Guardiano and Longobardi (2005, 2017), Hale (1987a,b, 1998, 2007, 2018), Longobardi and Guardiano (2009, 2017), Longobardi et al. (2013), Walkden (2014), and Willis (2011) are all, more or less, directly or indirectly, stimulated by Chomsky (1981) or Chomsky (1995), the two major landmarks in modern generative linguistics.

Syntax is assumed to be inherently vertically structured, and to reflect some biological language faculty, the UG. In principle, the old question of whether or not one should regard and analyze language as a response to culturally shaped communicative needs – the functionalist and relativist view – or as reflecting a genetic, pre-existing linguistic faculty which just needs some switch-setting during language acquisition – the formalist and absolute view – has not seen a definite answer so far.<sup>7</sup> In general, and due to the lack of clear empirical evidence in favor of one or the other view, it is legitimate to endorse each conception, and so also with a reconstructive agenda in mind.

An agnostic point of view is perhaps wisest, as it does not matter whether the proponents of generative grammar are right with their assumptions about some hypothetical innate endowment. What matters is that linguistic reconstruction needs a set of universal entities/units, as only then the Comparative Method can

---

<sup>7</sup>The evidence, insofar as it concerns language acquisition, is inconclusive. For instance, the existence of so-called *home signs*, i.e., individual gestural communication systems showing signs of syntax, morphology and prosody which are spontaneously developed by deaf children without linguistic input (on which see Begby 2017; Coppola and Newport 2005; Franklin et al. 2011; Goldin-Meadow, Brentari, et al. 2015; Goldin-Meadow, Namboodiripad, et al. 2015 or Wood et al. 2018) can be accounted for by both functional needs and/or a common genetic language faculty. I am not sure whether there will ever be a possibility for fact-based decision making in favor of the innate endowment vs. functional hypothesis.

be applied. Whether one assumes that all humans share some common genetic language faculty – irrespective of how exactly this may be encoded genetically and of how this faculty can account for the surface variation among the languages of the world – or whether all human languages are the result of common communicative needs resulting in (surprisingly often) similar types of communication systems/languages, is irrelevant. Relevant is the fact that we can analyze languages in the same terms, namely as being built around basic formal units.

But what are these basic units? In terms of the phonetic or rather phonological aspect of human languages, the universal comparability of all natural human languages can be explained by anatomic limitations (the human vocal tract allows for a certain set of sounds humans can produce). With syntax, things are different, as there is no obvious anatomical constraint shaping utterances. Syntactic limitations are, perhaps, determined by the basic processes of communication and/or cognition.

A prototypical message – at least a sufficiently complex one, beyond the level of interjections or purely deictic ones etc. – involves at least a prototypical action (= verb/predicate) and one or more prototypical arguments of this verb (in any case a subject/agent, and optionally one or more objects denoting benefactive or patient etc.). But how exactly these constituents interact and how they are realized in utterances, is open to discrete choices: Are there separate words for action and arguments or is there just one (polysynthesis)? If there are at least two separate words: Is the action (= verb/predicate) uttered first or last? Does the subject/agent occur sentence-initially or finally, preceding or following an object or the action/verb? Are there formal differences between verbs and nouns at all?

The idea that the range of syntactical possibilities may be delimited by a basic or ‘base-generated’ sentence structure implies purely logical constraints instead of anatomical limitations. Syntax, then, could vary among different speech communities as the result of logical choices, logical splits concerning these basic constituents. These constituents can be further segmented according to whether or not certain abstract features like person, number or gender are marked or not,

and if so, in what form etc. This results in a tree of logical choices, of proto-nodes which develop ever-finer logical ramifications. Eventually, this results in a set of syntactic segments built upon purely logical, discrete choices which can (but need not) be set by every natural human language in a synchronically arbitrary way.

These choices are called *parameters* in generative grammar, though the question of whether the primitive base-generated sentence may be the result of a genetically encoded language faculty or convergent evolution in terms of communicative needs is irrelevant for anyone interested in applying the Comparative Method to syntax. The result is the same: a set of logical, discrete choices according to which every human language can be classified. Generative grammar assumes that the base-generated sentence structure is the result of some innate endowment, and there are indeed certain observations indicating that this may be the case;<sup>8</sup> but for comparative and reconstructive purposes it does not matter, and one can employ parameters also with a functionalist mindset.

If one either intends to cross-linguistically compare and classify (typology) or to reconstruct syntactic aspects of a particular language family (historical-comparative linguistics), one will have no other choice than to rely on comparable units. Cross-linguistically comparable units, in turn, must be present in all natural human languages, must be discrete and finite in number. Whether one ascribes them to an innate ‘Universal Grammar’ or to a more complex interplay of genetically encoded abilities, mental processes and functional needs enabling

---

<sup>8</sup>A priori, nothing prevents humans from uttering sentences with a basic OSV order. Yet this type is extremely rare as a basic sentence order among the languages of the world, as visible in databases like WALS (see Dryer and Haspelmath 2013, feature 81A). How can this be explained, if not as the result of a strong cognitive barrier preventing humans from choosing such orders and hence human languages from evolving into this type? It appears to be less favored for humans to think of objects first, and this may give us a hint at the underlying mental processes, a clue to some base-generated sentence structure. Approaches arguing instead that languages and their grammatical structures are the result of culturally shaped communicative needs run into problems here, as it is not evident why there should be such a strong disfavor for object-first structures across the world. Why should that be the case? See also Riesberg et al. (2019) for a discussion of the universal agent-first tendency, particularly pp. 524–526 (with references).



a certain, but overall finite set of logical entities, of choices, of parametric settings – all that matters is that there are formal basic units that logically derive from some basic utterance and some very general cognitive principles (gender, number, person, animacy etc.).

These basic or universal categories are inspired by and inferred from typological observations. They are not invented or figments of our scholarly imagination. Categories like animacy, person, tense, mood etc. directly derive from empirical data, as does the archetypical proto-sentence, or, in other words, the unmarked basic utterance consisting of an action and its implicit (polysynthesis) or explicit agent argument. Obviously, humans think in these terms and shape their communicative systems (= languages) accordingly, even though with a great amount of variation (Is there a morphological noun-verb distinction at all? Overt or implicit subject?, etc.). Proponents of generative grammar just took these observations a step further and developed an approach to syntax based on logical ramifications of parameters.

The parameters of generative grammar may be the appropriate starting point for syntactic comparison and reconstruction. This does not mean that I endorse every view prevalent among generativists or that I adhere to generative grammar or Chomskyan linguistics in general. Quite the opposite, many tenets of this approach are highly questionable: For example, why should language change be restricted to first-language acquisition or rather inter-generational ‘catastrophic’ changes in the I-language (the central dogma of D. Lightfoot and his followers)? It is obvious that humans can change their grammar also during the course of their life; this extreme focus on ‘transmission problems’, so to say, strikes me as extremely unlikely. My adoption of parameters just means that I, as a historical linguist, need formal units in order to apply the Comparative Method, and that I make use of the fact that generativists have worked extensively on and with such units. That being said, I will now describe the status quo of research on parameters.

## 2.2 Parametric approaches to diachronic syntax

The search for parameters and higher-order principles originally started as so-called *Principles and Parameters* approach (P & P).<sup>9</sup> The pivotal publication, Chomsky's (1981) *Lectures on Government and Binding*, stimulated much research in the 1980s. The basic idea was and is that there is

“a set of open choices between presumably binary values (...) closed by each language learner on the basis of his/her environmental linguistic evidence. (...) grammar acquisition should reduce, for a substantial part, to parameter setting, and the core grammar of every natural language can in principle be represented by a string of binary symbols (...) each coding the value of a parameter in UG.” (Longobardi and Guardiano 2009, p. 1684)

Or, in other words, the classical P & P approach (not a theory, but an approach!) to syntax proposes that there is an innate set of universal *principles* shared by all natural human languages. Acquiring a particular language implies the setting of a finite number of *parameters* with smaller scope. As Chomsky (1995, p. 7) expressed, the P & P approach “(...) is in part a bold speculation rather than a specific hypothesis. Nevertheless, its basic assumptions seem reasonable (...) and they do suggest a natural way to resolve the tension between descriptive and explanatory adequacy”. In theory, morphosyntactic parameters, the “atoms” of language (Baker 2001), are characterized by discreteness, finiteness and universal occurrence or presence. These properties of Chomsky's approach indeed suggest, as proposed by Roberts (2007, 2021), to use parameters as a means to syntactic reconstruction.<sup>10</sup> Not phrases, not collocations, not sentences and not concrete utterances, but parameters, and thus abstractions. Writing in retrospect, Chomsky stated:

<sup>9</sup>See Ledgeway and Roberts (2017) and particularly Riolfi (2017).

<sup>10</sup>The basic idea as such can be found in and traced back at least to Roberts (1998, p. 396): “I submit that the principles-and-parameters model of cross-linguistic variation, as it has been developed since Chomsky (1981), is able to achieve this [= syntactic reconstruction].”

“Within the P & P approach the problems of typology and language variation arise in somewhat different form than before. Language differences and typology should be reducible to choice of values of parameters. A major research problem is to determine just what these options are, and in what components of language they are to be found.” (Chomsky 1995: 6)

Early comparative work inspired by these ideas resulted, *inter alia*, in the postulation of the null-subject or pro-drop parameter (see Rizzi 1982, especially Chapter 4), the head-directionality parameter (Hawkins 1983, Koopman 1984) or V-movement parameters (Emonds 1978, den Besten 1983 or Pollock 1989).

However, there are problems with the original idea (see, for a critique of classical P & P theory, Walkden 2014, pp. 19–23, Newmeyer 2004 or Hale 2007, 161 ff.). The most important one concerns the number and classification of parameters; in contrast to our knowledge about possible sounds human languages can consist of, classical P & P accounts failed to deliver anything like a ‘periodic system’ of parameters.<sup>11</sup> Connected to this is the apparent lack of clustering of surface phenomena of parameters predicted by the original P & P approach. It is likewise not evident how abstract features like parameters may be the result of some innate language faculty. For instance, the idea that every human being is born with the genetically encoded ability to choose whether or not subjects must be overtly expressed (the so-called pro-drop parameter) seems absurd.

The reaction to this criticism was a modification of the original idea; this modification is today known as ‘Borer-Chomsky Conjecture’.<sup>12</sup> As a result, parameters are largely decoupled from UG and are no longer conceptualized as being innate. Going back to Borer (1984, p. 29), and refined and fully articulated by Chomsky (1995), the ‘Borer-Chomsky Conjecture’ states that parametric

---

<sup>11</sup>Even though Baker (2001, p. 46) expressed his view that “linguistics (...) is ready and waiting for its Mendeleyev” – a point with which I strongly agree. Baker (2001, p. 183) himself attempted to present something like a ‘periodic system’ in form of a hierarchically organized tree.

<sup>12</sup>So termed by Baker (2008b, p. 156), as far as I know.

variation across languages is probably restricted to lexical items (more precisely: functional heads); cf. Ledgeway and Roberts (2017, p. 583):

“All parameters of variation are attributable to differences in the formal features of particular items (e.g. the functional heads) in the Lexicon.”

Formal features of heads can be, for example, categorial features (is item XY a noun, verb, etc.), abstract case features (NOM, ACC, etc.), movement-triggering features or  $\phi$ -features (person, number, gender, etc.). The great advantage as compared to the original concept lies in the fact that both synchronic and diachronic syntactic variation is rooted in the lexicon, not in the mind and genes, and that now only some very basic syntactic mechanisms, such as recursively combining two components – “merge” in minimalist terms –, are thought of as being genetically predetermined (= UG).<sup>13</sup>

Parameters are no longer assumed to be pre-specified in the innate endowment. Instances of greater or smaller parametric variation are part of the culturally transmitted and thus family-specific lexicon – synchronically chaotic, but diachronically explicable. Children acquire lexemes and they simultaneously internalize their correct, grammatical usage according to logical, binary oppositions (e.g., “XY is an adjective; it agrees with its head noun in terms of gender and number and should be placed before nouns, but after demonstratives and numerals”).

This parallels the status of phonemes: a priori, every healthy human being can learn to utter every sound occurring in any natural human language due to the presence of a vocal tract, but which concrete sounds are unconsciously abstracted as part of the native language during first language acquisition is determined by the input or rather the lexicon the child is exposed to. Parameter setting probably runs similar to this. According to the so-called ‘emergentist

---

<sup>13</sup>As are, perhaps, some very abstract generalizations which might not even be restricted to the human ability to think (e.g., we perceive our environment as consisting of changing subjects/agents and objects/patients which carry out or are subject to certain actions – the basic syntax, so to say, as described above).

view’ of Biberauer and Roberts (2017), parametric setting involves very few pre-specified components (= UG; basic mental processes), a large quantity of primary linguistic data (PLD; the lexicon) and functional aspects or those related to human cognition (optimization and economy principles; e.g., ‘stochastic’ setting of values or analogy).<sup>14</sup>

The parameters themselves are considered to be discrete choices which are the direct and entirely logical consequence of the nature of human cognition and communication. For instance, given the fact that a language has verbs (or rather prototypical words denoting actions), the speech community has a choice to encode all further information on this verb or not (what Baker calls the polysynthesis parameter). The language, or rather the speech community, can opt to reject polysynthesis on the verb (which then implies separate words denoting the verb and its arguments) and to place heads phrase- and clause-finally; further, that it has either ergative or nominative-accusative alignment.<sup>15</sup> Nothing of this is genetically encoded – there is no ergative-gene or the like –, but the information is collectively stored in the various items of the language’s lexicon, viz. in form of their functional heads, and children unconsciously abstract it.

The basic idea that humans probably think in terms of actions, agents and objects leads to the subsequent idea that human language in a general sense is built around these archetypes. All variation among languages is the result of ever-finer bifurcations of possible pathways and distinctions among verbs and their arguments. Within modern minimalist accounts of parametric variation, one thus no longer endorses the view that parameters are some exclusively innate phenomenon, but rather the result of an interplay of some very basic genetic factors (UG), cognitive mechanisms (‘feature economy’ and ‘input generalization’ in the terms of Biberauer and Roberts 2017) and culturally transmitted data (lexicon).

---

<sup>14</sup>What Biberauer and Roberts call “input generalization” and “feature economy”.

<sup>15</sup>See Baker (2001, p. 183) for the indicated hierarchy.

Syntactic change, in turn, occurs in form of small parametric changes; especially, though not exclusively, through reanalysis in syntactically (and/or pragmatically?) ambiguous structures and probably also by means of analogy. Without the assumption of parametric features being stored in the lexicon, it would be difficult to entertain the idea of parametric syntactic change at all. This was one of the weak points of the original P & P concept of innate parameters, as the lexicon can, of course, change over time (and hence parameters whose values are stored in the lexicon), whereas a set of fixed UG parameters should not change according to generative lore. If parameters are outsourced, so to say, from the mind (or the genome) to the lexicon, this will circumvent a significant problem.

Biberauer and Roberts have worked extensively on small parametric changes accumulating over time (see, among others, Biberauer and Roberts 2005, 2015, 2016, 2017; Roberts 2007, 2012, 2021). Slight parametric changes building on each other can (but need not!) result in the impression of ‘cascades of change’ or long-term ‘drift’. Working with the assumption that parametric values are encoded within the lexicon, this implies – at the very least – hundreds of different parametric values and settings<sup>16</sup> which can be classified according to their scope over the lexicon (Biberauer and Roberts 2017, p. 149):

“For a given value  $v_i$  of a parametrically variant feature F:

- a. **Macroparameters:** all heads of the relevant type, e.g. all probes, all phase heads, etc., share  $v_i$ ;<sup>17</sup>
- b. **Mesoparameters:** all heads of a given natural class, e.g. [+V] or a core functional category, share  $v_i$ ;<sup>18</sup>

<sup>16</sup>The opposite view, earlier endorsed by Baker (1996, 2001) – see also Huang (2015) – states that there are only a few principle-like macroparameters with far-reaching implications for the general structure of languages. Baker (2008a,b), however, acknowledges that there must exist many microparameters besides higher-order macroparameters.

<sup>17</sup>An example for a macroparameter is the head-directionality parameter determining whether or not a language is rigidly head-initial or head-final.

<sup>18</sup>An example for a mesoparameter is the pro-drop parameter defining whether a language must have an overt subject pronoun or whether marking on the verb is sufficient (= pronoun drop); pro-drop characterizes many Romance languages to this day.

- c. **Microparameters:** a small, lexically definable subclass of functional heads (e.g. modal auxiliaries, subject clitics) shows  $v_i$ ,<sup>19</sup>
- d. **Nanoparameters:** one or more individual lexical items is/are specified for  $v_i$ .”

This, in turn, implies parametric hierarchies and that syntactic change is more likely to affect parameters with small scope (nano-, and less likely micro-, and even less likely mesoparameters; macroparameters are expected to be diachronically rather stable, as they affect many lexical items). The more items are affected by one parameter, the likelier it is that the parametric value will remain stable over time.<sup>20</sup> Syntactic reconstruction, if carried out with parameters, can then be expected to show exactly this patterning: frequent changes at the bottom of the hierarchy, hardly any at the top.

As parameters are defined as discrete logical choices between opposite syntactic behaviors (Does a particular structure show behavior X: yes or no?), syntactic cognates may consist of individual lexical items (nanoparametric cognates) or whole domains (macroparametric cognates), or anything in between. For example, a (meso-)parameter could be posited for the possibility of languages

<sup>19</sup>Microparameters are, for example, responsible for subject-clitic systems of some Italo-Romance and Franco-Provençal dialects.

<sup>20</sup>This parallels phonological change, though on a very abstract level. As is well known, the more abstract and general a sound change is, and the less conditioned by specific contexts, the more likely it is to occur and to spread through the lexicon. But macroparameters pertain to a higher level: they should be compared with very general phonological settings (or rather distinctive features), such as: Does the language have a moderate/great/small vowel/consonant inventory? Are there distinctive back/front/tense/round etc. vowels? Is there a voicing contrast? Are there glottalized sounds, palatal, velar, uvular etc. consonants? Such features or basic settings are diachronically rather stable in a set of genetically related languages (e.g., note that the loss of a voice contrast in stops separates Tocharian from its IE sister branches – a major shift of a very general macrocategory or macrofeature like voice is very rare), whereas more specific differences, or rather the next level in the hierarchy (= meso- or microparameters) are more likely to change, i.e., individual sounds or subsets of sounds characterized by more fine-grained features than just voice or palatal, velar, uvular place of articulation – features that only pertain to a few phonemes and not most or all. This is the reason why specific vowel changes ( $*e > a$ ) or chain shifts in terms of plosives etc. (plosives  $*b, d, g >$  fricatives  $\beta, \delta, \gamma$ ) frequently occur, but the reduction of a former moderately complex vowel inventory to just one basic vowel or the complete loss of a macrocategory like voice (as in Tocharian) is extremely rare.

to place a noun before or after cardinal numerals. Would it be necessary to give examples with each cardinal showing the same behavior with different nouns (e.g., ‘three men’, ‘three trees’, ‘three books’ etc.  $\rightarrow$  NUM–N), or with the same noun and several cardinals (‘three men’, ‘four men’, ‘five men’ etc.) to determine the parametric setting of the language? *Sensu stricto* yes, but in practice a few or perhaps even one example may be enough. Parameters are by definition discrete choices which implies that a certain setting must always result in the same surface manifestations.<sup>21</sup> If we are interested in a macro- or mesoparameter, then a very great part of the language’s lexicon (all heads, or all heads of a given natural class) will be affected, so that we can safely generalize our conclusions drawn by one or two examples.

## 2.3 Noun Phrase parameters

As noun phrase parameters are the topic of this study, a word is required on the terminology and conceptualization. Since Abney (1987) at the latest it is common practice in generative studies to analyze nominal or noun phrases as actual or underlying determiner phrases, mirroring and duplicating clauses on the nominal level.<sup>22</sup> This view is very much based on detailed studies of modern (western) European languages which have developed article systems, though our growing understanding of languages which *prima facie* lack articles like Ossetic (see e.g., Erschler 2019a) suggests that a DP analysis may also be reasonable for languages without articles like Latin, Vedic or Avestan. However, for essentially traditional

---

<sup>21</sup>Note that the yes/no opposition is not equipollent, as the value ‘yes’ does not entail an either-or decision, whereas ‘no’ does. We will see several concrete examples of this in Chapter 3. Staying with the example of the main body, a parameter asking for the surface occurrence of the noun before numerals, if set to ‘yes’, will not imply that N cannot likewise occur *after* the numeral. Both surface manifestations are possible, N–Num and Num–N. If, on the other hand, the parameter is set to ‘no’, this implies that only an order Num–N is allowed.

<sup>22</sup>However, this DP-dogma is not universally agreed on, as there are a few critics. See, for a critique of a DP analysis of languages without articles, Fukui (1988), Corver (1992), Bošković (2005, 2007), Willim (2000), Baker (2003) or Bruening (2020), among others; cf. further Salzmann (2020) for an up-to-date discussion of the NP vs. DP debate.



reasons, I will nevertheless use the term ‘noun phrase’ throughout this study. Even though there may be good reasons to analyze NPs as DPs, the traditional terminology is not plainly wrong. Furthermore, there are some open questions which still need a satisfactory response<sup>23</sup> and until such is not presented, the DP hypothesis remains exactly this: a hypothesis.

NPs are easy to find even in premodern texts and at least some aspects of NPs are discussed – sometimes more exhaustive, sometimes less so – in many descriptive grammars of Iranian languages; cf. e.g., among many others, Durkin-Meisterernst (2014) on Middle Persian and Parthian, West (2011) on Old Avestan or the chapters in the *Compendium Linguarum Iranicarum* (Schmitt 1989b) and Windfuhr (2009b) on several other languages.

A focus on NPs is also justified due to the fact that lists of NP parameters already exist. Building upon recent literature on nominal categories and NP (= DP) structure within a generative and/or specific parametric tradition – see, among others, Alexiadou et al. (2007), Bernstein (1991, 2001, 2008), Corbett (1991), Ghomeshi et al. (2009), Giorgi and Longobardi (1991), Keenan and Paperno (2012), Paperno and Keenan (2012), Longobardi (2001a, 2005, 2008), Plank (2003) – a research group around G. Longobardi has compiled lists (or rather grids) of parameters relevant for the configurational syntax of noun phrases.

In more recent publications, this group has compiled a little less than 100 different parameters (Ceolin et al. 2020, 2021). Very unfortunately, however, the group has not published their questionnaire which defines the setting of these 94 parameters. Since the task of the present study was not to reinvent parametric linguistics, but to make use of what already existed – I am concerned with linguistic reconstruction, not with theoretical linguistics –, I adopted a smaller set

---

<sup>23</sup>For example, if NPs are, in fact, DPs which mirror clauses (mini clauses within clauses, in other words), how come most languages do not show two types of cases for arguments of nouns (equalling the nominative/accusative or absolutive/ergative systems) but only one (genitive)? Are there languages with a phrase-final determiner? Is it really true that, on the clausal level, only DPs can take over classical argument functions but determinerless NPs cannot? How come inherent grammatical features of the noun (e.g., gender) influence the morphological form of the determiner, but not vice versa?

of NP parameters. These were the parameters of Longobardi et al. (2013), whose definitions had been published in a publicly available questionnaire in an online appendix.<sup>24</sup>

I worked with most, though not all, parameters found there, using also the terminology – names and abbreviations – coined by Longobardi and colleagues; I only added three parameters pertaining to prototypical linker constructions (verbless relative clauses) and the *ezāfe*, because the issue is relevant for Iranian languages. In the remaining parts of this Chapter, I will describe in more detail the scope and nature of the individual parameters as well as implicational hierarchies. Table 2.1 gives a first overview of the NP features, the parameters and languages covered by my study.

Accordingly, the NP variation is covered by 53 parameters and encompasses 9 Iranian languages (3 from the Old Iranian period and 6 Middle Iranian languages). What looks a bit cryptic at first glance is in fact a highly logical system developed by Longobardi et al. For instance, the second parameter, called or abbreviated FGN, describes (or at this point: will describe) whether or not the language in focus has grammaticalized number. A priori, there is only a choice between yes/+ or no/-. According to generative grammar, however, this parameter is logically dependent on the very first one, called FGP and determining whether or not the language in focus has grammaticalized person (e.g., distinctions on anaphors between ‘I’/near, ‘you’/intermediate and ‘(s)he’/far deixis and on verbs between different persons). The second parameter determining number is dependent on a positive setting of the first parameter.<sup>25</sup> The second, in turn,

<sup>24</sup><https://benjamins.com/catalog/jhl.3.1.07lon/additional> (last accessed on 06-29-21).

<sup>25</sup>This is likely to be misunderstood. It is true that many languages, and so also Iranian ones, mark number but not person in terms of nominal morphology (e.g., Kurdish nouns inflected for plural but not for person). A priori, this may lead to the conclusion that there is no logical relationship between person and number marking. Yet this view appears to be premature. Firstly, the hierarchy presented here is based on the internal logic of generative grammar and its conceptualization of NPs/DPs, and this, in turn, was inspired by typological observations. It appears to be the case that every language which has number marking in some part of NPs also has different (near–far etc.) demonstratives as (facultative) constituents of NPs (and person marking

	Title/Description	Precondition(s)	Abbrev.	OAv.	YAv.	OP.	Kho.	Sgd.	Cho.	Bct.	Pth.	MiP.	Ved.	Gk.	Lat.
1	Gramm. Person		FGP												
2	Gramm. Number	FGP must be +	FGN												
3	Gramm. Gender	FGN +	FGG												
4	Feature spread to N	FGN +	FSN												
5	number on N (bounded nouns)	FSN +	FNN												
6	Gramm. partial definiteness		DGP												
7	Gramm. definiteness	DGP +	DGR												
8	Strong Person	FGP +, DGR +, NOD -	NSD												
9	Free null partitive Q	FNN +	DPQ												
10	Gramm. distal art.	FSN - or FNN - or DGR +	DDA												
11	Def.-checking N	DGR +	DCN												
12	Def. spread to N	DCN +, NSD -	DSN												
13	Def. on relatives	DGR +	DOR												
14	D-controlled infl. on N	FSN +	DIN												
15	Plural spread from Cardinals	FSN +	CPS												
16	Gramm. boundedness		CGB												
17	Strong article	DGR +, FNN +, CGB -	CGR												
18	Bounded-checking N	CGB +	CCN												
19	null-N-licensing article	FSN - or FNN - or DCN -, NOD + or NSD +	DNN												
20	Structured Adjective Phrases		AST												
21	Feature spread to structured APs	FSN +, AST +	FFS												
22	Feature spread to predicative APs	FGN +	FSP												
23	D-controlled infl. on adjectives	NSD -, FFS +	ADI												
24	DP over relatives		ADR												
25	RelCl. extrap.	ADR -	AER												
26	Free reduced RelCl	ADT +	ARR												
27	Adpositional Genitive		GAD												
28	Free Gen.		GFR												
29	Uniform Gen.	GFR +	GUN												
30	DP over free Gen.	GFR +, ADR +	GPR												
31	GenO	GUN must not be +	GFO												
32	Gen-feature spread to N		GFS												
33	D-checking possessives	DGR +, NSD + or CGR not +	PDC												
34	Adjectival poss.		APD												
35	Post-affix poss.	DCN +	PAP												
36	Clitic poss.		PCL												
37	N-feature spread to pron. poss.	FFS + or AST -, PAP + or PCL +	PHS												
38	N-feature spread to free Gen.	FFS +, GFR +, PHS must not be -	GSP												
39	Adjectival Gen.	APD +	AGE												
40	Poss.-checking N	GFS -	GCN												
41	Strong partial locality	FSN - or FNN +, CGR must not be +	TPL												
42	Strong locality	TPL must not be -	TSL												
43	D-checking demonstratives	FSN - or DGR +, TPL must not be -	TDC												
44	N over Demonstratives	FGP +	NDE												
45	N over Cardinals		NOC												
46	N over Ordinals	NOC -	NOO												
47	N over adjectives	NOO - or NGS -, NPP -	NOA												
48	N over GenO	GFO must not be -, NOA - or AST -	NGO												
49	N over external arguments	NGO - or (GFO -, NOA - or AST -)	NOE												
50	Definiteness on APs	DGP +, postnominal APs	DOA												
51	Verbless RelCl	FLI must not be +	VRC												
52	Facultative linker	GLI must not be +	FLI												
53	Generalized linker		GLI												

Table 2.1: Parametric Feature Matrix: NP parameters in Old and Middle Iranian languages and selected other early attested IE languages; the parameters, their abbreviations, and the logical relationships are based on Longobardi et al. (2013).

in the verbal domain). At least I do not know of a counterexample, i.e., of a language in which there is number marking in NPs, but absolutely no person marking in demonstratives and/or in the verbal domain – whereas the opposite is found, for instance, in Muskogean languages which have person, but no obligatory number or gender marking in NPs (see Mithun 1999, p. 464 or e.g. Hardy 2005, pp. 102–104 on NPs in Alabama). Languages with nominal person, but no number marking thus exist, whereas those with number but not person marking seem to be absent. The hierarchy proposed by generative grammar (number marking only if the language also has person) therefore appears to be substantiated by typological accounts. Secondly, the person marking does not have to be on nouns, but it can occur on other parts of a noun phrase, i.e., on determiners or function words. In many European languages, this role is taken over by facultative demon-

determines whether or not the third one (topic: grammaticalized gender) can be set at all. And this goes on and on, resulting in a rather intricate system of inter- and crossdependencies. See Figure 2.1 for a graphic representation of these most fundamental parameters.<sup>26</sup>

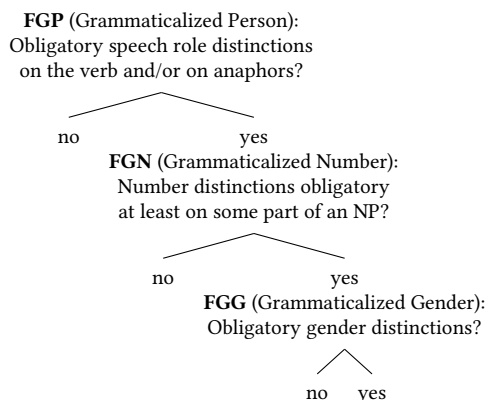


Figure 2.1: The three most basic parameters FGP, FGN and FGG.

Many other parameters logically depend on a positive setting of these parameters. As these parameters are completely self-explanatory, I will skip detailed examples, particularly as all Iranian languages have person (i.e., different anaphors and verbal person) and nominal number marking.

Gender marking presents a division among modern Iranian varieties. Some Iranian languages have preserved gender, whereas others have lost it, and the same situation is already visible in Middle Iranian times with, for example, Parthian and Middle Persian having lost gender distinctions while Khotanese

---

stratives or obligatory articles, whereas number marking either occurs also on determiners (i.e., articles) or directly on the noun. The same is also true for Iranian languages. In my sample, every language has person marking and also number marking, and the parametric hierarchy proposed by Longobardi et al. is at least not contradicted by my data. Languages like Mandarin, Japanese and Korean, on the other hand, all have negative values for FGP, FGN and FGG.

<sup>26</sup>The defining questions are abbreviated versions of the larger commentaries given by Longobardi et al. in the online appendix as per FN 24. The branching, however, is my attempt to visualize the information, particularly the logical dependencies, of the original parametric grid in such a way that parametric hierarchies become obvious.

preserved them. The individual situation will be discussed in Chapter 3 for each language of my sample.

### 2.3.1 The basic NP/DP structure

According to generative and typological studies (see Alexiadou et al. 2007, Bernstein 1991, 2001, 2008, Cinque 2005, Corbett 1991, Dryer 2018, Ghomeshi et al. 2009, Giorgi and Longobardi 1991, Keenan and Paperno 2012, Paperno and Keenan 2012, Longobardi 2001a, 2005, 2008 or Plank 2003, among others), there may be a basic or ‘base-generated’ order of nominal arguments. If correct, this may be the unmarked or default order of NPs/DPs, valid for human language(s) in general. Once again, proponents of generative grammar have worked extensively on this ‘base-generated’ order, but the idea itself does not necessarily rely on a generative mindset, as one could also advocate this idea based on typological data alone (see Cinque 2005 or Dryer 2018). The central idea is that nominal arguments, including numerals, genitives and attributive adjectives, “occupy universally fixed positions in the nominal structure with N [= the noun] parametrically taking different orders with respect to such positions” (Longobardi 2001b, p. 580). In other words: there is an unmarked, ‘base-generated’ fixed order with nouns at the right edge, and observable differences among the languages of the world are due to the option to ‘raise’ nouns (N) to the left.<sup>27</sup> This, then, can be described in terms of different parametric settings or values. The ‘base-generated’ order, with determiners (D) being leftmost, followed by two types of genitives (called *genitivus subjectivus*, GenS, and *genitivus obiectivus*, GenO), cardinal numerals (Card), ordinal numerals (Ord), several different types of adjectives (Adj, more

---

<sup>27</sup>This is the terminology of generative grammar, directly emerging out of a DP analysis with the determiner as head or highest element. Vice versa, i.e., within an NP analysis and with nouns as universal heads of nominal phrases, it would, *sensu stricto*, be incorrect to speak of leftward ‘raising’ (‘lowering’ would then be correct). A more neutral terminus could be ‘movement’ or ‘moving’, but this term is already occupied within Chomskyan linguistics/Minimalism for another process. In what follows I will adopt the generative terminology simply to avoid confusion, but the reader may note that this does not necessarily imply simultaneous acceptance of the DP analysis.

on which below), and finally the noun can be outlined as follows:

$$(D) > (\text{GenS}) > (\text{Card}) > (\text{Ord}) > (\text{Adj}) > (\text{GenO}) > (N)$$

Figure 2.2: Base-generated DP/NP structure.

Note that the different constituents of a DP/NP can, but need not be overtly expressed, including the noun (here marked by brackets); yet if several different arguments are expressed, they will often reflect certain ordering principles (e.g., determiners will occur before numerals, if they appear together, or numerals before adjectives).<sup>28</sup> This is precisely the reason why this universal or basic DP/NP order has been proposed in the literature mentioned above. In Longobardi's own words (2001b, p. 562–563), “[t]he first thing to be observed is that within DPs the principal arguments of the head noun are hierarchically ordered in a way roughly similar to that found in clauses: thematic subjects (e.g. agents) are higher than direct objects (e.g. themes) and other complements”. This is marked here by symbols/arrows pointing right, stating that everything within a DP/NP occupies a hierarchical position. The central idea of the DP hypothesis is that nominal phrases constitute structural mirrors of clauses; the contentious issue is whether one regards the determiner as head (this would be the DP analysis) or the noun (NP) and whether the hierarchy is oriented left- or rightwards.

For the purposes of this study this is irrelevant, as the parameters only require the base-generated DP/NP structure outlined above, without any explicit head. Once again, we can be agnostic with regard to the theoretical interpretation or framework. The parameters ask for features like definiteness or different genitives, and such features do not depend on the question of whether determiners are the archetypical heads of nominal phrases or rather nouns. Only the basic structure is relevant, as some parameters ask how far to the left nouns can be raised (or moved). If we can agree on the basic elements and order, this will provide a good basis for all subsequent analyses and parametric comparisons.

---

<sup>28</sup>Note that I follow Longobardi in assuming that Cardinals rank higher than Ordinals, whereas Cinque (2005, p. 328) assumes the opposite.

I am aware of the dangers of circular reasoning, but the partly hypothetical, partly typologically vindicated base-generated or universal DP/NP structure may indeed resemble sentences, if one assumes a basic SOV (subject before object, or agent before patient) order for transitive clauses; even more so, if one assumes that verbs can be raised to the left, whereas objects will usually show a strong tendency of not doing so.<sup>29</sup> NPs as mirrored clauses are theoretically appealing, and they are at least not *prima facie* disproved by the available typological data. Personally, I would tend to regard nouns as universal heads of nominal phrases, parallel to verbs in verbal phrases – mostly because of the fact that the inherent grammatical/functional properties of verbs and nouns determine the morphological marking on function verbs or nominal determiners –, but as just mentioned, this is actually a negligible point. Relevant is the fact that there may be a basic order both in the verbal and in the nominal domain.

As part of this hypothetical mirror structure of clauses, proponents of generative grammar posit two basic types of so-called ‘functional’ genitives (see the discussion in Longobardi and Silvestri 2012, pp. 93–104), viz. a ‘subject’ and ‘object’ genitive (GenS and GenO). These genitives are well known from the classical languages Latin and Greek (e.g., *amor patris* ‘[the] love of the father’ or ‘[the] love for the father’), but they also occur in modern languages like German. Generativists assume that there is a hierarchical difference in such a manner that one of these genitives is a ‘higher’ one occurring before structured adjectives (termed *genitivus subiectivus*/GenS by Longobardi) – Germanic prenominal -s would be an example – and a ‘lower’ one after structured adjectives (called *genitivus obiectivus*/GenO). The same appears to be the case with NPs: almost certainly there are (rare) examples (after all, it is a logical option to either have object-subject or subject-object), but I do not know of a language with a basic GenO–GenS order, whereas GenS–GenO orders are found, for instance, in German (with N raised to the left, thus GenS–(Adj)–N–GenO): *Marias*

<sup>29</sup>Again, I wish to refer to WALS (Dryer and Haspelmath 2013) and the conspicuous rarity of object-subject basic sentence orders (feature 81A: 564 times SOV, 488 times SVO, 95 times VSO, compared with 25 times VOS, 11 times OVS and 4 times OSV).

*sorgfältige Beschreibung Ottos* ‘Mary’s accurate description of Otto’ (Longobardi 2001b, p. 567).

There is a noticeable processing bias to favor subjects/agents over objects in basic transitive clauses (see e.g., Bornkessel-Schlesewsky and Schlewsky 2009, Bickel et al. 2015 or Riesberg et al. 2019), despite the fact that there is no logical reason to favor an S–O order over an O–S one. Subjects, and even hypothetical sub-sentence (i.e., phrasal) subjects or agents (for the bias discussed here both notions apply), tend to precede objects (S–O),<sup>30</sup> and variation in terms of word order comes in form of a movable (or ‘raisable’) tertium comparationis, either a verb or a noun. The ‘base-generated’ SOV-clausal and GenS–GenO–N orders on the NP level seem plausible to me (with the option to move N or V to any other position, resulting in SVO, VSO or GenS–N–GenO, N–GenS–GenO; for simplicity’s sake, we will ignore VOS, OVS and OSV here).<sup>31</sup>

Assuming that all human beings share the same agent-first preference, the ‘base-generated’ order of subject–object represents the unmarked and most basic verbal as well as nominal syntax. All languages will tend to fall back to this order over time, if not already evincing it, due to the processing bias associated with it. This may be the reason why there appear to be strong cross-linguistic statistical tendencies in terms of word order, indicating that object-initiality is disfavored and that subjects tend to appear to the left of objects. The most natural order seems to be one in which the verbal or nominal head (or, more neutrally, element) is placed to the right of both the subject and an object, but with the cross-linguistically frequently used option to ‘raise’ it to a position between subject and object. This results in a basic SOV or GenS–GenO–N order. Deviations from this ‘base-generated’ orders, such as S–V–O, V–S–O or GenS–N–GenO are marked in this model – even more so, if the most basic S–O order is violated (which is, after all, a logical parametric option and is thus sometimes selected by

<sup>30</sup>This extension to phrase-internal ordering mechanisms would then be an enlarged version of what Riesberg et al. (2019) call the “universal agent-first hypothesis”.

<sup>31</sup>Unfortunately, WALS does not differentiate sub-clausal/phrasal subjects and objects, and only speaks of a general genitive (feature 86A). Regarding O–S orders, see FN 27 of Chapter 4.



speech communities, but only sporadically). This theoretical concept elegantly solves a number of syntactical problems, notably in a way which seems to be empirically substantiated.

Apart from genitives, different types of adjectives have been posited by proponents of generative grammar (see Crisma 1993), viz. subject or speaker oriented adjectives (“high adjectives”), two types of manner adjectives and argument adjectives. I will not adopt this differentiation in my data sample and subsequent analyses, as it is a primarily semantics-based distinction. It is highly difficult to arrive at a clear-cut scholarly consensus on the exact meaning of words which may be hapax legomena in fragmentarily attested premodern languages like Chorasmian or Bactrian. Very often there is no consensus at all, and to base judgements about subtle parametric differences on opaque material is methodologically questionable. It is usually easy to discern adjectives as such, but I would not go any further in premodern languages with (comparatively) sparse attestation.

### 2.3.2 NPs: core and periphery, arguments and adjuncts

Current generative interpretations of NP/DP structure assume a distinction between a phrasal core (everything between D and N in the ‘base-generated’ phrase) and facultative elements to the left or right of this core, reminiscent of the difference between arguments and adjuncts of predicates. The nominal adjunct elements involve a different, non-functional type of genitive, relative clauses and non-structured adjectives which can be used instead of relative clauses (participles). The core or phrasal nucleus ends in and with the noun in the ‘base-generated’ order, but since N can be ‘raised’, this can lead to a situation in which an object genitive or structured adjectives form(s) the right edge of the nucleus. Cf. the already mentioned German NP *Marias sorgfältige Beschreibung Ottos* ‘Mary’s accurate description of Otto’ which has the order GenS–Adj–N–GenO. In cases like this the phrasal core ends with a GenO. An-

other example (see Longobardi 2001b, p. 579) would be Italian *il vestito azzurro* ‘the blue dress’ with a so-called ‘restrictive’ reading of the adjective and the order D–N–Adj.<sup>32</sup>

Proponents of generative grammar assume a different type of genitive outside of the core. Apart from two types of ‘functional’ and phrase-internal genitives, a so-called ‘free’ genitive is posited which occurs outside the actual phrase. Functional genitives (GenS and GenO) are always non-adpositional, whereas so-called free genitives can be either adpositional or inflectional.<sup>33</sup> The main difference between these genitives concerns, as their name suggests, their boundedness to structural positions: Functional ones are bound to certain archetypical positions within a DP/NP, free genitives are not.

Free genitives can either precede the determiner (if the language has an overt one) or follow the noun and/or structured adjectives; they are, in any case, phrase final (final here means both before and after the structured core as sketched above). Furthermore, they can be freely iterated, which sharply distinguishes them from functional genitives (which cannot); this recursion is, in fact, the diagnostic criterion to separate free from functional genitives. An illustrating example of recursive free genitives can be seen in “Leonardo’s famous portrait *of* Monna [sic!] Lisa *of* the Louvre Museum” (Longobardi and Silvestri 2012, p. 94). Note that in this example “Leonardo’s” is a GenS and thus a functional genitive, whereas “of Monna Lisa” and “of the Louvre Museum” are free genitives. Schematically, one could depict this as in Figure 2.3.

GenS > Adj > N ... FreeGen–FreeGen–(...)

Figure 2.3: Example for functional and free genitives.

<sup>32</sup>Compare this with an ‘appositive’ reading of the adjective which results in a different order D–Adj–N in Italian (as in other Romance languages): *l’azzurro vestito*.

<sup>33</sup>See also the appendix (p. 16 ff.) to Longobardi et al. (2013) as per footnote 24.

In general, this might be symbolized as follows:

$$(\dots) - (\text{FreeGen}) - (\text{FreeGen}) \dots (D) > (\dots) > (N) \dots (\text{FreeGen}) - (\text{FreeGen}) - (\dots)$$

Figure 2.4: Free genitives, ‘base-generated’ order.

An example from a modern Iranian language (Pashto) would be the following:

- (6) Pashto recursive free genitives (cf. Robson and Tegey 2009, p. 755)
- |           |            |           |              |           |                |           |               |             |               |            |
|-----------|------------|-----------|--------------|-----------|----------------|-----------|---------------|-------------|---------------|------------|
| <i>de</i> | <i>tor</i> | <i>de</i> | <i>tarik</i> | <i>de</i> | <i>xowunki</i> | <i>de</i> | <i>məshər</i> | <i>wror</i> | <i>xâysta</i> | <i>kor</i> |
| of        | Tor        | of        | history      | of        | teacher        | of        | older         | brother     | beautiful     | house      |
| Free Gen  |            | Free Gen  |              | Free Gen  |                | Free      | Gen           | Adj         |               | N          |
- “Tor’s history teacher’s older brother’s beautiful house.”

Here the actual NP consists of an adjective and a noun, *xâysta kor*, and several recursive free genitives at the left margin.<sup>34</sup> If the language makes use of the parametric option to ‘raise’ the noun, this will lead to a situation in which adjectives (or object genitives) constitute the phrasal edge:

$$(\dots) - (\text{FreeGen}) - (\text{FreeGen}) \dots (D) > (\dots) > (N) > (\text{Adj}) \dots (\text{FreeGen}) - (\text{FreeGen}) - (\dots)$$

Figure 2.5: Free genitives in a language with ‘raised’ N.

A parallel phrase-finality is assumed to be valid for relative clauses. Relative clauses can surface before or after the phrasal core (preceding relative clauses are frequent in older IE languages, most modern Iranian languages instead consistently place relative clauses after their head noun/NP). They are in any case assumed to be phrase-final; see Figure 2.6.

$$(\text{RelCl}) \dots (D) > (\dots) > (N) \dots (\text{RelCl})$$

Figure 2.6: Relative clauses.

This pertains also to a special type of adjectives, namely those with the functions of relative clauses (participles). They, too, can surface to the left or

<sup>34</sup>For the sake of completeness: I omitted a locative phrase *pə kəli ke* ‘in the village’ which likewise surfaces phrase-finally and even precedes the free genitives, cf. the original phrase in Robson and Tegey (2009, p. 755).

right of the phrasal core. In generative tradition, these adjectives are called non-structured adjectives, because they occur outside the structured phrasal core (see Baker and Vinokurova 2009, Cinque 2010, Kayne 1994 or Sleeman 2011).

## 2.4 Parametric hierarchies

In what follows, I will give a short description of several parameters and their characteristics – though not all, because many are self-explanatory – in the typical tree or branching format showing logical dependencies and implications perhaps a bit better than the grid. Note that all parameters are macro- or mesoparameters and thus affect a great or even the greatest part of the lexicon (micro- and nanoparameters would go beyond the scope of my study). There are parameters covering many, though not all, aspects relevant for cross-linguistic diversity in terms of NP features: of course nouns and their surface manifestations (where exactly in the NP), adjectives, genitives and possessives and their placement, definiteness and boundedness markers, relative clauses and finally *eṣāfe* parameters. With that being covered, there is an enormous amount of information available to the researcher. Whether one uses this for phylogenetic purposes or for actual linguistic reconstruction is then open to personal preferences.<sup>35</sup>

---

<sup>35</sup>Though I am skeptical whether syntactic data on their own are sufficient to shed light on phylogenetic aspects, as is the central assertion of Longobardi, Guardiano etc. For instance, Longobardi et al. (2013, pp. 134–135, 143, 146) themselves acknowledge that Farsi appears as an outlier in their sample of IE languages (see particularly Fig. 16 of the aforementioned paper). In other articles, e.g. Guardiano and Longobardi (2017, p. 15) with Fig. 16.4, this becomes even more evident, as here Farsi clusters together with non-IE languages like Basque, Turkic languages or Uralic varieties – obviously a wrong phylogenetic classification. If syntactic parameters or values of a group of genetically related languages are taken as the basis of comparison for a reconstruction of the likewise syntactic values of their last common ancestor, this will be a valid procedure in my eyes. However, any attempt at elucidating genetic relationships of languages should take into account the whole picture, i.e., phonological, morphological or lexical innovations beside syntactic aspects.

The most general parameters are those determining whether or not a given language has grammaticalized person, number and gender, i.e., whether or not it displays obligatorily

- speech role distinctions on the verb and/or on anaphors,
- number distinctions at least on some part of the DP/NP,
- gender distinctions.

The three parameters related to this have been described above in the context of the introduction of the (blank) parametric grid. I would therefore like to proceed to parameters which are relevant for word order patterns; this is what most people would probably regard as actual syntax. Building on the concept of a universal ‘base-generated’ order (with D-Num-GenS-Adj-GenO-N), different parameters define the extent of N-raising to the left. From a purely logical point of view, and with the generative account of a universal, cross-linguistic, basic DP/NP structure in mind, only the leftmost grammatical setting of a noun within a DP/NP is relevant, and every placement afterwards is then rendered irrelevant, because it is predictable.

In plain language: How far to the left can nouns be ‘raised’? If the language allows for nouns to be raised to a position preceding ordinal numbers, but not to one preceding cardinal numbers, implying N-ORD and CARD-N surface orders (Farsi/Modern Persian would be such a language), the values or settings of related ‘raising’ parameters will be predictable. Nouns which can precede ordinal numbers must also be able to precede adjectives in a particular language according to the generative model of universal DP/NP structuring. Consider Figure 2.7 for a tree-like representation of N-raising parameters.

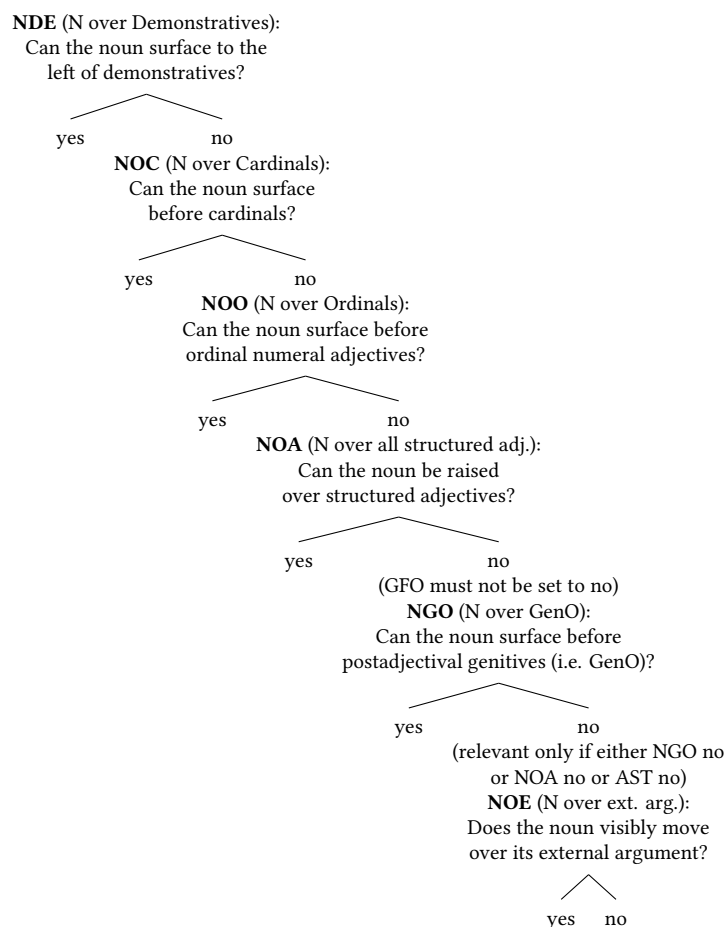


Figure 2.7: Noun-raising parameters.

Two examples (Farsi and Pashto) may illustrate this concept of a parametric hierarchy and predictability. Farsi regularly places cardinal numbers before their head nouns,<sup>36</sup> but ordinal numerals follow the noun they qualify in an *eṣāfe* con-

<sup>36</sup>Here and also in the following I will ignore the fact that Farsi makes use of numeral quantifiers, as do several other modern Iranian languages. Even though the 53 parameters of my study cover a great deal of NP-internal variation, there are many other aspects not dealt with, among which numeral quantifiers are a rather prominent example. Case marking beyond genitives in general is another point not dealt with, and the same pertains to differentiations in terms of number marking (dual, plural, paucal, ‘numeral’ etc.). The languages of the world have many more means to express syntactic differences than those 53 parametric settings I focus on.

struction like ordinary adjectives. Compare the following NPs (all from Windfuhr and Perry 2009, pp. 470–477):

- (7) Farsi NPs, extent of N-raising
- a. Demonstrative  
*ān mard*  
 that man  
 D N  
 “that man”
  - b. Cardinal  
*do kilo gušt*  
 two kilo meat  
 Card (Class) N  
 “two kg of meat”
  - c. Ordinal  
*sāl-e sevom*  
 year-EZ third  
 N Ord  
 “third year/grade”
  - d. ‘Possessor’/subject genitive  
*ketāb-e Hasan*  
 book-EZ Hasan  
 N Poss  
 “Hasan’s book”
  - e. Adjective  
*ketāb-e bozorg*  
 book-EZ big  
 N Adj  
 “big book”

The cutoff point clearly is between NPs with ordinals and cardinals. Note, however, that the N-raising parameters only determine whether or not the noun *can* be ‘raised’, i.e., placed before the mentioned argument, and that this is not

meant to be understood as an either/or decision (such as that nouns must *always* be raised over a particular argument, given a positively set parameter, and that they cannot be placed after the argument). For instance, if nouns can be raised over ordinals, the parameter NOO has a positive value, but this does not exclude the possibility that N can likewise be placed after ordinals. A positive value of NOO just means that nouns can, but need not, surface preceding ordinal numbers.

Vice versa, things are different. A negative value implies that nouns can never occur before the particular argument; this is an exclusion criterion. This situation is visible in Farsi (cf. Windfuhr and Perry 2009, p. 446). The normal pattern for NPs with ordinal numbers is an *ezāfe* construction like *sāl-e sevom* ‘third year/grade’ (with *sāl* ‘year’ being the noun), thus N–Ord with a ‘raised’ N. But there is also the possibility to build an *ezāfe* construction with reversed roles. In that case, the ordinal is substantivized as in *avval-e xordād* ‘the first of Xordād/May’ (with *avval* ‘first’ and Ord–N). Ordinals precede their head nouns also in ‘focused position’ (Windfuhr and Perry 2009, p. 475), as in *dovvom-in sāl-gard* ‘the second anniversary’ (lit. ‘second-the anniversary’) vs. ordinary *sāl-gard-e dovvom*.<sup>37</sup> Likewise, NPs with a possessor (i.e., a genitivus subjectivus) are usually *ezāfe* constructions, cf. *ketāb-e Hasan* ‘Hasan’s book’ (Windfuhr and Perry 2009, p. 474). Note, however, that the possessor can also precede its head noun in which case the noun has an anaphoric clitic as in *Hasan ketāb-aš* (lit. ‘Hasan book-his’).<sup>38</sup>

<sup>37</sup>The same applies to superlative adjectives (cf. *bozorg-tar-in māšin* ‘the largest car’ with Adj–N), and in Tajiki, the variant of Persian spoken in Tajikistan, even to other adjectives beyond superlatives. Clearly, topicalization is a relevant factor, as is the cross-linguistically well-known tendency of superlatives to act like ordinals. For this reason it would actually be advisable to differentiate between several types of adjectives (both formally and semantically), but as stated above, the situation concerning premodern languages is often not as straightforward as with Farsi. For this reason, only one category ‘adjectives’ is used in this study, with necessary shortcomings (such as the behavior of superlatives in contrast with ordinary adjectives).

<sup>38</sup>This is called ‘inversion’ in standard grammars of Persian, but from a generative point of view, this may rather be understood as a reflex of the unmarked, ‘base-generated’ order, in which case the *ezāfe* construction would entail an inversion. Thus, the Persian ‘inversion’ might be a reversion to a more natural NP order, which is apparently not one with an *ezāfe*-type order of



In essence, there is variability with ordinals, (superlative) adjectives and possessors. The normal pattern is an *ezāfe* construction (noun–argument), but ‘focus’ is possible (argument–noun). With cardinals, however, this variability is not found. Cardinals are always prenominal, and the noun cannot be ‘raised’. This means that the N-raising parameters of Farsi fit in with the generative concept of a ‘base-generated’ order. The values are all predictable, bearing the hierarchy of Figure 2.7 in mind – all arguments lower in the hierarchy than cardinal numbers usually show up after their ‘raised’ head noun, but they can also surface to its left. The parameters are all positively set in this case, whereas the ones pertaining to the behavior of nouns with cardinals and demonstratives are negatively set. Strictly speaking, the only relevant parameters are the ones with a negative value, or, in other words, the parameters which define with which arguments N-raising is *not* possible anymore, and here particularly the last one with a negative value (from a top-down perspective; or the first one from a bottom-up perspective). Every parameter further down in the hierarchy from this point on will be positively set, and every parameter further up in the hierarchy negatively. The parametric values are predictable in both directions, as long as one knows the cutoff point, and the Farsi data illustrate this perfectly.

Farsi as a right-branching language in terms of nominals (but not in the verbal/clausal domain) has its cutoff point between cardinals and ordinals. The question arises how rigidly left-branching languages will be classified and analyzed. Pashto is a good example of this type of languages. Pashto shows the following patterns (Robson and Tegey 2009, pp. 754–755):

- (8) Pashto NPs, extent of N-raising
- a. Demonstrative, adjective and free genitive
 

<i>de</i>	<i>asad agha loy kitâbuna</i>
of	Asad those big books
FreeGen	Dem Adj N
	“those big books of Asad”

---

constituents.

- b. Cardinal, adjective and two free genitives

<i>de</i>	<i>asad</i>	<i>de</i>	<i>plâr</i>	<i>lə</i>	<i>tsaloro</i>	<i>dero</i>	<i>xâysta</i>	<i>luṇo</i>
of	Asad	of	father	with	four	very	pretty	daughters
FreeGen		FreeGen		(Adp)	Card	(Adv)	Adj	N

*sara*  
with  
(Adp)  
“with Asad’s father’s four very pretty daughters”

Apparently every nominal argument (and adjunct) precedes the head noun. Note that Pashto has no inflectional, and hence also no functional genitives. Genitives are always free (adpositional) and recursive which is well illustrated by the second example (showing also the phrase-finality, as the core NP is enclosed by an ambiposition *lə... sara* ‘with’). The generative concept predicts that if nouns cannot be raised over adjectives, they will not have the possibility to be raised over higher arguments like numerals or determiners, too. This situation is met in Pashto. All N-raising parameters are negative, there are no exceptions. In principle, it would have been sufficient to determine only the behavior of nouns concerning adjectives. If N-raising is not possible here, all higher parameters must also be negative by necessity.

The original N-raising hierarchy of Longobardi and colleagues lacked a parameter defining whether or not the language can raise nouns to a position over demonstratives. The concept of a structured ‘base-generated’ NP order implies that nouns should have at least a theoretical possibility to be raisable over demonstratives, resulting in a surface order N–Dem. To my knowledge, several non-IE languages like Basque or Hebrew have this capability, but the actual reason why this gap in the system was noticeable to me was the patterning or behavior of Old Avestan. Demonstratives usually surface to the left of their NP(s), cf. e.g., Y. 36.6: *imā raocā* ‘these lights’. However, there are several exceptions, e.g., Y. 32.13; 45.3, 45.4 *aṇhōuš ahiiā vahištəm* ‘the best one in this world’ with a N–Dem order. Old Avestan thus has the option to raise the noun over demonstratives. West (2011, p. 121) ascribes this to a deictic sense of the demonstrative, but the N-raising per se seems beyond doubt.

This suggested to me that it was reasonable to include a parameter determining a language's possibility to raise nouns over demonstratives (abbreviated NDE). Concerning the hierarchy, it made sense to place this parameter at the top, over/before NOC, but the same language that was responsible for the inclusion of parameter NDE, Old Avestan, has but one example of a noun phrase with a cardinal number. If parametric values are predictable according to the hierarchies, Old Avestan must have had the option to raise nouns also over cardinal numbers (as well as any other nominal arguments). However, the only example in the corpus shows a Card–N order: Y. 44.18: *dasā aspā* 'ten mares'. Maybe this is just a transmission problem, and the (non-poetic/vernacular) language actually had the option to raise nouns over cardinal numbers.

The parametric hierarchy predicts that an N–Card order must have been legitimate in Old Avestan, but the lack of more textual material prevents us from getting clarity on the real scope of Old Avestan N-raising. An alternative view might explain the Old Avestan examples with N–Dem order as due to poetic freedom. However, I do not think that the poetic nature of the OAv. corpus or metrical considerations should lead us to believe that otherwise invalid orders could become acceptable. What is preserved in the corpus reflects the syntactical rules of Old Avestan, and if there are several examples of one (presumably hierarchically higher) order, but only one of another (presumably lower) one, this can reflect a lack of data. Either way, I included parameter NDE.

Next, I would like to discuss the parameters related to Number (see Figure 2.8). Parameters FGN and CPS should be self-explanatory, but the other parameters require some comments. Parameter FGN has a very general scope (number marking anywhere on the NP/DP), and parameters FSN (number marking on the head noun) and FNN (morphological number marking on N) further delimit the number marking. Parameter DIN, as defined by Longobardi and colleagues, basically asks whether or not all nouns, when not introduced by a definite determiner, take a particular suffix (like the nunation of Arabic), precisely because such a determiner in the D area may be missing. DPQ, in turn, asks whether or

not count/mass distinctions can be encoded by means of differential case marking (as in Finnish). Since Semitic and Uralic languages (among others showing positive values of these parameters) are or were among the contact languages of Iranian varieties, it seemed reasonable to include these parameters.

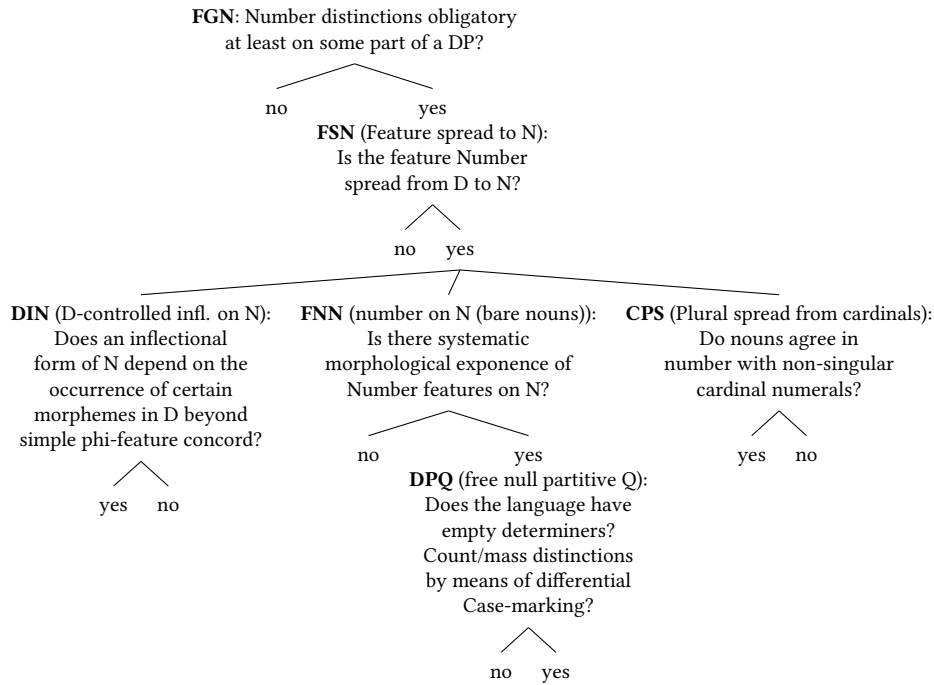


Figure 2.8: Number-related parameters.

I will illustrate these parameters and their settings with some examples from modern Iranian languages (again Farsi and Pashto). Consider the following NPs and the parametric settings they reflect:

(9) Number marking in Iranian

- a. Farsi (based on Windfuhr and Perry 2009, p. 483, but with corrected gloss)
 

*deh-hā xeyli rāh nistand*  
 village-PL much way not.be.3PL  
 “The villages are not far.”

- b. Farsi (Windfuhr and Perry 2009, p. 537)

*in dah ruz-rā*  
these ten days-DO  
“these ten days [direct object]”

- c. Pashto (own example)

*paxt-un, paxt-ânə*  
Pashtun-DIR.SG.M Pashtun-DIR.PL.M  
“The Pashtun, the Pashtuns.”

- d. Pashto (Robson and Tegey 2009, p. 769)

*tsalor wâr-a pəx-e ye*  
four all-PL.F foot-PL.F his  
“all four of his legs”

The two Farsi examples show (a) that number marking occurs (FGN → yes), (b) that number marking ‘spreads’ to the noun (from a DP-point of view, so FSN → yes), (c) that number marking is morphologically explicit (FNN → yes), (d) that a definite count noun denoting a plural requires a plural marker, but no case marker (thus parameter DPQ is set negatively), and, finally, (e) that nouns do not agree in number with non-singular cardinal numerals (thus CPS → no).

The two Pashto examples demonstrate (a) that parameters FGN, FSN and FNN are set as in Farsi (number is morphologically marked on the noun), but (b) that nouns agree in number with non-singular cardinal numerals (CPS → yes).

The parameters pertaining to number are, as a rule, not very difficult to determine. I will give one or two examples in Chapter 3 for every language of my sample in order to illustrate the individual settings. The only parameters requiring a closer look are DIN and DPQ, because they have the potential to blur an otherwise clear genealogical picture (both parameters negatively set) due to contact effects with languages (Uralic and Semitic) which have positive settings.

The next important aspect is definiteness, i.e., whether or not the language can specify nouns and if so, in what form and where exactly. Consider the parameters of Figure 2.9 (note that the bracketed comment above parameter DDA

is to be read as: The following two conditions must likewise be met, together with parameter DGR being positively set: either parameter FSN or parameter FNN must have a negative value).

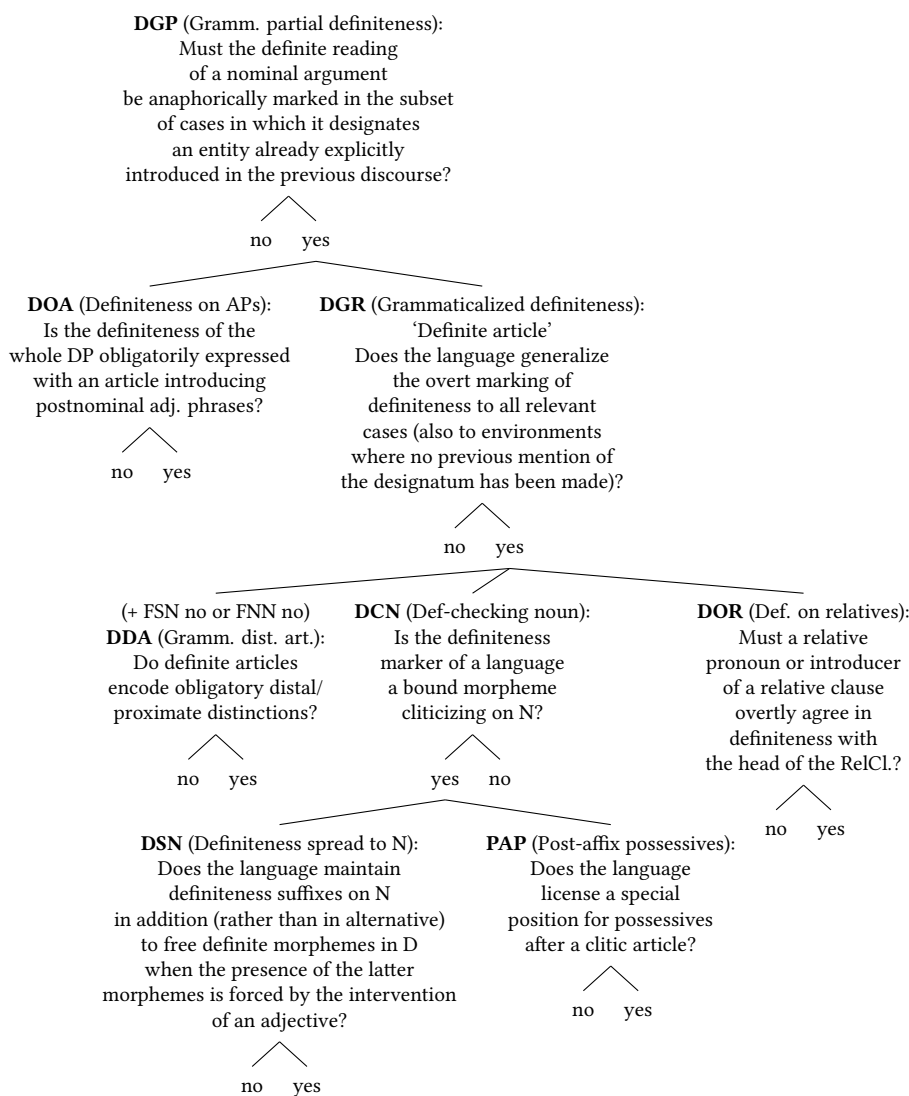


Figure 2.9: Definiteness-related parameters.

Once again, I would like to illustrate these parameters with examples from modern Iranian languages. It is of crucial importance to understand the pre-

dictive value of parametric hierarchies. Parameter DGP asks whether NPs that refer to a previously mentioned entity must always be marked for definiteness by means of anaphoric pronouns. This is not the case in Farsi, parameter DGP has a negative value. This becomes clear from examples like the following one:

- (10) Lack of obligatory anaphoric definiteness marking in Farsi (Windfuhr and Perry 2009, 537f.)

- a. *va sag-hā-ye mahalle ham dar ayyām-e, Korde be kuh zuze*  
 and dog-PL-EZ place also in days-EZ Korde be kuh howl  
*be-keš-and va ow'ow kon-and*  
 SBJV-continue.PRS.-3PL and bark SBJV.make-3PL  
 “and also if the dogs of the village howl during the days of *Korde be kuh* and bark (...)”
- b. *va agar sag-hā ow'ow kard-and*  
 and if dog-PL bark make.PRET.-3PL  
 “(...) and once the dogs barked (...)”

The dogs mentioned just one sentence earlier are not introduced by an anaphoric pronoun, even though they entail a definite reading. What does this imply? In short, that all other definiteness parameters are irrelevant. Farsi does not meet the first (or better: highest) requirement for them to be potentially relevant. Of course, speakers of Farsi *can* use anaphoric pronouns to emphasize the definiteness of some discourse topic. It is also true that Farsi has rather fine-grained specificity distinctions concerning objects (cross-linguistically, this is known as Differential Object Marking; see Bosson 1985 or e.g. Witzlack-Makarevich and Seržant 2018; for Persian, see Samvelian 2018). Consider the following examples from Windfuhr and Perry (2009, p. 486):

- *gorg-rā koštand* ‘they killed the wolf’,
- *(yek) gorg-i-rā koštand* ‘they killed a certain wolf’,
- *(yek) gorg-i koštand* ‘they killed some wolf’,
- *gorg koštand* ‘they killed wolves, were wolf-killing’,

- *gorg-hā koštand* ‘they killed many wolves’.

But this does not mean that previously mentioned entities must always be marked as definite. They can, but the anaphoric marking is facultative, not obligatory. Hence parameter DGP has a negative value. If DGP is set negatively, all other parameters further down in the hierarchy must be irrelevant according to the hierarchy or rather generative model. They could only become relevant if DGP had a positive setting. Parameters DOA, DGR, DDA, DCN, DOR, DSN and PAP are therefore not set. Definiteness parameters in general, even though this may be contrary to what our intuition tells us, can only be focused on if the logical precondition is given. I will mark this the same way as Longobardi and colleagues, viz. with a ‘0’ in the parametric feature matrix to be presented in the appendix.

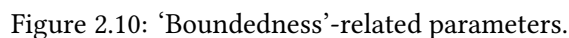
The situation in Pashto is identical. Entities mentioned earlier can, but need not be specified by means of anaphors:

- (11) Lack of obligatory anaphoric definiteness marking in Pashto (Robson and Tegey 2009, 769f.)
- a. *yaw ghaṭ wux*  
     one.M huge camel  
     “a (lit. one) huge camel”
  - b. *mo wux khlâs kəray shu*  
     we.A camel released made can.PST.PFV.3SG.M  
     “(…) we managed to free the camel.”

The English translation adds an article (*the* camel) which is lacking in the original. There is neither an article nor a demonstrative/anaphoric pronoun. Hence also in Pashto parameter DGP is set negatively, and all other definiteness parameters become irrelevant. As Longobardi and colleagues have worked out, parameter DGP is set positively in modern Romance, Germanic and Celtic languages, but not in the Slavic (with the exception of Bulgarian) and Indo-Iranian ones of their sample. This presents an interesting division among the



The same predictive capabilities of parametric hierarchies become evident with parameters related to so-called ‘boundedness’, consider Figure 2.10.



‘Boundedness’ means that languages can be classified according to whether or not bare singular count nouns in argument contexts have an unbounded denotation (CGB +) or a bounded one (CGB -). This can be illustrated by bounded ‘*John sells an apple*’ or ‘*Henri reads a book*’.

Farsi is a language in which the highest parameter CGB is set positively. The unbounded reading of nouns in Farsi, and the accompanying requirement to use an overt determiner (either *yek* or clitic *-i*, or pleonastically, both of them) to signal boundedness of the denoted noun, becomes clear from examples like the following ones:

(12) Boundedness in Farsi (Windfuhr and Perry 2009, 485f.)

- a. *sib mi-xar-am*  
apple IMPF-buy.PRS-1SG  
“I’ll buy (some) apples.”
- b. *gorg koštand*  
wolf kill.PST.3PL  
“they killed wolves, were wolf-killing.”
- c. (*yek*)                    *gorg-i koštand*  
(one/a/some) wolf-INDF kill.PST.3PL  
“they killed some [or: one] wolf.”

A bare count singular noun has an unbounded reading in Farsi (e.g., *sib* ‘apple(s)’ or *gorg* ‘wolf(es)’ of the examples above), and an overt determiner is necessary to get a bounded indefinite denotation. Parameter CCN defines whether a language with unbounded bare nouns (CGB positive) makes use of a clitic signaling indefiniteness. Farsi has such a marker in the form of *-i*, but it simultaneously also has the possibility to use *yek*, etymologically the word for ‘one’ (but also used with plurals, and thus decoupled from its numerical semantics), in the determiner position to mark a noun as indefinite. Since the clitic is the default marker, the value of parameter CCN is positive for Farsi. Again, I would like to emphasize that this parameter entails no either-or deci-

sion, but that it only asks for a clitic indefiniteness marker on the noun. It does not influence the parametric setting, whether or not additional means exist to mark indefiniteness. Since CGB was set positively, the parameters depending on a negative setting of CGB (such as parameter CGR and others further down in the hierarchy) have become irrelevant.

In Pashto, the situation is different; except for a few mass nouns (see Robson and Tegey 2009, p. 729), a reading of a singular-inflected noun as unbounded is not possible due to obligatory inflectional affixes signaling *inter alia* countability. Pashto thus has a negative setting of parameter CGB. Since CGB is set negatively, parameter CCN becomes irrelevant, and CGR instead needs to be investigated. As becomes clear from Figure 2.10, there are two preconditions which must be met for parameter CGR to become relevant: parameters DGR and FSN must be set positively, if parameter CGR is to be set. Recall that Pashto had a positive value for parameter FSN, but a negative value for DGP (obligatory anaphoric definiteness marking). A negatively set DGP, in turn, neutralized the hierarchically lower parameter DGR, and hence the logical implication is that parameter CGR cannot be set.

The parametric hierarchies with their logical structures suggest a top-down approach. The primary focus must be on the highest-ranking parameters of each branch or tree, but if possible, one may descend, step-by-step, from these highest-ranking parameters to the lowest-ranking ones. In doing so, preconditions and interdependencies must be taken into account at each step.

The next, small hierarchy is also related to definiteness parameters (see Figure 2.11). Parameter NSD basically asks whether or not the language can place referential nominal material (e.g., names) in the D-area or not. This can also occur via a surrogate, a “filler” or so-called expletive (as in Italian *Roma antica* vs. *l’antica Roma* ‘ancient Rome’). DNN, on the other hand, describes whether or not the language introduces definite nounless arguments/adjuncts (a genitive or adpositional argument or a relative clause) via a determiner identical with an article (DNN +) or by means of another determiner (e.g., demonstrative: in

such a case, the parameter is set negatively, DNN -). Both parameters depend on grammaticalized definiteness marking (DGP and DGR +).

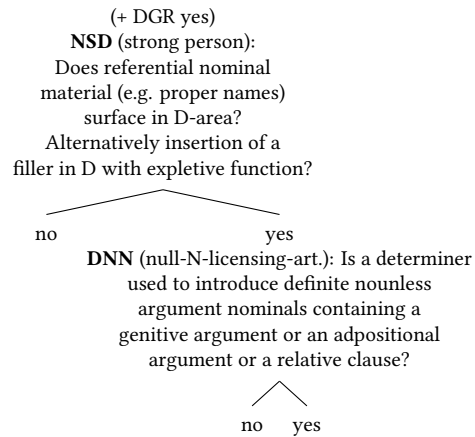


Figure 2.11: Further definiteness-related parameters.

Adjectives require a comment with regard to parameter ADI (see Figure 2.12).

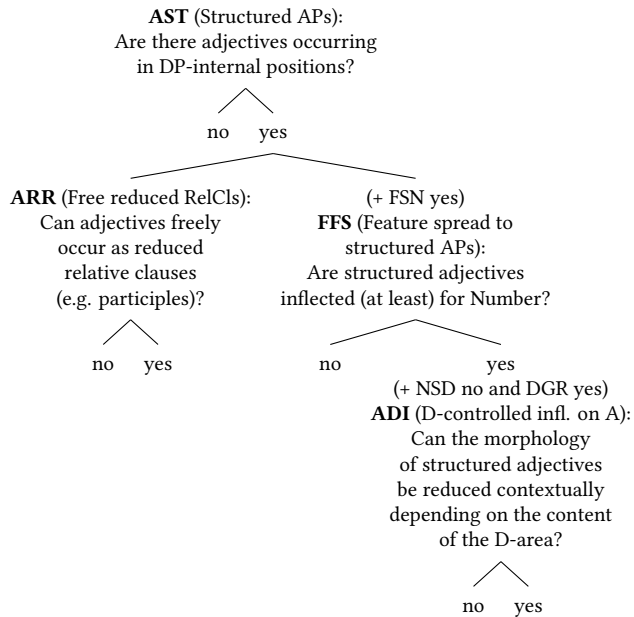


Figure 2.12: Adjectival parameters.

Recall the generative concept of a DP/NP core (defined as everything in between ‘base-generated’ determiner and noun) and periphery (everything outside), and note that adjectives inside the core (‘arguments’, so to say) are called ‘structured’ for this reason: they surface inside the DP/NP structure. Adjectives outside the core (‘adjuncts’, in other words) can occur as participles not bound to any fixed positions, and there is one parameter defining whether this is possible (ARR). Parameter FFS should be self-explanatory, only parameter ADI might be unclear. ADI states whether or not the morphology of structured adjectives can be reduced in special contexts. This is, for instance, the case with the Germanic weak inflection.

There should be no difficulties in defining the values of basic adjectival parameters. Another parametric hierarchy pertaining to adjectives is given in Figure 2.13. Parameter APO defines whether or not possessives resemble (or rather behave like) adjectives. If so, parameter AGE must be set. It relates to the presence or mere possibility of adjectival genitives, i.e., it distinguishes languages which can use adjectives in the place of genitives from those that cannot do this. Both Farsi and Pashto have negative values for APO, and hence AGE, depending on a positive setting of APO, is irrelevant for these languages.

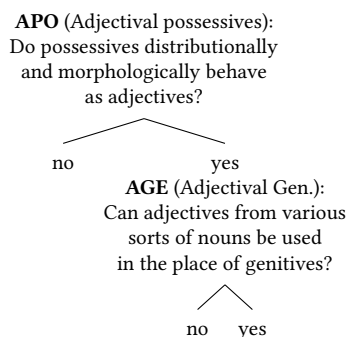


Figure 2.13: Further adjectival parameters.

Genitives within a generative framework (GenS, GenO, free Gen.) have been discussed above; consider Figure 2.14 for a parametric hierarchy related to genitives.

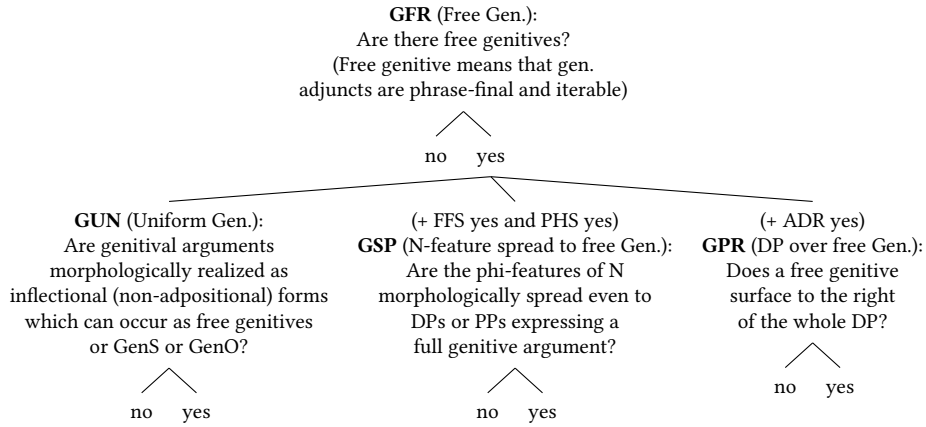


Figure 2.14: Genitival parameters.

There are two positions for structured genitives inside the ‘base-generated’ DP/NP core, to which must be added a free and recursive genitive outside of this core. However, Longobardi et al. assume that there is a different type of genitive present in old Indo-European languages, a so-called ‘uniform’ genitive. All genitival arguments, free and structured ones, fall together in just one category in languages like Latin or Ancient Greek. ‘Uniform’ genitives can surface in each of the structural configurations for the arguments of nominal heads, inside or outside the core, and they are always morphologically realized as inflectional (i.e., non-adpositional) forms. Furthermore, they unite the features of otherwise different genitives (that is, different in the generative concept): free genitives are recursive, phrase-final and usually adpositional, whereas uniform genitives can surface between a determiner and an adjective, and they are potentially recursive. Three examples from Old Avestan may illustrate this:

- (13) Old Avestan, uniform genitives (see also West 2011, p. 40)

- a. Placement between D and N, recursivity (Y. 51.10)

*huuō      dāmōiš      drūjō      hunuš*  
 he.NOM.SG creator.GEN.SG wrong.GEN.SG son.NOM.SG  
 “He (is a) son of the creator of wrong.”

- b. Recursivity, surface manifestation on left edge of NP (Y. 45.8)

*vaṇhōuš mainiiōuš šiiaoθnahiiā uxδaxiiā-cā*

Good.GEN.SG Will.GEN.SG deed.GEN.SG utterance.GEN.SG-and

*vīduš*

knowing

“knowing (something) of the Good Will’s deed and utterance.”

- c. Recursivity, surface manifestation on right edge of NP (Y. 53.3)

*yeziiuī dugədrqm zaraθuštrahē*

youngest.NOM.SG daughter.GEN.PL Zarathushtra.GEN.SG

“youngest of the daughters of Zarathushtra.”

The parameter defining whether or not the language has several different types of genitives or just one uniform genitive is termed or abbreviated GUN (see Figure 2.14). The uniform genitive is, in principle, an enlarged or enhanced version of the free genitive; a free genitive with no restrictions on placement, blurring the adjunct/argument distinction. Parameter GSP, in turn, asks whether the phi-features of the noun are morphologically encoded also on free genitives (as is the case e.g. in Hindi, but not in Old Avestan). Parameter GPR defines the surface manifestation of free genitives, either at the right or left periphery of an NP.

Two other genitival parameters describe whether or not genitival arguments are head-marked and govern phi-feature agreement with their noun (instead of the opposite). This is a typical condition of Semitic languages, but not of Indo-European ones; nonetheless, since Semitic languages were and are contact languages of Iranian varieties, it seemed reasonable to include these parameters as well (see Figure 2.15).

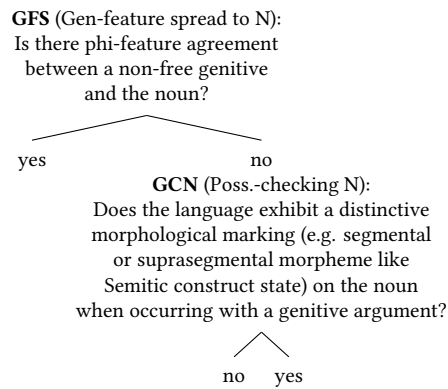


Figure 2.15: Further genitival parameters.

As regards relative clauses, two parameters define whether or not N can be raised over a relative clause (RelCl) – which would manifest itself in a surface order N–RelCl – or, if that is not normally the case (as in modern Indo-Aryan languages), if there is at least the possibility that in a subset of cases relative clauses can be extraposed to the right of the noun. Consider Figure 2.16.

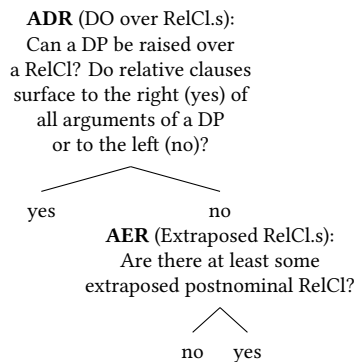


Figure 2.16: Relative clause parameters.

Possessives and the generalized linker so typical of many (though not all) modern Iranian languages are the final two subsets or hierarchies to be presented here. Figure 2.17 depicts hierarchies related to possession.



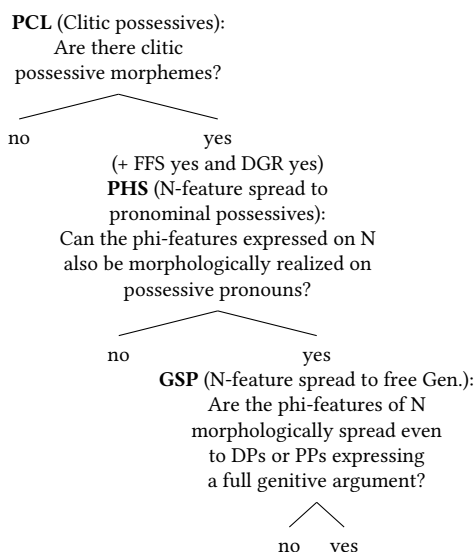


Figure 2.17: Parameters with regard to possessives.

Concerning parameters related to possessives, distinctions can be made between languages that have clitic possessive morphemes (such as Farsi, e.g., *Hasan ketāb-aš* ‘Hasan’s book’, lit. ‘Hasan book-his’) and languages that do not have such morphemes. If the relevant parameter PCL is set positively, the dependent parameter PHS can be set, defining whether or not the phi-features expressed on the noun (gender, number etc.) can also be morphologically realized on possessive pronouns. Note, however, that this parameter is only relevant in languages in which also structured adjectives morphologically encode the phi-features of N (parameter FFS must be positively set). If this is also answered in the affirmative, parameter GSP asks the same question with regard to genitives.

The last hierarchy to be presented here concerns the scope of linkers. Traditionally, such linkers are known in Iranian linguistics as *eẓāfe* (vel sim., i.e., *eẓafe*, *izafa*, *izafat* etc., from Arabic *’d’ft* ‘connection’). Several examples of Farsi using *eẓāfe* constructions to link nominal arguments to the particular head noun have already been presented in the discussion of N-raising parameters. In general, NPs constructed with linkers are typical of several western Iranian languages

like Farsi, Kurdish or Zazaki, whereas eastern Iranian languages usually do not have them.<sup>39</sup> The three parameters associated with linkers, not found in Longobardi et al. (2013), but posited by myself, define their nature and scope: VRC states whether or not a language allows for verbless relative clauses to further specify a Noun Phrase. This is, for instance, the case in Old Avestan. Such verbless relative clauses can be considered as proto-linker constructions out of which the later *ežāfe* emerged. Consider the following two examples:

(14) Verbless relative clauses ('proto-linker'), Old Avestan

a. Y. 31.6 (West 2011, p. 147)

*mąθrəm yim hauruuatātō ašahiiā*  
 spell.ACC.SG REL.ACC.SG.M integrity.GEN.SG Aša.GEN.SG  
*amərə<ta>tātas-cā*  
 immortality.GEN.SG-and  
 "a spell which (is one) of integrity, Aša and immortality."

b. Y. 35.4

*tāiš šiiaoθənāiš yāiš vahištāiš*  
 this.INS.PL deed.INS.PL REL.INS.PL best.INS.PL  
 "with those deeds that (are) the best."

It is commonly claimed that this was the archetype of the later western Iranian linker constructions (be they inflected, as in Zazaki, or not, as in Farsi), cf. Seiler (1960), Haider and Zwanziger (1984) or Yakubovich (2020). In Old Avestan, the relative pronoun, functioning as a relative connector, is still inflected, as becomes clear from the two examples. This stage can also be seen in Old Persian (though here the relative connector is an extended version of the old relative pronoun), with examples like DB. I 80–81 *xšačam taya Babirau* 'the kingship in Babylon', XPf. 30 *dārayavauš haya manā pitā* 'Darius, my father', DB. I 88–89 *avam kāram tayam Nadintabairahyā* 'that army of Nadintu-Bel'. The relative

<sup>39</sup>Though superstrate effects or rather the prestige of Persian – in the East in the form of the Tajiki or Dari variants – must always be taken into account. For instance, modern eastern Iranian languages like Waxi have the option to use clearly borrowed Persian-type *ežāfe* constructions side by side with their native NP structures.

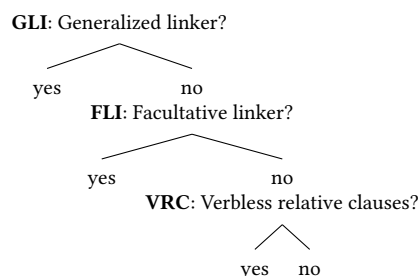
connector agrees in terms of gender, number and case with its head noun; note, however, that only linker constructions with NOM.-ACC. forms are attested in the Old Persian corpus.

The next stage in the process of grammaticalization is a loss of the inflectional capability or rather flexibility of this proto-linker or relative connector. There is only a reduced inventory of forms. For instance, the invariable form Young Avestan *yaṭ*, originally being the nom.-acc.sg. neuter form of the relative pronoun, can be used in NPs where it should not be applicable if agreement would still apply. It has become a proto-linker:

- (15) Young Avestan, facultative (proto-)linker (Yt. 5.18)

*puθrām      yaṭ      pourušaspahe*  
 son.ACC.SG.M REL.ACC.SG.N Pourušaspa.GEN.SG  
 ‘The son of Pourušaspa’.

Since this proto-linker is facultative and not obligatory, and since it still displays a reduced form of agreement in some NPs (namely those with a head noun in the nominative or accusative cases) instead of being reduced to its linker function only, it seemed reasonable to include a parameter describing this strategy. I termed it ‘facultative linker’ and differentiated it from a generalized linker not showing any signs of agreement which is used as a default strategy to link nominal arguments to their head noun (as in modern Persian). Farsi has a generalized linker, but other Modern Iranian languages like Kurdish and Zazaki probably would require a further hierarchical layer (obligatory, but inflected linker). This would result in a four-tier system of (a) verbless relative clauses (e.g., Old Avestan), (b) facultative linkers with partial inflection (Old Persian, Young Avestan), (c) obligatory linkers with (partial or general) inflection (Kurdish, Zazaki) and obligatory and (d) invariant linkers (Farsi). However, the premodern Iranian languages of my sample do not require such a fine-grained distinction. Two or perhaps three layers and the accompanying parametric categories will suffice. Consider Fig. 2.18.

Figure 2.18: Linker/*ezāfe* parameters.

In my view, there is a logical relationship between all three parameters. Since the diachronic pathway is clear in Iranian – the modern *ezāfe* is based on the old relative pronoun and verbless relative clauses – this can be reformulated in terms of a parametric hierarchy. The first question relates to the presence of a generalized linker, and if this is answered negatively, a facultative linker may still be present. Only if the particular language does not have a generalized or facultative linker, the presence of verbless relative clauses becomes relevant. This may not be true for linkers in other language families, but a parametric hierarchy building upon relative clauses makes sense for Iranian.

That Farsi indeed has a generalized linker is shown by the fact that it is not restricted to genitives, but that it can also be used to link adjectives or even clauses/nominalized phrases to a head noun, as in the following example:

- (16) Farsi, *ezāfe* construction with phrasal scope (Windfuhr and Perry 2009, p. 471)

*pesar-e az madrase gorixte*  
 boy-EZ from school flee.PTCP

“boy (which) fled from school (i.e., truant boy).”

Several standalone parameters remain which are not clearly integrated into any of the formerly mentioned hierarchies. Parameter FSP distinguishes languages in which adjectival phrases in predicative constructions agree at least in number with their governing subject noun (FSP +) from those in which they re-

main uninflected (FSP -). Modern Persian has uninflected predicative adjectives, whereas Old Persian had inflected ones:

(17) Inflection of predicative adjectives

a. Old Persian (XPh 47-48)

*šiyāta*            *ahaniy*            *jīva*            *utā marta*  
 happy.NOM.SG be.SBJV.PRS.1SG alive.NOM.SG and dead.NOM.SG  
*artāvā*            *ahaniy*  
 blessed.NOM.SG be.SBJV.PRS.1SG  
 “I want to be [or: let me be] happy (while) alive and blessed (when) dead.”

b. Farsi (Windfuhr and Perry 2009, p. 487)

*man gorosne (am)*  
 I hungry be.1SG  
 “I am hungry.”

Another parameter not discussed so far is DOA, defining definiteness on postnominal adjectival phrases. If postnominal adjectives must obligatorily be marked for definiteness by means of an introducing article, this parameter must have a positive setting (this is, for instance, the case in Classical Greek, as in *ὁ ἀνὴρ ὁ ἀγαθός* ‘the good man’, or ‘the man, the good one’). In order to become relevant at all, the language must have definiteness marking (parameter DGP must be positively set), and it must allow for postnominal adjectives, i.e., N-raising at least over adjectives.

Parameter GAD asks for the presence of an adpositional genitive in contrast with an inflected one. If a language has adpositional genitives, this parameter will have a positive value. Pashto examples in the text above demonstrated that even otherwise inflectional languages can have adpositional genitives. GAD does not contradict or render irrelevant the settings of other genitive-related parameters; it is a genuine standalone parameter. Speaking of genitives, another parameter, GFO, defines whether or not a language has an overt functional (i.e., postadjectival) *genitivus obiectivus*. Note that in many languages a GenO, if present, may surface postnominally due to N-raising as discussed above. In other

languages, namely those which have a so-called uniform genitive (the usual state for ancient Indo-European languages), this parameter becomes irrelevant, as the uniform genitive occupies structural positions otherwise reserved for object genitives. This parameter thus stands in a mutual relationship with GUN which defines the presence or absence of uniform genitives.

Parameter PAP depends on the presence of suffixed definite articles (as e.g., in Norwegian or Swedish). It requires a positive setting of DCN as a precondition, and asks whether or not possessives surface after this suffixed article. This is irrelevant for Farsi, as the necessary logical precondition is not met, but this does not mean that this is a priori also valid for other Iranian languages.

The last parameter to be discussed, PDC, defines whether or not possessives can be used as definite determiners without an additional or preceding article/determiner (e.g., English *my book* in contrast with Italian *il mio libro*). Note that PDC depends on a positive setting of DGR (full grammaticalization of definiteness).



## Chapter 3

# Noun Phrases in Iranian and beyond

### 3.1 General remarks

The following survey-like compilation of noun phrase properties in premodern Iranian languages has been compiled to the best of my knowledge. An exhaustive description of the whole intra-family variation is, of course, beyond the scope of this dissertation; it would be presumptuous to promise such. Many Iranian varieties have died out completely without leaving behind any noteworthy written testimonies, and even in those cases in which at least some written sources have overcome the centuries or millennia, either quantitative or qualitative limitations of the literary evidence (repetitive material, formulaic phrases or translations from other languages) may distort our view on the particular premodern language in focus. Sufficient accounts of syntactic properties may also be missing in the available grammatical descriptions.

Generally speaking, there was and is a surprising lack of literature on Iranian NPs within a decidedly comparative frame. Large-scale comparisons taking into account several Iranian languages like Dabir-Moghaddam (2006), Windfuhr (2009a) or Stilo (2005, 2012) are supplemented by a few others describing aspects



of two or slightly more languages like Brunner (1977), Heston (1976) or Skjærvø (2009b); apart from that, only grammatical sketches or descriptions of individual languages exist, such as Degener (1993) for Khotanese or Holmberg and Odden (2008) for modern Hawrami. In many cases, a coherent description of the noun phrase syntax of individual languages was missing, and the data had to be assembled either from a variety of papers, grammars and grammatical sketches or by means of a careful analysis of preserved (or available) texts.

Nevertheless, the following Chapter makes the attempt to describe the noun phrase variation in premodern Iranian languages in a typologically utilizable manner and to present data which can subsequently be used for reconstructive purposes. Note, once again, that aspects of “traditional syntax”, as Hale (2018) called it, are not covered by this Chapter. “Traditional syntax” aims at describing and reconstructing the morphosyntax of a language. Neither the number of cases nor of genders, neither their morphological markers nor traditional case functions (*dativus finalis*, *dativus commodi*, *dativus possessivus* etc.) will be compared and reconstructed by me.

Such morphosyntactic facts are already known, and their reconstruction for the Proto-Iranian, Proto-Indo-Iranian and Proto-Indo-European levels is largely uncontroversial.<sup>1</sup> It is common handbook knowledge to reconstruct eight or nine cases for PIIr. and PIE, three genders and three to four numbers.<sup>2</sup> The morphological markers encoding these features are safely reconstructible,<sup>3</sup> so that a mere description of the morphosyntax of various Iranian languages and yet another reconstruction of the morphosyntax of PIE or Proto-Indo-Iranian would just duplicate well-known facts. Instead, I will focus on what Hale called “config-

---

<sup>1</sup>See e.g., Mayrhofer (1989), Jügel (2017), Skjærvø (2017), Kümmel (2018), Hale (2018), Keydana (2018) or Lundquist and Yates (2018).

<sup>2</sup>Depending on whether or not one wants to reconstruct a separate directive or allative case, as evidenced by Anatolian, and whether or not a fourth number is reckoned with, namely a separate collective.

<sup>3</sup>Perhaps with the exception of the *b<sup>h</sup>-m*-cases (dative, ablative and instrumental dual and plural, and partly also singular), on which see Bonmann (2017), Lundquist and Yates (2018) or Hill (2022).

urational syntax”, i.e., on parametric values determining placement and ordering rules besides general features like the presence or absence of definiteness markers, free or functional genitives and the like, as stated in Chapter 2. I intend to describe the ‘real’ syntax, the set of rules of how to configure noun phrases, largely ignoring morphology and hence also morphosyntax.

Concerning Iranian, I will limit this survey to Old and Young Avestan, Old and Middle Persian, Parthian, Bactrian, Chorasmian, Sogdian and Khotanese, trying to accurately define their particular parametric settings. Table 3.1 lists the main grammars and grammatical descriptions relevant for these languages.

Group	Language	Glottocode	Main Sources (Grammars, Grammatical Descriptions, Gramm. Sketches)
?	Old Avestan	aves1237	West (2011), Skjærvø (2009b), Kellens (1989), Hoffmann and Forssmann (2004), Sokolov (1979b)
?	Young Avestan	aves1237	Skjærvø (2003, 2009b), Hoffmann and Forssmann (2004), Kellens (1989), Sokolov (1979b)
SW	Old Persian	oldp1254	Skjærvø (2009b), Kent (1953), Schmitt (1986, 1989a, 2008), Brandenstein and Mayrhofer (1964), Sokolov (1979a)
NE	Khotanese Saka	khot1251	Degener (1993), Emmerick (1968, 1989, 2009), Konow (1941)
NE	Sogdian	sogd1246	Yoshida (2009), Sims-Williams (1979, 1982, 1989), Gershevitch (1961)
NE	Chorasmian	khwa1238	Durkin-Meisterernst (2009), Èdel'man (2008), Humbach (1989)
SE	Bactrian	bact1239	Gholami (2014), Gholami (2011), Steblin-Kamenskij (1981)
NW	Parthian	part1239	Durkin-Meisterernst (2014), Skjærvø (2009a), Sundermann (1989b), Brunner (1977), Rastorgueva and Molčanova (1981a)
SW	Middle Persian	pahl1241	Durkin-Meisterernst (2014), Skjærvø (2009a), Sundermann (1989a), Brunner (1977), Rastorgueva and Molčanova (1981b)

Table 3.1: Premodern Iranian languages and their main grammatical descriptions.

## 3.2 Notes on the classification of Iranian and on the choice of languages

Before presenting the data for Iranian noun phrases, it may be reasonable to briefly introduce the Iranian language family and its subdivisions. Generally

speaking, research on Iranian languages can be carried out in two directions, namely once by taking Iranian as a family on its own and (for the argument's sake) ignoring external genetic relatives, and once by acknowledging the fact that Iranian is just a part of the Indo-Iranian and even larger Indo-European language families. Both approaches are inherently comparative, differing only in their scope. Concerning the issue of nominal syntax, I will try to find a holistic solution by looking at the problem from different angles, namely by focusing at first only on Iranian languages and by adopting a broader Indo-Europeanist point of view only in a second step.

That being said, several phonological, morphological and lexical innovations set the Iranian languages apart from their closest relatives, the Nuristani and Indo-Aryan languages (Mayrhofer 1989). Together, these innovations justify the postulation of a separate Iranian (sub-)branch within Indo-Iranian and the assumption of a last common proto-language of all Iranian languages: Proto-Iranian (PIr.). Traditionally, this Iranian (sub-)branch of Indo-Iranian is conceived of as a family with a rather straightforward phylogeny, namely as a tree with four branches. There is a major division between an eastern and a western group – marked primarily by shared phonological innovations<sup>4</sup> – and within each group a subdivision into a northern and southern subgroup is proposed (see Windfuhr 2009a). This results in a north-western, a south-western, a north-eastern and a south-eastern group of Iranian languages, not each of which has representatives in Old or Middle Iranian times. A traditional family tree for Iranian would therefore look like the one of Figure 3.1.

However, there are good reasons to abandon this phylogeny. Various phonological and morphological isoglosses, intersecting and blurring clear-cut divisions between groups, strongly indicate that a simplistic tree model fails to account for Iranian (see Sims-Williams 1996, Jügel 2014 or Korn 2016, 2019 for

---

<sup>4</sup>Such as PIr. initial *\*b*, *\*d*, *\*g* being retained, at least in early Middle Iranian times, in the western Iranian languages, whereas they developed into fricatives *\*β*, *\*δ* and *\*γ* in the eastern ones; or the clusters PIr. *\*-ft-* and *\*-xt-* which became voiced in eastern Iranian languages (in traditional notation, *\*-βd-*, *\*-γd-*), whereas this voicing did not occur in western languages.

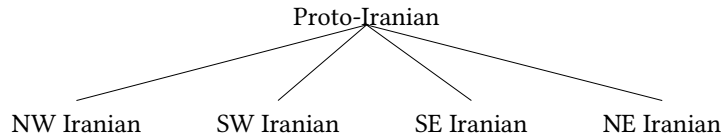


Figure 3.1: Traditional phylogeny of Iranian.

similar views) and that a quadruple division, even though in theory appealing, misses the facts. Iranian in its entirety can better be understood not as a continuum of four greater groups, but of several solitary languages (like Pashto or Balochi) and smaller subgroups (e.g., Tatic, Yidgha-Munji, Shughni-Roshani etc.), all with the same Proto-Iranian roots, but forming individual and multifariously interacting subbranches resulting in several smaller and three greater areal clusters (see also the discussion in Tremblay 2005). All of these subbranches were and are subject to intersecting phonological, morphological, lexical, and, as we will see, syntactical isoglosses. For roughly two and a half millennia, i.e., from Proto-Indo-Iranian times around 2000 BCE (Sintashta and Andronovo cultures; see Section 4.1 below) until ca. 500 CE, when intruding Turkish elements definitely broke up the former continuum and forced the last remnants of the steppe Iranians into mountainous refugia, the members of this large Iranian network were never really separated from each other.

In my view, the Iranian languages can best be described in terms of a ‘linkage’ (on which see Ross 1988 or François 2014). The ‘linkage’ concept was inspired by Austronesian languages, many of which were and are separated from each other by dozens or even hundreds of kilometers of open sea.<sup>5</sup> Nevertheless, there was contact between Austronesian languages, and the separation due to the geographical circumstances did not prevent the speech communities from sharing linguistic innovations. Secondary areal innovations independent of and superimposing themselves on earlier genetic subdivisions resulted in a set of smaller and greater networks, or rather ‘linkages’.

<sup>5</sup>Or even thousands, as is the case with the extreme outliers Malagasy and Rapanui.

The Iranian situation is not very different. Here, too, contact and mutual exchange played a major role. Iranian languages are characterized by an interplay between greater and smaller linguistic networks. Contact and hence contact phenomena arose either within the context of regular trade (most prominently the Silk Road) or through population movements that gave the continuum a good stir (for example, the Parni moving from the steppe periphery to the southwest and seizing power in Iran – forming the Arsacid Empire –, or the movements of the Balochi, Zazaki or Sangesari speech communities into their modern areas). Secondary areal features or innovations could overlap common genetic heritage and/or previous individual developments – the Pamir languages (cf. e.g. Edelman and Dodykhudoeva 2009; Morgenstierne 1973; Paxalina 1969; Payne 1989) are the prime example. The traditional tree model of historical-comparative linguistics assumes that a proto-language breaks up into several daughter languages or nodes and that these daughters evolve independently from each other. This cannot have been the case with Iranian, permanent intra-familial contact must have characterized the situation for millennia, and innovations like the change of  $*s > *h$  (on which see e.g. Mayrhofer 1989, p. 7) must have spread in the form of waves long after the breakup of Proto-Iranian.

A rather large core area reaching from Anatolia all the way to Chorasan, Transoxiana and Chorasm once constituted the central part of the Iranian network or linkage. In antiquity and well into the early Middle Ages the Iranian area extended, beyond the core territory, even further to Europe (Ukraine and the Pannonian Plain) and north-western China with the Scythians, Sarmatians and Sakas dominating the steppe corridor. Nowadays it is much smaller, but the core area, the central part of this continuum, encompassed and still encompasses what is traditionally termed north-western Iranian (Median; Parthian; Kurdish, Zazaki, Taleshi, Tati, Caspian languages, Balochi) and most of the eastern Iranian languages (Chorasmian, Sogdian; Bactrian; Pashto, most modern Pamir languages, Parachi-Ormuri). All of these languages, irrespective of many later and highly individual developments, participate in one common phonological devel-

opment: PIE  $*\hat{k}u$  and  $*\hat{g}^{(h)}u$  developed into sequences  $*-sp-$  and  $*-zb-$ , respectively (cf. Avestan *aspa-*, Sogdian  $\langle 'sp-\rangle$ , Ormuri *yâsp* ‘horse’ < PIE  $*h_1ek_{\hat{u}}o-$ ; Old Avestan *hizuuā-/hiuū-*, Parthian  $\langle 'zb'n\rangle$ , Sogdian  $\langle 'z\beta'k\rangle$ , Pashto *žeba*, Yazghulami *zveg* ‘tongue’).<sup>6</sup> This fundamentally diverges from the development found in the Sakan languages ( $\rightarrow \check{s}, \check{z}$ , as in Khot. *aśśa-* [aša-], Wakhi *yaš* ‘horse’; Khot. *biśāa-* [βižā] ‘tongue’ < seemingly PIE  $*ui\hat{g}^{(h)}ueh_2-$ )<sup>7</sup> and in the languages of the south-west, of which Persian is the prime example, which changed these sequences into simple sibilants *s* and *z* (cf. OPers. *asa-* ‘horse’ and *hi/azā-* ‘tongue’ [ $\langle h-z'-n-m\rangle$ , Acc.SG]).

Apart from this shared phonological innovation, however, the central Iranian languages appear to form several originally individual subbranches of Iranian (e.g., Proto-Balochi, Proto-Pashto, Proto-Kurdish, Proto-Parachi-Ormuri etc.) which secondarily interacted and converged in many ways. This, then, can easily mislead us to wrongly assume a common proto-language (‘Proto-Central-Iranian’ vel sim.),<sup>8</sup> whereas in reality there just was and is a bundle of overlapping innovations, a very large Sprachbund. The development of PIE  $*\hat{k}u$  and  $*\hat{g}^{(h)}u$  could be a very early innovation which spread like a wave among those languages, whereas it did not affect the already then peripheral ancestors of the later Persianide and Sakan languages.<sup>9</sup> Some innovations affected larger areas of the continuum (such as the defining western vs. eastern isogloss constituted by the change of PIr.  $*-ft-$  and  $*-xt-$  >  $*-\beta d-$ ,  $*-\gamma d-$  in eastern territories) and some smaller (e.g., Kurdish and some other languages of the north-western periphery of this core area developing *ežāfe*-morphemes inflected for case and gender), superimposing themselves on earlier genetic divisions. In my view, ‘Central’

<sup>6</sup>The PIE archetype of ‘tongue’ is difficult to reconstruct, see Mayrhofer (1992–2001), I: 591–593. These Iranian forms seem to continue a form with initial  $*h-$ , as if from PIE  $*s-$  or rather <  $*si\hat{g}^{(h)}ueh_2-$  (?).

<sup>7</sup>Note the apparent difference between the Sakan archetype with initial  $*u-$  and that of all other Iranian languages with initial  $*s-$  in this word.

<sup>8</sup>Korn (2016) subsumes all these languages under the heading ‘Central Iranian’.

<sup>9</sup>The alternative view would reckon with an early genetic split between three Iranian subbranches.

or ‘Core Iranian’ thus rather denotes a macro-area within Iranian than a phylogenetic node, very much like ‘Eastern Iranian’ is an areal rather than genetic statement (cf. Sims-Williams 1996).

Iranian languages were once spoken also on the vast Eurasian steppe. These Iranian languages did in many ways not participate in several innovations and developments circulating in the core area of the Iranian linkage, at least as far as we can judge from what scanty remains survived. Quite the opposite, the steppe Iranian languages went their own way in many aspects (most prominently perhaps by partly developing a new plural based on a feminine abstract/collective), so that it may indeed be reasonable to speak of a northern or north-eastern periphery of the continuum or linkage, as e.g., Jügel (2014) does. The only modern survivors of the steppe Iranian languages are Wakhi (cf. e.g. Bashir 2009; Grünberg and Stéblin-Kamensky 1988; Morgenstierne 1973; Steblin-Kamenskij 1999) and Ossetic (cf. e.g. Abaev 1964; Erschler 2019b; Thordarson 1989), both spoken in mountainous refugia (the Caucasus and Pamir ranges, respectively). Even though we may never know exactly the amount of diversity,<sup>10</sup> what has been preserved of the Middle Iranian representatives Khotanese and Tumshuqese, and what has survived today in the form of Wakhi and Ossetic, shows us that these steppe Iranian languages did indeed form a periphery and that any attempt to reconstruct features of Proto-Iranian must take into account these peripheral languages. Often they alone preserve something lost everywhere else within Iranian (e.g., Wakhi *maž* < PIIr. *\*maǰ<sup>h</sup>ia* ‘I.DAT.SG’). This could be similar with syntactic aspects.

The status of Chorasmian and Sogdian (and modern Yaghnobi)<sup>11</sup> is some-

---

<sup>10</sup>We should not yet give up hope that one day Scythian-Sarmatian ‘Iguvine Tablets’ or at least a ‘Tabula Bantina’ of the Graeco-Scythian colonies along the northern and north-eastern coast of the Black Sea will be unearthed. Recently, the so-called ‘unknown Kushan script’ of ancient Central Asia was deciphered by Jakob Halfmann, Natalie Korobzow and myself (Bonmann et al. 2023) and turned out to represent a previously unknown Middle Iranian variety.

<sup>11</sup>Yaghnobi, spoken in Tajikistan (see e.g. Bielmeier 1989; Xromov 1987), is a surviving modern Sogdian dialect. As such, it is not directly a steppe Iranian language, but a descendant of a peripheral core Iranian variety.

what janus-faced, as these languages on the one hand participate(d) in several typical western Iranian developments or innovations (e.g., PIE  $*\hat{k}_u$  and  $*\hat{g}^{(h)}_u$  >  $*sp$  and  $*zb$ ; or the verb  $*gaub-$  ‘to speak’ – OPers. *gauba-* ‘id.’, Sogd. <γwβ>, Cho. <γwy> ‘to praise’ –; or the compounding of a preposition and pronoun as in OPers. *hačāma*, Sogd. <c’mk>, Cho. <cm(y)k> ‘from me’), whereas they, on the other hand, also show steppe-affiliations (or, in other terms, ‘Scythian’ features like the *t*-collective/plural or the numeral ‘16’ with *r* as in Cho. <’xrδys>, Sogd. C <xwšrts>, Oss. *æxsærdæs* <  $*xšardas$  compared with Av. *xšuuazdasa*, Khot. *kšasu*, MiPers. <š’zdh> and Ved. *śóḍaśa*).<sup>12</sup> There appears to be a rather smooth transition between steppe or north-eastern/northern periphery and the Iranian languages of the core area, in line with the concept of a linkage.

The Iranian continuum was and is delimited by another periphery towards its southern or south-western end, with the Fars province at its centre (Persian and related languages). Here, too, individual developments took place which did not spread beyond the south-western boundaries (e.g., to give a phonological example, PIE  $*\hat{k}$  and  $*\hat{g}^{(h)}$  > Persian *h* and *d*). On the other hand, Persian in its different stages and guises (spoken or written, liturgical or literary) always had a major impact on other Iranian languages due to its status of being the *lingua franca* (or one among the *linguae francae*) of the *Ērān-šahr*. We thus find clearly secondary south-western or Persianide features in otherwise divergent languages like the *ežāfe* construction in Wakhi, which we need to recognize as such secondary overlaps before we reconstruct the respective feature for Proto-Iranian. South-western peripheral Iranian must always be reckoned with as a secondary overlay due to the prestige of Persian.

To sum up, Iranian is characterized by two peripheries which are today represented by Persian and other south-western varieties (Larestani, Sivandi, Lori, Bashkardi, Fars dialects, Kumzari etc.) on the one hand and Wakhi and Ossetic

<sup>12</sup>Pay attention to the difference between Ossetic (= Scythian-Sarmatian) and Khotanese (= Sakan) ‘16’ which is one of many hints indicating that also within the steppe Iranian languages there must have been a good amount of variation in terms of lexical, morphological, phonological and syntactical innovations and retentions.



(and perhaps also Yaghnobi) on the other (and in premodern times by Old and Middle Persian and Khotanese and Tumshuqese, respectively), and a great bulk of other languages in between these ends which sometimes are conservative, but more often innovative. The same is true for premodern languages. Even though we only have substantial textual remains of 3 Old Iranian languages (Old and Young Avestan as well as Old Persian) and 6 Middle Iranian ones (Khotanese, Sogdian, Chorasmian, Bactrian, Parthian, Middle Persian), these languages evince a phonological, morphological and lexical variation among each other similar to that of modern Iranian languages. A linkage of closely related and multifariously interacting sister languages, with several phonological, morphological and lexical isoglosses separating and uniting the individual languages in changing subsets or groups, suggests to include all of them in any attempt to reconstruct features of their last common proto-language.

Unfortunately, premodern languages like Tumshuqese or Sarmatian/Alanic were not sufficiently well attested to include them. There was not enough textual material to plausibly determine their parametric settings, and the same pertains to several other Old and Middle Iranian languages which undoubtedly existed, but of which next to nothing is known (like Median, Old Kurdish, Old Balochi, Old Pashto, predecessor languages of the modern Pamir languages etc.). Thus, the reconstruction of the NP parameters of Proto-Iranian to be carried out in Chapter 4 rests on 9 premodern Iranian languages, namely on those with enough data to ascertain parametric values.

In addition to the Iranian varieties, data for 3 old Indo-European languages (Vedic, Archaic Greek and Old Latin) will also be presented. They shall serve as the basis for a comparison and reconstruction beyond the Iranian level, aiming at PIIr. (Vedic) and PIE (the others). In case of doubt on the setting of parametric values, the reader is encouraged to directly study either the written testimonies themselves in the case of premodern languages and/or the cited reference works (grammars, grammatical sketches etc.). With that being said, let us start with the oldest Iranian languages first.

### 3.3 Old Iranian languages

Old Iranian languages are by definition those languages that are attested until the downfall of the Achaemenid empire as a result of Alexander's campaign. Since Proto-Iranian must have been spoken in the second millennium BCE (see Section 4.1), this is a rather large chronological period, and the three known Old Iranian languages represent, in fact, three different strata. The oldest attested language, Old Avestan, spoken in the 2nd millennium BCE, was a rough contemporary to Hittite, Palaic and Luvian in Anatolia, Mycenaean Greek in the Aegean, and the oldest Vedic represented by the *R̥gveda*, spoken in southern Central Asia and later in the Punjab. Young Avestan, spoken in the 1st half of the last millennium BCE, could be compared with and probably was spoken more or less contemporaneously to those varieties of Vedic which were preserved in form of the three younger *Samhitās* (which are dated to ca. 1000–800 BCE, see Jamison 1991, pp. 1–16). Likewise, it could be compared chronologically with Homeric Greek. Old Persian, the most recent of the three preserved Old Iranian languages, was the language of the ruling dynasty of the Achaemenid empire and can be compared with archaic Latin (or, perhaps, the younger Latin of Plautus and Terence) and classical Greek.

#### 3.3.1 Old Avestan

The extant Old Avestan (OAv.) texts – i.e., the *Ahuna Vairiia* prayer (Y. 27.13), the *Gāthās* (Y. 28–34, 43–51, 53), the *Ā Airiia ma Išiiā* or *Airiāman* prayer (Y. 54.1) and the *Yasna Haptaŋhāiti* (Y. 35.2–41) – appear to have been composed around 1500 BCE somewhere in southern Central Asia (see e.g., Degener 2002; Hintze 2015) and predate the Younger Avestan texts by several centuries. Old Avestan is the oldest Iranian language we know of and a close relative of the oldest discernible Indo-Aryan language, the roughly contemporary Vedic. In some aspects (e.g., the distinction between a genitive and locative dual), OAv. appears to be even more archaic than Vedic, and as such it seems to be the best approximation as

to what Proto-Indo-Iranian may have looked like in terms of morphology and perhaps also syntax.

In general, Old Avestan is a typical ancient Indo-European language. Noun phrases are characterized by (always) suffixal and polyfunctional inflectional markers indicating gender, number and case. As a consequence, adverbs and adpositions are, in general, facultative. *Prima facie*, word order is rather free, both on the sentence and phrasal level, tmesis and hyperbata are common, and other categories or phenomena (e.g., definiteness, animacy, differential object marking) characteristic for many modern Indo-Iranian languages are no relevant factors in the morphosyntactic system of Old Avestan. A typical example of the structure and interplay of OAv. noun phrases can be seen in the following (poetical) sentence:

(18) Old Avestan (Y. 47,6)

- a. *tā*            *dā*                            *spəntā*        *mainiiū*,  
 DEM.INS.SG lay.2SG.IND.AOR.ACT holy.INS.SG mind.INS.SG  
*mazdā*            *ahurā*,  
 Mazda.VOC.SG lord.VOC.SG  
 “Through that bounteous/holy will thou didst establish, Lord Mazdā,”
- b. *āθrā*        *vaŋhāu*        *vīdāitīm*            *rānōibiiā*  
 fire.INS.SG good.LOC.SG allocate.VN.ACC.SG contestant.DAT.DU  
 “the allocation of the good by fire to the two contestants”
- c. *ārmatoiš*                            *dəbazaŋhā*                            *ašaxiiācā*.  
 righteousness.GEN.SG reinforcement.INS.SG Aša.GEN.SG-and.  
 “with the reinforcement of righteousness and Aša/right.”

This single sentence (cf. also West 2011, pp. 18–19) contains just one verb, namely *dā*, but four basic NPs (the addressee in the vocative; a first instrumental; the direct object – a verbal noun – in the accusative; and a second instrumental), with the direct object *vī.dāitīm* and the final instrumental *dəbazaŋhā* having further dependent NPs each (genitives, a dative and an instrumental).

As indicated in the preceding section, I do not intend to reconstruct the morphosyntax of Proto-(Indo-)Iranian. If this were my goal, the statement that Old

Avestan had eight cases, three genders and three (or four) numbers – and that Proto-(Indo-)Iranian may not have been much different in this regard – could suffice. But morphosyntax is not syntax in the generative sense; and the statement that Old Avestan had a rather free word order, as is common practice to state, is theoretically underspecified. Word order is just the superficial representation of underlying syntactic rules, and these rules can be formalized in parametric terms. I am sure that detailed studies of the ‘configurational syntax’ of old Indo-European languages will reveal that the superficially free word order of Old Avestan, Vedic or Ancient Greek is (or was) in fact a rather rigid system of parametric settings interacting on both a higher sentence or clause level (= a VP with arguments) and several smaller levels with NPs (and their arguments) and other individual phrases leaving not much room for free placement.<sup>13</sup>

With that being said, I will now describe the nominal macro-, meso- and microparametric values of Old Avestan, leaving aside verbal aspects for future studies. Noun phrases of Old Avestan are defined (or rather restricted) by the following parametric settings:

1. Gramm. Person (FGP): set to yes. Cf. the distinction between 1st, 2nd and 3rd person pronouns (Skjærvø 2009b, p. 81) and the three-way contrast in terms of deictic demonstratives. See also Hoffmann and Forssmann (2004, pp. 158–169).
2. Gramm. Number (FGN): yes. There is difference between singular, dual and plural, and traces of the PIE collective. Cf. e.g., the dative dual in the example above (Y. 47,6), *rānōi-biiā* ‘to/for the two contestants’, or the

<sup>13</sup>The surface word order would then be (or have been) either the result of parametric settings interacting with situational information structure (concerning Greek – which has received much more attention than old (Indo-)Iranian languages –, see, e.g., Dik 1995, Matić 2003, Allan 2012, 2014, Bertrand 2013, Goldstein 2016 or Beschi 2018) or, perhaps, complex parametric interferences which result(ed) in orders that can be misinterpreted by modern observers as being primarily pragmatically motivated, whereas they are (or were) exclusively syntactic in nature and origin. For instance, from a pragmatic point of view, we may analyze a clause (or phrase) as showing fronting or left dislocation, but from a syntactic one the same surface order could be interpreted as pragmatically neutral N (or V) raising.

various nominal paradigms exemplifying number marking on nouns and adjectives in Hoffmann and Forssmann (2004).

3. Gramm. Gender (FGG): yes. There is a distinction between masculine, feminine and neuter gender. Cf. e.g., the near-deictic demonstrative pronoun *ima-/a-* with its nom. sg. masculine *aīam*, fem. *iīam* and neuter *imat* (Skjærvø 2009b, p. 83).
4. Feature spread to N (FSN): yes. Cf. again the dative dual in the example above (Y. 47,6), *rānōi-biiā*, or other examples like Y. 43.15 *aṭ tōi vīspāṇg aṇgrāṇg ašāunō ādarā* which I would translate as “thus, they call all the evil ones righteous (ones)”. Here, plurality (= number) is directly morphologically marked (in conjunction with accusative case) on the head noun *aṇgrāṇg* ‘evil (ones)’, on the accompanying preceding pronominal adjective *vīspāṇg* and on the following predicative (or appositive?) adjective *ašāunō*.<sup>14</sup>
5. Number on N (bounded nouns) (FNN): yes, see the preceding example of (FSN).
6. Gramm. partial definiteness (DGP): set to no. There is no obligatory anaphoric marking of nouns mentioned previously in the discourse. Neither does Old Avestan have specialized articles nor definiteness affixes on nouns (or nominal arguments); definiteness *can* be expressed via the relative pronoun + N (usually understood as verbless relative clauses), but it need not. Consider for instance the followings NPs: Y. 46.1 *daṁiīuš yōi sāstārō drəguuāntō* ‘the wrongful governors of the land’ or Y. 46.3 *yōi uxšānō asnqm* ‘the oxen of days (= sunrisings)’ (West 2011, p. 82). In these

<sup>14</sup>Interestingly, Skjærvø (2009b, p. 132) translates this passage as “thus, [they say/claim] (that) all the evil ones (are) followers of Order”, whereas West (2011, p. 120) has “they have declared all the righteous their enemies”. Apparently, a translation depends on the personal interpretation of the structure of the clause, or, in other words: Is there just one NP, or are there two separate ones?

NPs, the relative pronoun seems to indicate definiteness. However, there are many more instances in which the relative pronoun does not occur, whereas the NP must be understood as being definite, e.g., Y. 51.15 *garō dāmānē ahurō mazdā jasaṭ pouruiiō* ‘the lord Mazda enters the house of song as the first one’, Y(H) 39.2 *ašāunqm āaṭ urunō* ‘the souls of the righteous’, Y. 44.20 *karapā usixšcā* ‘the Karpan and the Usij’, Y. 30.4 *acištō drəguuatqm aṭ ašāunē vahištəm manō* ‘the worst (existence shall be) of the wrongful ones, but best thought for the righteous one’ or Y. 32.14 *nī kāu-uaiiascīṭ xratūš nī dadaṭ varəcā.hīcā* ‘the very Kavis give up their intellects and dignities’. Hence anaphoric definiteness marking is **not** obligatory and parameter DGP must get a negative value. Negative setting of (DGP) renders 14 parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA) irrelevant or rather neutralizes them. If neutralized, the respective parameters will be left out of the discussion.

7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking. There is a partitive genitive (e.g., Y. 32.8 *gāuš bagā* ‘portions of the cow’), but this does not permit a positive setting of this parameter, because genitives have several other functions in Old Avestan.
8. D-controlled infl. on N (DIN): no. Old Avestan has nothing reminiscent of the Arabic nunation.
9. Plural spread from Cardinals (CPS): yes. There is only one example in the corpus, but this shows plural spread from cardinals: Y. 44.18 *dasā aspā* ‘ten mares’, with (fem.) nom.-acc. pl. ending *-ā* (or *-ā̃*).
10. Grammaticalized boundedness (CGB): no. Nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. A singular count noun is inflected for singular and refers to a singularic entity, cf. Y. 51.5 *yaθā ... gqm vīdaṭ vāstriiō šīiaoθanāiš ərəšuuō*

‘how a herdsman upright in his actions obtains a cow’, not \*‘cows’. The negative value of CGB renders the logically dependent parameter CCN irrelevant.

11. Structured Adjective Phrases (AST): yes. Cf. Y. 47.2 *ahiiā mainiīuš spāništahiiā* ‘of this most bounteous will’, showing Dem–N–Adj with an inflected adjective (phi-feature concord), in this case in the genitive. Another example: Y. 34.15 *vahištā srauuāscā šiiāoθanācā* ‘the best things to be known for (lit. “fames”) and deeds’.
12. Feature spread to structured APs (FFS): yes, see AST directly above.
13. Feature spread to predicative APs (FSP): yes, cf. Y. 29.3 *hātqm huuō aojištō* ‘he (is) strongest of all beings’ with a genitive, 3rd person pronoun (Nom. Sg.) and inflected (Nom. Sg.) predicative adjective. Or see Y. 27.14 *ašm vohū* ‘right (is) good’.
14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. See the sections in Skjærvø (2009b, 154ff.) or West (2011, 77ff.). One example: Y. 51.8: *huuō zī mθrā šiiātō yā vīdušē mrauuaitī* ‘that very prophet is happy who speaks to one who knows’. Positively set ADR neutralizes the dependent parameter AER.
15. Free reduced RelCl (ARR): set to yes. OAv., like other ancient IE languages, has free adjectives functioning as reduced relative clauses,<sup>15</sup> i.e., participles. Cf. e.g., Y. 31.17 *vīduuā vīdušē mraotū* ‘let the knowing one speak to the knowing’, or other examples from West (2011, 72ff.).
16. Adpositional genitive (GAD): no, genitives are inflectional.

<sup>15</sup>See Baker and Vinokurova (2009), Cinque (2010), Kayne (1994), Sleeman (2011) or several articles in Le Feuvre et al. (2017) for the theoretical background of such free adjectives.

17. Free Gen. (GFR): yes, there are free, recursive genitives, e.g., Y. 43.9 *rātqm nəmanhō Ašahiia* ‘the tribute of reverence of Truth/Right’.
18. Uniform Gen. (GUN): yes, there is a uniform genitive being inflectional and not bound to certain structural positions. Take the example from parameter GFR or, this time preceding the noun (and being recursive), Y. 45.11 *dāng patōiš ... uruuaθō* ‘the ally of the master of the house’. Genitives are extremely free and can surface variously inside or outside the NP core, see the discussion in West (2011, pp. 36–40, 123).
19. DP over free Gen. (GPR): yes, (free) genitives can occur both before and after the rest of the NP/DP, so in generative terminology, DPs/NPs can be raised over free genitives. See GFR for an example. Since a uniform genitive is present, it can also take over the role of an object genitive, as in Y. 31.10 *vanhāuš fšānghīm manahō* ‘the cultivator of good thought’ (with Gen–N–Gen). This neutralizes parameter GFO.
20. Gen-feature spread to N (GFS): no. OAv. is a double-marking and no head-marking language in the sense of Nichols (1986).
21. Adjectival poss. (APO): yes. This parameter asks whether possessives distributionally and morphologically behave as adjectives or not. This is the case in OAv., cf. Y. 37.3 *ahmākāiš azdabīš* ‘with our bones’ (Adj.–N) or Y. 51.13 *x<sup>v</sup>āiš šiiaoθanāiš hizuuascā* ‘through his own actions and his tongue’s’ (the first word is a possessive adjective; but as the last word, ‘tongue’, shows, possession can also be expressed via a genitive).
22. Clitic poss. (PCL): yes, OAv. has the opportunity to express possession by means of dative enclitic pronouns, e.g., Y. 32.2 *spəntqm vā ārmaitīm vanuhīm varəmaidī; hā nā aṇhaṭ* ‘we adopt your liberal piety, (being) good; it shall be ours.’ The dative pronouns are *vā* and *nā*, respectively.
23. N-feature spread to free Gen. (GSP): no. Genitival arguments are NPs of their own, without phi-feature agreement to N. See, for example, Y(H) 35.8



*kahmāicīṭ hātqm* ‘for any of the (living) beings’ (with N in the singular, but the genitive marked for plural).

24. Adjectival Gen. (AGE): no. Genitival relations are either expressed via inflectional endings on nouns, pronouns and adjectives or via possessive adjectives (West 2011, p. 36).
25. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Old Avestan.
26. Strong partial locality (TPL): no. TPL asks whether or not a subset of demonstrative pronouns, viz. deictic ones, systematically surface in the D-area. This is not the case. Even though demonstratives usually introduce NPs, there are a few examples (Y. 32.13; 45.3,4 *aṇhōuš ahiiā* ‘this world’) with deictic demonstratives and a N–Dem order. Thus there is no *systematic* spread of demonstratives to D. This neutralizes the subsequent parameter TSL.
27. N over Demonstratives (NDE): yes. Determiner elements – OAv. has no true article, but demonstratives – usually surface to the left of their NP(s), cf. e.g., Y. 36.6: *imā raocā* ‘these lights’. However, there are exceptions, e.g., Y. 32.13; 45.3,4 *aṇhōuš ahiiā* ‘of this world’ with N–Dem. Old Avestan thus has the option to raise the noun over determiner elements.
28. N over Cardinals (NOC): unclear. There is only one example of an NP with a cardinal in the OAv. corpus, and this shows a Card–N order: Y. 44.18: *dasā aspā* ‘ten mares’. This would imply a negative setting of this parameter. The parametric hierarchy, however, indicates that nouns should have the possibility to be raised over cardinals, since OAv. nouns can also be raised over demonstratives (D-elements) and, as the next parameter shows, ordinals. Probably this is just a transmission problem.

29. N over Ordinals (NOO): yes, Old Avestan has the option to raise a noun over an ordinal number, cf. e.g., Y. 28.11 *aṇhuš paouruiiō* ‘the first existence’ (N–Ord).
30. N over adjectives (NOA): predictably yes, and this is confirmed e.g., by *spādā anaocaṇhā* ‘the two hostile armies’ with N–Adj. West (2011, pp. 121–122) states that adjectives following their head noun often have more syllables and thus are longer than the noun. This may be the case (and it is an interesting observation),<sup>16</sup> but it does not affect the argument that nouns can be raised over adjectives or ordinals.
31. N over GenO (NGO): predictably yes, and this is confirmed by Y. 28.4 *aēšē Ašahiiā* ‘in search of Truth/Right’ (N–GenO).
32. N over external arguments (NOE): yes, examples under parameters NOO, NGS, NOA and NGO.
33. Verbless relative clauses (VRC): yes, see example 14 of Chapter 2 and the discussion in Reichelt (1978, 370ff.) and Seiler (1960). Note also Probert (2015, pp. 407–414) for a discussion of verbless relative clauses in other ancient Indo-European languages.
34. Facultative linker (EZ1): no. OAv. has only verbless relative clauses with fully inflected relative pronouns/connectors.
35. Generalized linker (EZ2): no.

As has been discussed in Chapter 2, the most interesting observation is the amount of N-raising (even over demonstratives). Old Avestan generally shows rather great freedom in terms of raising of any type of syntactic element, cf. e.g.,

---

<sup>16</sup>See Truckenbrodt (2007) or Ahn (2016) for modern works on the ‘syntax-phonology interface’, but one should also mention Behaghel (1909) here. The idea, in short: syntactic surface structure may be dependent on phonological properties of the lexemes involved.

## (19) Old Avestan, raising of elements

## a. Raising of a verb (Y. 46.2)

*ahmī mazdā anaēšō*  
 be.IND.PRS.1SG Mazdā.VOC.SG weak.NOM.SG  
 ‘I am, O Mazdā, weak.’

## b. Raising of an object in a transitive clause with normal SOV order (Y. 31.2)

*yaθā ratūm ahurō vaēdā*  
 how model.ACC.SG lord.NOM.SG know.PRF.1SG  
 ‘How the Lord knows the model (to be)’.

To sum up, only 35 of the 53 parameters required a discussion or had to be set. Particularly parameter DGP turned out to be highly influential, neutralizing, due to its negative value, 14 other parameters.

### 3.3.2 Young Avestan

Young Avestan (YAv.) is not a direct descendant of Old Avestan, but rather a closely related cousin. Young Avestan was probably spoken in the first half of the last millennium BCE and likewise somewhere in southern Central Asia (Yaz II culture?). Apart from some minor differences (somewhat reduced case inventory etc.), the general picture of the language is still very much that of an ancient Indo-European language. An illustrative example of Young Avestan syntactical properties may be seen in the following sentence:

(20) Young Avestan (Y. 57.11-12)

*vispae<sup>i</sup>biiō haca ar<sup>ə</sup>zaē<sup>i</sup>biiō vauuanuuā pa<sup>i</sup>ti*  
 all.ABL.PL from battle.ABL.PL win.PART.PRF.NOM.SG.M back  
*jasā<sup>i</sup>ti*  
 come.IND.PRS.3SG  
 “Having won, he comes back from all battles.”

As can be seen, case endings combined with an adposition signal an ablatival relationship of a nominal adjunct (*pa<sup>i</sup>ti jāsā<sup>i</sup>ti* would suffice as a clause), and besides a regular predicate, surfacing at the right end of the utterance and consisting of a verb and a preverb, there is a participle functioning simultaneously as (overt) subject and as a reduced clause further specifying this subject. Typical for an ancient Indo-European language, the NP *vispae<sup>i</sup>biiō haca ar<sup>ə</sup>zaē<sup>i</sup>biiō* evinces a hyperbaton (as in Latin *magnā cum laude* ‘with great praise’) and the language is a pro(noun)-drop language, with the verbal ending *-ti* signaling the person (3rd in this case) and grammatical subject. The information of 2 clauses (‘He has won/is victorious. He returns from all battles.’) is stored or conveyed in just 6 words.

The parametrical values do not significantly differ from Old Avestan and can be defined as follows:

1. Gramm. Person (FGP): set to yes. Cf. the distinction between 1st, 2nd and 3rd person pronouns (Skjærvø 2009b, 81, 116ff.) and the three-way

contrast in terms of deictic demonstratives. See also Hoffmann and Forssmann (2004, pp. 158–169).

2. Gramm. Number (FGN): yes. There is difference between singular, dual and plural, and traces of the collective (see Skjærvø 2009b, p. 103). For number marking in general, see the various nominal paradigms exemplifying number marking on nouns and adjectives in Hoffmann and Forssmann (2004).
3. Gramm. Gender (FGG): yes. There is a distinction between masculine, feminine and neuter gender. Cf. e.g., the near-deictic demonstrative pronoun *ima-/a-* with its nom. sg. masculine *aīam*, fem. *iīam* and neuter *imat* (Skjærvø 2009b, p. 83).
4. Feature spread to N (FSN): yes. Cf. Yt. 5.93 *auuā zaoθrā* ‘these libations’ (number is marked, notwithstanding the demonstrative, on the noun).
5. Number on N (bounded nouns) (FNN): yes, cf. the preceding example of (FSN).
6. Gramm. partial definiteness (DGP): set to no. There is no obligatory anaphoric marking of nouns mentioned previously in the discourse, only a facultative one (among others by means of the so-called relative particle or relative connector; see below, parameter FLI). NPs need not be marked for definiteness, as can be seen, i.a., in Yt. 13 which centers around the Fravashis. They are mentioned early on in the hymn and in the course of the text, they are mentioned again and again, but without definiteness marker. Consider e.g., Yt. 13.31 *ašāunqm vaṇuhīš sūrā spəṇtā frauuašaiiō yazamaide* ‘We worship the good, strong, beneficent Fravashis of the faithful’. The Fravashis here are clearly definite, but there is no obligatory anaphoric marking. As with OAv., negative setting of (DGP) renders 14 parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA) irrelevant or rather neutralizes them.

7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking. There is a partitive genitive, but this does not permit a positive setting of this parameter. Even though cardinal numbers higher than ‘100’ require a following noun to be marked for genitive plural (Yt 1.19 *hazaṇrām narqam* ‘thousand men’, compared with V. 18.31 *caθβārō aršāna* ‘four males’), this does not permit a positive setting of this parameter.
8. D-controlled infl. on N (DIN): no. Young Avestan has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. Yt. 14.44 *catanrō parānā* ‘four feathers’ or the examples of parameter DPQ.
10. Grammaticalized boundedness (CGB): no. Singular count nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. An example can be seen in Yt. 13.7 *huuaspāi naire barəmnāi* ‘a man riding on a good horse’ [DAT.SG], not to be understood as \*‘(several) men’. Negatively set CGB neutralizes CCN.
11. Structured Adjective Phrases (AST): yes. Cf. Yt. 13.129 *vīspəm ahūm astuuantəm* ‘the entire bony existence’.
12. Feature spread to structured APs (FFS): yes, see AST directly above, where singular and accusative are spread from the noun (*ahūm*) to the adjectives (*vīspəm ... astuuantəm*).
13. Feature spread to predicative APs (FSP): yes, cf. V 5.59 *yaṭ ahmi nmāne ... nā<sup>i</sup>rika daxštauua<sup>i</sup>ti aṇhaṭ* ‘when a woman becomes menstruant in this house’, Yt. 5.129 *yaṭ asti baβriš sraēšta* ‘because the female beaver is the most beautiful’ or Yt. 19.12 *bun gaēθā amaršantiš* ‘the creatures (shall become, aorist subj.) indestructible’.

14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. See Skjærvø (2009b, pp. 154, 158). One example: Y. 19.1 *ciṭ **auuat** vacō ās ahura.mazda **yaṭ** mē frāuuaoṇō* ‘which was **that word**, Ahura Mazda, **which** you said forth to me?’ (Skjærvø 2009b, p. 158). Since ADR is set positively, parameter AER becomes irrelevant.
  
15. Free reduced RelCl (ARR): set to yes. YAv., like other ancient IE languages, has free adjectives functioning as reduced relative clauses, i.e., participles. Cf. e.g., P26 *zazuš vīspaēšu vanhušu* ‘having won in all good things’ with a participle and two dependent locatives.
  
16. Adpositional genitive (GAD): no, genitives are inflectional.
  
17. Free Gen. (GFR): yes, there are free, recursive genitives, e.g., Y. 9.13 *tūm (...)* *zaraθuštra nmānahe pouruṣaspahe* ‘you (are) Zarathustra of the house of Pouruṣāspa’. Cf. also Vd. 2.33 *taoxma upa.barat pasuuqmca staoranqmca mašiianqmca sūnqmca vaiiqmca āθrqmca suxrqm saociṇtqm* ‘he brought the seeds of sheep and oxen, of men, of dogs, of birds, and of red blazing fires’. *taoxma* ‘seed.ACC.PL’ is the head noun, *upa.barat* the verb, and all other words are either direct genitival arguments of the noun or indirect ones (2nd order recursive genitives) in the case of *suxrqm saociṇtqm* ‘red.GEN.PL, blazing.GEN.PL’ further defining 1st order *āθrqm* ‘fire.GEN.PL’.
  
18. Uniform Gen. (GUN): yes, there is a uniform inflectional genitive not bound to certain structural positions, i.e., functioning inter alia as subject and object genitive. Cf. Y. 57.3 *ahe yasna yazatanqm* ‘by his (GenS) sacrifice of/to the gods (GenO). Since this uniform genitive occupies also the functions of an object genitive, as can be seen by the example given here, parameter GFO is neutralized.

19. DP over free Gen. (GPR): yes, (free) genitives can occur both before and after the rest of the NP/DP, so in generative terminology, DPs/NPs can be raised over free genitives. See GFR for an example.
20. Gen-feature spread to N (GFS): no. There is no feature spread from genitives to their head nouns, both show phi-feature spread to dependent adjectives or demonstratives, but do not agree with each other. Consider e.g., Yt. 19.35 *miθrəm vīspanqm daxiiunqm dañhupaitīm yazamaide* ‘We sacrifice unto Mithra, the lord of all countries’ (lit. Mithra, of all countries the “countrylord”).
21. Adjectival poss. (APO): no. As stated in Skjærvø (2009b, p. 119), YAv. has no genuine possessives, but instead makes use of the genitive of the personal pronouns. Young Avestan differs from Old Avestan in this regard. Since this parameter has a negative value, AGE is neutralized.
22. Clitic poss. (PCL): yes, YAv. has the opportunity to express possession by means of dative enclitic pronouns, cf. e.g., the Wackernagel-position clitic in V 2.5 *azəm tē gaēθā frāðaiieni* ‘I shall further your herds’.
23. N-feature spread to free Gen. (GSP): no. There is also no feature spread from the noun to genitives. See GFS.
24. Poss.-checking N (GCN): no. Distinctive morphological marking (head marking, such as the Semitic construct state) does not occur on YAv. nouns when occurring with a genitive argument.
25. Strong partial locality (TPL): no. Demonstratives frequently surface in the D-area, introducing NPs. Cf. e.g., V. 2.11 *imqm zqm* ‘this earth’. However, there are also rare counterexamples (see NDE), so that this parameter must get a negative setting. This neutralizes TSL.
26. N over Demonstratives (NDE): yes. Demonstratives mostly surface to the left of their nouns, but there are cases like Y. 19.3 *baya aēša ās (ahunahe*



*vairiiehe*) ‘This was the part/piece (of A. V.)’ evincing an order N–Dem–Verb (answering to Y. 19.1 *ciṭ auuaṭ vacō ās* ‘what saying/adage was this?’ with Dem–N). Even though rarely used, the language nevertheless could raise nouns over demonstratives.

27. N over Cardinals (NOC): yes. Quite clearly, the normal pattern for YAv. is an order Card–N, e.g., N. 84 *hapta hanti ... ratauuō* ‘seven are ... the models’, V. 18.31 *caθβārō aršāna* ‘four males’ or V. 7.51–52 *duua maniu* ‘two spirits’. However, when using the proto-*ezāfe* or proto-linker construction, YAv. also evinces an N–Card order, cf. e.g., Yt. 8.33 *auui karšuuqn yāiš hapta* ‘to the seven parts of the earth’, lit. ‘to the parts of the earth with the seven’ (vel sim.), or Y. 72.5 *vīspāiš haca karšuuqn yāiš hapta* ‘from all seven parts of the world’. Compare this with Yt. 10.85 *hapta karšuuqn* ‘seven parts of the world’.
28. N over Ordinals (NOO): (no). The normal pattern is Ord–N, cf. Yt. 17.57 *paoiriiqm gərəzqm* ‘the first complaint’. The NPs with ordinals in Vd. 18.39–40; 18.45–46 etc. have no overt nouns. I found no examples with ordinals and N–Ord order in the whole corpus (ordinals 1–20), and Reichelt (1978) has not a single line on the internal order of NPs with ordinals. The parametric hierarchy would predict that YAv. nouns could be raised over N.
29. N over adjectives (NOA): yes, cf. Vr 2.5 *narəm ašauuanəm* ‘the righteous man’ with N–Adj.
30. N over GenO (NGO): yes, cf. the example of parameter GUN.
31. N over external arguments (NOE): yes, examples under parameters NOO, NGS, NOA and NGO.
32. Verbless relative clauses (VRC): irrelevant due to positive setting of FLI.

33. Facultative linker (FLI): yes. As described in Chapter 2, Young Avestan has a connecting relative. It is certainly not a generalized linker like the marker of the modern Persian *ezāfe* construction, as the YAv. pronoun is still declinable (when the antecedent is in the nominative or accusative, otherwise invariably the NOM.-ACC.SG. neuter form *yaṭ* occurs). On the other hand, it has properties of a linker (Skjærvø 2009b, pp. 100–101), connecting a noun with dependent adjectives or genitives (see Seiler 1960, pp. 134–170 or Haider and Zwanziger 1984). The relative particle/connector is facultative in Young Avestan and not fully grammaticalized, hence parameter FLI gets a positive value. Cf.

(21) Young Avestan prototype of a linker

a. nominative (H. 2.2)

*vīspəm imat̃ yaṭ juiiō*  
 all.NOM.SG this.NOM.SG.N REL.NOM.SG.N living.NOM.SG  
*aṇhuš*  
 being/existence.NOM.SG  
 ‘All this, this living existence.’

b. uninflected *yaṭ* (Yt. 5.18)

*puθrəm yaṭ pourušaspahe*  
 son.ACC.SG.M REL.ACC.SG.N Pourušaspa.GEN.SG  
 ‘The son of Pourušaspa.’

34. Generalized linker (GLI): no.

### 3.3.3 Old Persian

Old Persian (OPers.), the language of the official inscriptions of the Achaemenid dynasty, is the oldest representative of the south-western group of Iranian languages. Certain developments clearly distinguish it from Old Avestan (reduction of the case inventory etc.), but the language still has a rather complex morphological repertoire compared with Middle Persian. Syntactically, Old Persian resembles Young Avestan, particularly concerning the emergence of a generalized (though not obligatory) linker and similar phenomena. See the following sentence:

- (22) Old Persian, basic transitive sentence (DB 2.25f.)

<i>kāra</i>	<i>haya</i>	<i>manā</i>	<i>avam</i>	<i>kāram</i>	<i>tayam</i>
army.NOM.SG	REL.NOM.SG	I.GEN.SG	that	army.ACC.SG	REL.ACC.SG
<i>hamiçiyam</i>	<i>aja<sup>n</sup></i>	<i>vasiy</i>			
disloyal.ACC.SG	crush.IPRF.3SG	much			

‘My army utterly destroyed the rebellious army’.

The subject, *kāra* ‘army’, is further specified by a genitive pronoun linked by means of a relative connector (*haya manā*) to its head noun. The direct object, in turn, consists of basically two words, *avam kāram* ‘that army’, which again are further specified by another relative connector, this time linking an adjective to the noun: *tayam hamiçiyam*.

In terms of NP parameters, Old Persian can be defined as follows:

1. Gramm. Person (FGP): set to yes. Cf. the distinction between 1st, 2nd and 3rd person pronouns (Skjærvø 2009b, 81, 116ff.) and the two-way contrast (near–far) in terms of deictic demonstratives.
2. Gramm. Number (FGN): yes. There is a morphosyntactically overtly marked difference between singular and plural (and vestiges of the old dual). See Kent (1953, p. 57).

3. Gramm. Gender (FGG): yes. There is a distinction between masculine, feminine and neuter gender. Cf. e.g., the near-deictic demonstrative pronoun *ima-/a-* with its nom. sg. masculine *aīam*, fem. *iīam* and neuter *imat* (Skjærvø 2009b, p. 83).
4. Feature spread to N (FSN): yes. DNa 16 *imā dahayāva* ‘these lands’ (Dem–N) with the noun marked for plural number.
5. Number on N (bounded nouns) (FNN): yes, cf. the preceding example of (FSN).
6. Gramm. partial definiteness (DGP): set to no. As with Old and Young Avestan, there is no obligatory anaphoric marking of nouns or rather NPs mentioned previously in the discourse in Old Persian. OPers. does not have specialized articles or definiteness affixes on nouns (or nominal arguments); though specificity may be marked by means of a relative pronoun used as a connector (DB 1.44 *gaumāta haya maguš* ‘Gaumata the Magus’). The lack of anaphoric definiteness marking can be seen in DB 1.51–52 *avahyarādiy **kāram** avājaniyā mātyamām xšnāsātiy* ‘therefore he would smite **the army** so that it will not recognize me’ – the army clearly is definite, but unmarked. Again, the negative setting of (DGP) neutralizes 14 parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA).
7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking.
8. D-controlled infl. on N (DIN): no. OPers. has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. DSf 26, even though the cardinal is not written out: *XL arašnīš baršnā* ‘forty cubits in depth’; *arašnīš*, the noun, is marked for plural number (and accusative case).

10. Grammaticalized boundedness (CGB): no. Singular count nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. Consider DB 1.48 *naiy āha martiya* ‘there was no man’. Since CGB has a negative value, CCN is neutralized, as in Old and Young Avestan.
11. Structured Adjective Phrases (AST): yes. Cf. DZc 6 *bumiyā vazrakāyā* ‘great earth’ (with N-Adj), DE 1 *baga vazraka auramazdā* ‘A(h)ura Mazda (is) a great god’ or DB 2.18 *kāra pārsa* ‘the Persian army’.
12. Feature spread to structured APs (FFS): yes, see AST directly above.
13. Feature spread to predicative APs (FSP): yes, cf. DB 1.85–86 *utā abiš nāviyā āha* ‘and it was navigable throughout its waters’, with *nāviyā* being inflected for gender, number and case (NOM.SG.F., referring to a female river name).<sup>17</sup>
14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. See Skjærvø (2009b, pp. 154–155, 159–160). One example: DB 3.74–75 *martiyā tyaišaiy fratamā anušiya āhatā avāja* ‘he killed **the men who** were his foremost followers’. This parametric value neutralizes AER.
15. Free reduced RelCl (ARR): set to yes. Old Persian, like other ancient Indo-European languages, has free adjectives functioning as reduced relative clauses, i.e., participles or verbal adjectives. Cf. e.g., DB 2.75, 89–90 *duvarayā-maiy basta adāriya* ‘he was held captive (lit. bound, *basta*) at my gate’. Another example: DB 2.38 *hamiçiyā hagmatā* ‘the conspirators, having come together’.

<sup>17</sup>Skjærvø (2009b, p. 114) has a different translation, viz. ‘and it (= the Tigris) was \*deep [with waters] (= in spate)’.

16. Adpositional genitive (GAD): no, OPers. genitives are inflectional.
17. Free Gen. (GFR): unclear. Old Persian probably had recursive free genitives, but I found none in the texts surveyed by me (or overlooked them). Neither Kent (1953) nor Brandenstein and Mayrhofer (1964) nor Skjærvø (2009b) give any information concerning recursive genitives. There are only appositives to genitives being themselves genitives, like e.g., DB 3.58–59 *dārayavahauš xšāyaθiyahyā* ‘of Dareios the king/of king Dareios’. Another example can be seen in DNa 10 *xšāyaθiya dahyūnām vispazanānām* ‘king of countries of all peoples’.<sup>18</sup> Perhaps there really are none in the corpus – we should not forget that the Old Persian corpus is, in general, not very great as compared with, say, the Young Avestan one.
18. Uniform Gen. (GUN): (yes), there is a uniform inflectional genitive not bound to certain structural positions. Cf. DPi *dārayavahauš Xšhyā viθiyā* ‘in the house of king Dareios’ (Gen–Gen–N), DNa 43–44 *pārsahyā martiyahyā ... arštiš* ‘the spear of a Persian man’ (Gen–Gen... N), XV 1–2 *maθišta bagānām* ‘the greatest of the gods’ (N–Gen), or DNa 7–8 *aivam parūvnām framātāram* ‘one lord of many’ (Num–Gen–N). However, this genitive is not recursive, so strictly speaking not all preconditions for this parameter are given; probably this is just an artifact of the limited text corpus. The presence of a uniform genitive would render parameter GFO irrelevant.
19. DP over free Gen. (GPR): unclear due to a lack of unambiguous examples of recursive free genitives (probably yes).
20. Gen-feature spread to N (GFS): no. As in Old and Young Avestan, Old Persian genitives constitute dependent NPs on their own, optionally taking

<sup>18</sup>Or is this a recursive genitive in DNa 10? Is *vispazanānām* just an attributive adjective to *dahyūnām*, or an NP on its own, being dependent on the latter? Compare also XV 11–12 *xšāyaθiya dahyūnām paruv zanānām* – Note in passing that *vispazanānām* is no OPers. form (which would be \**vispadanānām*).

adjectives and demonstratives which then must agree in terms of case, number and gender; but there is no feature spread of genitives to their head noun. See DNa 43–44 *pārsahyā martiyahyā ... arštiš* ‘the spear of a Persian man’.

21. Adjectival poss. (APO): no, possession is either expressed via clitic possessive pronouns or via genitives, particularly also with genitives of personal pronouns (see Skjærvø 2009b, p. 119). One example: DBa 14–15: 2-2-2-2 *manā taumāyā* ‘8 of my family/dynasty (were kings before me)’, another one can be seen in XV 18 *manā pitā* ‘my father’. A negative value of APO neutralizes parameter AGE due to an implicational relationship.
22. Clitic poss. (PCL): yes, Old Persian has the possibility to express possession by means of genitive-dative enclitic pronouns, e.g., DNa 53–55 *utā-maiy viθam utā imām dahayāum* ‘both my house and my (lit. this) land’.
23. N-feature spread to free Gen. (GSP): no. Genitives constitute dependent NPs with internal phi-feature spread to adjectives or other elements, but genitives do not agree with their own, higher-ranking head nouns.
24. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Old Persian.
25. Strong partial locality (TPL): yes, demonstratives systematically surface in the D-area, introducing NPs. Cf. e.g., XPa 11–13: *imam duvarθim* ‘this gate’.
26. Strong locality (TSL): yes, apparently all demonstratives occur in the D-area; there are no differences. Cf. XV 2–4 *haya imām būmim adā haya avam asmānam adā* ‘who created this earth, who created yonder sky’.

27. N over Demonstratives (NDE): no, nouns apparently cannot be raised over demonstratives. Demonstratives consistently surface to the left of N, e.g., XV 24 *imām dipim* ‘this inscription’ or XV 12–13 *xšāyaθiya ahyāyā būmiyā* ‘king in/on this earth’ (N–[Dem–N]<sub>loc</sub>).
28. N over Cardinals (NOC): no (probably), but there are not many examples of cardinals at all. One of these is DB 1.37–38 *Viyaxnahya māhyā XIV raucabiš θakatā āha* ‘14 days of the month Viyaxna were past’.
29. N over Ordinals (NOO): unclear, the only two instances of an ordinal, DB 1.10, and DBa 16–17, contain ‘ninth’ in predicative function (*adam navama* ‘I am the ninth’). There are not enough data.
30. N over adjectives (NOA): yes, cf. the famous passage DB 5.22 *sakā tyaiy xaudām tigrām baratiy* ‘the Sakas who wear a pointed hat (lit. hat–pointed)’, or DB 2.18 *kāra pārsa* ‘the Persian army’.
31. N over GenO (NGO): yes, cf. DNa 10 *xšāyaθiya dahyūnām vispazanānām* ‘king of the countries of all peoples’.
32. N over external arguments (NOE): yes, examples under parameters NOO, NGS, NOA and NGO.
33. Verbless relative clauses (VRC): irrelevant due to positive FLI.
34. Facultative linker (FLI): set to yes. As described in Chapter 2, Old Persian has verbless relative clauses with inflected relative pronouns (of the type *gaumāta haya maguš* ‘Gaumāta the Magus’). This type of NP modification involves no generalized linker like the marker of the modern Persian *ežāfe* construction, as the Old Persian pronoun/connector *haya-* is still declinable for gender, number and case. Going back to verbless relative clauses, constructions like DNa 21 *dātam taya manā* ‘my law’ (lit. the law that is mine’) or DB 2.25 *kāra haya manā* ‘my army’ have some properties of a linker (see Skjærvø 2009b, pp. 100–101, Seiler 1960, pp. 134–170,



Haider and Zwanziger 1984 or Yakubovich 2020) as they can connect a noun with dependent adjectives or genitives, but they are not comparable to the modern situation, as *haya-/taya-* is facultative and not obligatory (and still declinable). There are many NPs built without these facultative or proto-linkers, and the internal order is still flexible – for instance, a Farsi-type linker construction should have the order *\*dātā tayanā manā* (ignoring the inflection of the linker); attested, however, is DB 1.23 *imā dahyāva tayanā manā dātā apariyāya* ‘these lands behaved according to my law [ABL-INST]’.

35. Generalized linker (GLI): no. OPers. has no generalized linker.

### 3.4 Middle Iranian languages

The periodization of Iranian languages into three major eras (Old, Middle, Modern/New) is based on a correlation with major cultural turning points. Middle Iranian languages are traditionally defined as Iranian varieties which were spoken between the end of the Achaemenid empire and the advent of Islam, i.e., roughly from the 4th century BCE to ca. 800/900 CE (depending on when exactly Islamization occurred in the eastern territories). Again, there are significant differences in the exact frames of attestation, with some languages being preserved earlier and some later.<sup>19</sup> All six substantially preserved Middle Iranian languages show typical features of modern Iranian languages spoken in their individual regional contexts (phonological developments, morphological innovations/archaisms, lexical peculiarities), thus evincing the same variation which still characterizes the Iranian linkage. Nevertheless, they clearly represent an older stage in the development, still having (at least the eastern languages) rich sets of nominal endings.

#### 3.4.1 Khotanese

Khotanese (Khot.), like the closely related Tumshuqese (Tumsh.), was a Sakan variety once spoken in the Tarim basin in what is now northwestern China. Khotanese was the native language of the kingdom of Khotan and is chiefly preserved from the 7th to the 10th centuries in Buddhist texts (and Brāhmī script), but several legal documents and letters are also known. The only extant Sakan language is Wakhi (on which see Morgenstierne 1973, Lorimer 1958, Paxalina 1975, Steblin-Kamenskij 1999, Bashir 2009), but Wakhi is not a direct descendant of either Khotanese or Tumshuqese. Neither of the latter two has any living descendants.

---

<sup>19</sup>Chorasmian is an extreme case, as it is mainly preserved in texts dating to the 11th, 13th and 14th centuries. Note also the fact that centuries after the fall of the Sasanian empire Middle Persian was preserved, taught and written as a liturgical language in the Tarim Basin, similar to Latin in Europe.

Of all Middle Iranian languages, Khotanese (and Tumshuqese) Saka shows the greatest resemblance to Old Iranian representatives. Khotanese nouns and nominal arguments have different sets of endings for masculine and feminine gender and there are remnants of the old neuter. Singular and plural number are morphologically marked, and in the oldest documents six cases (in the singular) can be distinguished, namely nominative, accusative, genitive-dative, instrumental-ablative, locative and vocative. Khotanese has the richest repertoire of nominal inflectional endings of all Middle Iranian languages, making it much more similar to the morphosyntactic systems of Young Avestan or Old Persian than to those of Middle Persian or Parthian. This is an interesting observation, given the fact that the Khotanese documents mostly date to the second half of the 1st millennium CE. The earliest Khotanese texts and the youngest Young Avestan ones, despite being separated by a gap of more than 1000 years, have more in common than Khotanese texts with contemporary Bactrian ones, due to a comparably complex morphological inventory in Khotanese. We are probably observing a *Randsprachenarchaismus* here (though a contact-induced retention is also a possibility, given the contact with Tocharian).

A typical transitive sentence has an SOV order and marks the different roles by means of endings:

(23) Khotanese, basic SOV order (Z 13.109)

*balys-ä*            *dāt-u*            *hvat-e*  
 Buddha-NOM.SG law-ACC.SG speak.PST-3SG.M  
 ‘The Buddha proclaimed (lit. spoke) the law’.

The subject is marked by an ending in the nominative case, the direct object receives an accusative ending, and the verb signals tense and person (and in this case also gender, being a past 3rd sg. ending). As regards the parametric settings, there are only minor differences compared with Young Avestan:

1. Gramm. Person (FGP): yes, Khotanese has verbal person and a three-way contrast in terms of demonstratives (Emmerick 2009, p. 387).

2. Gramm. Number (FGN): yes, there is a contrast between singular and plural (and traces of the dual). See the inflectional paradigms of Emmerick (1968).
3. Gramm. Gender (FGG): yes, masculine and feminine gender are systematically distinguished, and there are remnants of the neuter. See Emmerick (2009, p. 384) or Emmerick (1968, p. 249).
4. Feature spread to N (FSN): yes, nouns are marked for number and gender, and these features are thus ‘spread’ from D to N. Cf.

(24) Khotanese, feature spread to N (Z. 22.292; after Emmerick 2009, p. 405)

*ka nā parrījīyā dukh-yau jsa*  
 if them rescue.OPT.3SG woe-ABL.PL from  
 ‘If (only) he may rescue them from woes!’

5. number on N (bounded nouns) (FNN): yes, see FSN above.
6. Gramm. partial definiteness (DGP): no, even though demonstratives are frequently used, they are not obligatory. See, for instance, Z. 23.28 *balysu vāte jūhāte āṇi* ‘he is yearning for (*vāte*) the Buddha (*balysu*)’ – there is only one Buddha, this noun must be definite. Yet there is no definiteness marker. Another example: Z 23.22 *ku nā sata štā haṃbaḍa mātā* ‘when the full moon has not risen (lit. when not risen is full moon)’ – the moon is inherently definite, but there is no definite marking. Anaphoric marking of entities introduced earlier is possible, but not mandatory. Hence 14 other parameters are neutralized, viz. DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA.
7. Free null partitive Q (DPQ): no. There is no count/mass distinction by means of differential case marking. But note the fact that the nominative (instead of the accusative?) can be used to express measure (Emmerick 1989, p. 223).

8. D-controlled infl. on N (DIN): no, there is nothing reminiscent of nunation.
9. Plural spread from Cardinals (CPS): yes, cf. Šgs 3.5v3: *haudyau ratanyau* ‘seven jewels’, in which *-yau* is a plural ending (Degener 1993, p. 47).
10. Grammaticalized boundedness (CGB): no, singular count nouns do not appear to be unbounded. There is an indefinite article, but this does not necessarily imply that singular count nouns are unbounded when not used together with the indefinite article. Cf. e.g., Z. 23.131 *kye rre päte rrīṇa nā māta* ‘(nuns) whose father (is) a king, their mother a queen’, or (even though one could doubt that this is actually a count noun) Z. 22.235 *śśūjīye ggūšto hvaḍāndā* ‘they ate one another’s flesh (*ggūšto*)’. Count nouns with an unbounded reading get a plural ending/marking, cf. Z. 22.265 *daśa pale kṣattrā drjsāre* ‘they will hold banners, flags, umbrellas’ or Z. 23.126 *striye atā irate śśaṭṭhe* ‘women are very cunning, deceitful’. A negative value of CGB neutralizes parameter CCN.
11. Structured Adjective Phrases (AST): yes, cf. Sgh 7b1 *ttātāṇa balysāṇa rahāśśa* ‘in this Buddha-like sphere’ (with Dem–Adj–N).
12. Feature spread to structured APs (FFS): yes, cf. Sgh 36a5: *ttā biśśi dīra kädātāne* ‘all evil deeds’ (with Dem–Adj–Adj–N, and all elements inflected or marked for nom. pl.).
13. Feature spread to predicative APs (FSP): yes, cf.
  - (25) Khotanese, feature spread to predicative adjectives (Z. 4.33)
 

<i>āṣṣei’ṇā</i>	<i>hātenai</i>	<i>śśīyā</i>	<i>hamā</i>
blue.NOM.SG.M	red.NOM.SG.M	white.NOM.SG.M	same.NOM.SG.M
<i>hamau</i>			
goblet.NOM.SG.M			
‘blue, red, white (is) the same goblet’.			

14. DP over relatives (ADR): yes, the NP can be raised over relative clauses, i.e., relative clauses surface to the right of NPs. See Z. 23.23 *kṣīrā ku ne rre nāštā* ‘the land where there is no king’, and Emmerick (2009, p. 406) for other examples. This neutralizes parameter AER.
15. Free reduced RelCl (ARR): yes, there is the possibility of a genitivus absolutus, i.e., a free reduced relative clause formed with a participle in the genitive. Cf.

- (26) Khotanese, genitivus absolutus (Sgh § 43.1)  
*ttätäye hvatye hvanai*  
 this.GEN.SG spoken.GEN.SG speech.GEN.SG  
 ‘this spoken speech (= by/through/after this speech having been spoken)’.

16. Adpositional genitive (GAD): no, genitives are inflectional.
17. Free Gen. (GFR): yes, cf. the following example:

- (27) Khotanese, free recursive genitives (Z. 24.643)  
*mārīñi hīñi tcabaljāka*  
 Māra.GEN.SG army.GEN.SG scatterer.NOM.PL  
 ‘scatterers of the army of Māra’.

18. Uniform Gen. (GUN): yes, cf. Sgh 7a5: *gyastānu gyastā* ‘the god of the gods’ (Gen–N), compared with Z. 22.276: *śśāśiña balysä* ‘the Buddha’s Śāsna’ (N–Gen). Apparently, inflectional genitives are not bound to certain structural positions. Since this uniform genitive can also occur in object genitive position and function – GenO, cf. Z. 23.18: *dātā pvā’mata* ‘the hearing of the law’ (with Gen–N). –, parameter GFO is neutralized.
19. DP over free Gen. (GPR): unclear (probably no). I did not find a single example of a free, recursive right-branching genitive in the Khotanese texts surveyed for this study (and the grammars completely ignore this issue).

20. Gen-feature spread to N (GFS): no. In line with the Old Iranian languages (and Indo-European patterns in general), Khotanese genitives do not agree with their head nouns.
21. Adjectival poss. (APO): no. Possession is expressed via the existential verb + clitic forms of the personal pronouns (Emmerick 2009, p. 404), e.g., *ništā-mā* ‘I don’t have’.
22. Clitic poss. (PCL): yes, there are clitic forms of the personal pronouns used to express possession. See parameter APO.
23. N-feature spread to free Gen. (GSP): no. Nouns do not agree with genitives in terms of number or gender. See the example of GFR.
24. Adjectival Gen. (AGE): (yes). Khotanese can use adjectives formed with the suffix *-ĩña-* instead of genitives (Emmerick 1989, p. 225). Besides that, the later texts have the frequently used possibility to reinforce a genitive by means of *hĩvia-* ‘suus, -a’, e.g., in *śakrrā hĩvī bārai* ‘the steed of Śakra’, but the genitive itself must nonetheless be marked on the dependent nominal argument (Emmerick 1989, p. 223). But since APO is set negatively, AGE, strictly speaking, becomes irrelevant.
25. Poss.-checking N (GCN): no. There is no head marking on nouns when they occur with genitival arguments.
26. Strong partial locality (TPL): yes, demonstratives usually surface in the D-area, introducing NPs. Cf. e.g., Śgs 2.5r5: *ṣā aysmū-vaṣṭāmata* ‘this Samādhī’.
27. Strong locality (TSL): no, not all demonstratives occur in the D-area (Degener 1993, p. 48). Cf. Sgh 7a6: *biśśi ttā uysnora* ‘all (these) beings’.
28. N over Demonstratives (NDE): no, demonstratives always precede their nouns; N cannot be raised (Degener 1993, p. 47). See TPL.

29. N over Cardinals (NOC): yes, *contra* Degener (1993, p. 47), nouns can be raised over cardinals. It is true that many NPs evince the order Card–N, such as Šgs 3.5v3: *haudyau ratanyau* ‘seven jewels’. However, there are counterexamples. Consider e.g., Z. 22.262: *gyastānu rrunde tcohora* ‘the four kings of the gods’ (with Gen–N–Card), or Z. 22.162 *ratanyau haudyau* ‘(full of) seven jewels’.
30. N over Ordinals (NOO): yes, apparently, cf. Z. 22.157 *hīñe ratani haudamä* ‘the seventh jewel of the army’ (with Gen–N–Ord).
31. N over adjectives (NOA): yes, nouns can be raised over adjectives, cf. Nanda 7: *hārū ye mistä* ‘one/a great merchant’ (with N–Idf–Adj). The other order, Adj–N, is attested, for instance, in Nanda 42–43 *avījsyaca strriya* ‘blind woman’. See Degener (1993, p. 47) or Heston (1976, pp. 6, 35). An illustrating example, showing both orders, can be seen in Z. 22.333: *hvataimū baśde māštä kye māstu byanu yanīndi* ‘I have told you of the great evils (*baśde māštä*, N–Adj) which cause great obstruction (*māstu byanu*, Adj–N).’
32. N over GenO (NGO): yes, even though only rarely attested, the mere possibility to raise N over objective genitives is there. See Emmerick (2009, p. 400). An example is Z. 23.11 *paramārthä dātä* ‘the Paramārtha of (= pertaining to) the Law’.
33. N over external arguments (NOE): yes, N can be raised over some arguments (adjectives, genitives, and also appositional material, such as names or titles, e.g., *gyastä balysä* ‘the Lord Buddha’).
34. Verbless relative clauses (VRC): no, apparently relative clauses are consistently built with finite verbs (even the copula or negated copula).
35. Facultative linker (FLI): no.
36. Generalized linker (GLI): no.



### 3.4.2 Sogdian

Sogdian (Sogd.) was the native language of Sogdia or Sogdiana, the ancient territory around Samarkand and the Zerafshan river. Sogdian was, however, also a *de facto* lingua franca of large parts of Central Asia. It was once in wide-spread use along the north-eastern part of the Silk Road(s), and most texts of the language are from Sogdian merchant colonies in the Tarim Basin and the Gansu corridor; the most prominent documents from Sogdiana proper are those from Mount Mugh. As a result of the cultural impact of the earlier Achaemenid Empire, Sogdian texts are mainly written in three scripts which are all derived from the Aramaic consonantal script, namely the Sogdian, Manichaean and Syriac scripts. There were different literary traditions, among which a native Sogdian one, a Christian one, a Manichaean and a Buddhist one of which most material is preserved (prominently in Turfan and Dunhuang). The rather uniform Sogdian preserved in these different literary traditions has no living descendants, but the modern language Yaghnobi, spoken in the Yaghnob valley in Tajikistan and differing in a few points from its Middle Iranian cousin Sogdian (Yaghnobi has an imperfect stem built with an augment *a-* and a 3rd plural ending *-or* not found in Sogdian), is a close relative.

Sogdian has three grammatical genders (masculine, feminine and remnants of the old neuter) and two regular numbers marked in NPs (singular and plural). Additionally, there is a separate number, the so-called numerative (etymologically the old dual, but decoupled from its restricted function), which is only used with elements immediately following numerals higher than ‘one’ (Sims-Williams 1979). The most prominent feature of Sogdian concerning its nominal morphosyntactic properties is its double system of ‘light’ and ‘heavy’ stems (on which see Sims-Williams 1982). The ‘light’ stems distinguish – at least in the singular – six cases (nominative, accusative, genitive-dative, instrumental-ablative, locative, vocative), whereas the ‘heavy’ stems only have a contrast between a direct and an oblique case. Yet in light of the parametric values covered by my study, this distinction between two inflectional patterns appears to be irrelevant,

as the parameters do not ask for case distinctions but other features.

Sogdian markedly differs from all three Old Iranian languages and from its contemporary Middle Iranian relative Khotanese by having developed a rich system of articles similar to that of modern western European languages (see Wendtland 2011 for an exhaustive discussion), or similar to the system of Classical Greek. Indefinite nouns can be used without an article, as in (ə)pyār xwaβn-u wēt(u)δāram ‘last night I saw a dream’ (Vessantara-Jātaka, fol. 3,1; see Benveniste 1946), in which the ‘dream’, *xwaβnu*, receives an accusative ending but no article or determiner. However, if the noun is already known from previous discourse and/or definite, it will obligatorily receive an article. Cf. the following sentence (in transliteration:

(28) Sogdian articles (Yoshida 2009, p. 305)

<i>xu</i>	<i>zātē</i>	<i>əwən</i>	<i>əptre</i>	<i>andēk</i>	<i>δāre</i>
the.NOM.SG	son.NOM.SG	the.GEN.SG	father.GEN.SG	custom	have.OPT.3SG

‘The son should have his father’s characteristics’.

As regards the parametrical settings, the major difference as compared with the three Old Iranian languages and Khotanese concerns parameter DGP, as it is not set negatively in Sogdian and hence does not neutralize 14 other parameters.

1. Gramm. Person (FGP): yes, verbal person and a three-way contrast in terms of determiners/demonstratives (Wendtland 2011) show that Sogdian has grammaticalized person.
2. Gramm. Number (FGN): yes, Sogdian distinguishes between singular and plural and thus has two numbers. Sims-Williams (1989, p. 183) asserts that Sogd. has three numbers, but the so-called numerative, going back to the old dual, is only used in a very restricted sense, namely with nouns following numerals. It is thus not a general number like singular or plural which apply to all nouns or adjectives.

3. Gramm. Gender (FGG): yes, Sogdian has two genders (masculine and feminine) and vestiges of a third (neuter). See Sims-Williams (1989, p. 183).
4. Feature spread to N (FSN): yes, plurality is marked on (and hence ‘spread’ to) nouns. Cf. the following clause from the ‘Story of the Pearl-Borer’:

(29) Sogdian, morphological plural marking on nouns, ‘Story of the Pearl-Borer’, clause 11, Manichean version (following Yoshida 2009, p. 332)

*art-šu čakanāč piḍār maryār-t nē framāye suḅte*  
 and-him from.what sake pearl-PL not order.PRS.2SG bore.INF  
 ‘So why did (lit. do) you not order him to bore (the) pearls’.

As is well-known, the Sogdian plural morpheme *-t-* goes back to an old collective or abstract built with *\*-tā-* (Sims-Williams 1989, p. 183), thus a synchronic plural *βrātart* ‘brothers’ originally meant ‘brotherhood, fraternity; Bruderschaft’. There are only vestiges of the old plural, e.g., in *βayān* ⟨βγ’n⟩ ‘of/from the gods’ (old gen. pl. < *\* bagānām*). In any case, number is ‘spread’ to nouns, i.e., marked and overtly expressed.

5. number on N (bounded nouns) (FNN): yes, morphological marking by means of *-t-*, see FSN above.
6. Gramm. partial definiteness (DGP): yes, in environments where the designatum has previously been mentioned, articles (with a three-way contrast in terms of deixis) are used to signal (anaphoric) definiteness. See Wendtland (2011) for an exhaustive treatment of Sogdian weak demonstratives and articles. For example, consider the main part of the ‘Story of the Pearl-Borer’ in Henning (1945, pp. 466–469) (reproduced in Yoshida 2009, 330ff.), in which ‘the judge’, when first mentioned, surfaces as a bare noun *əxtu* (⟨xtw⟩ in Manichaean script [l. 16], ⟨’γtw⟩ in Sogdian script [l. 2]), whereas in a later part of the text, he receives an article (of the 2nd-person/*Du*-deixis set), viz. *šu* (ə)*xtu* ‘the judge’ (Manichaean ⟨-šw xtw⟩ [l. 34], Sogdian ⟨-šw ’γtw⟩ [l. 14]). In doing so, Sogdian differs from the three

Old Iranian languages and from Khotanese, and the 14 parameters that were otherwise neutralized must be discussed and set.

7. Gramm. definiteness (DGR): yes, cf. the example above (nr. 28) in the introductory remarks. Articles in Sogdian can function as definite generics ('the son', 'the father's') or can be used also without previous mentioning of the designatum to signal its definiteness; as is the case, for instance, in the main part of the 'Story of the Pearl-Borer' (following Henning 1945, pp. 466–469), where right at the beginning of the text, in line 15–16 of the Manichaean Sogdian version, 'the hired laborer' surfaces as Sogdian *xu marāz martī* (<-xw mr'z mrtty>; lit. 'the laborer man').
8. Strong Person (NSD): yes, this seems to be the case (the parameter asks whether referential nominal material, such as proper names, surfaces in the D-area). See Yoshida (2009, p. 314), with the example *smēr yari* 'Mount Sumeru'.
9. Free null partitive Q (DPQ): no, there are no count/mass distinctions by means of differential case marking.
10. Gramm. distal art. (DDA): yes, there is a three-way distinction in definite articles (Yoshida 2009, 291ff.). There is likewise one in demonstratives, but according to the defining question of this parameter, parameter DDA must get a positive setting. This is an interesting observation, linking Sogdian to Wolof or Basque rather than to most modern western Indo-European languages with articles (but note the Armenian article *-s/-d/-n* with a three-way contrast).
11. Def.-checking N (DCN): no, the definiteness marker is a separate article surfacing in the D-area and not a bound morpheme cliticizing on N. A negative value of DCN neutralizes the dependent parameter DSN.
12. Def. on relatives (DOR): no, the introducer of a relative clause, i.e., a relative pronoun, must not agree in definiteness with its nominal head.

See e.g., *yunē čakraβart čintāmani dārani ke-ti əzu parβērāt-δārām* ‘this Čakravart Čintamani spell **which** I (have) explained’ (Yoshida 2009, p. 318).

13. D-controlled infl. on N (DIN): no, nothing like the Arabic nunation exists in Sogdian.
14. Plural spread from Cardinals (CPS): (yes). As stated in the introductory remarks, Sogdian has a special numerative (Sims-Williams 1979). The defining question of this parameter is whether or not nouns, when introduced by a non-singular cardinal numeral, must agree with it or not. Since the feature of number leads to a special marking of a noun (or adjective etc.) following a cardinal numeral higher than ‘one’, I would opt for a positive value of this parameter, even though Longobardi et al. termed it ‘Plural spread from Cardinals’. I would rather interpret this in a more abstract fashion, viz. whether or not the feature of non-singular number is marked on elements following cardinal numerals. This is, then, the case in Sogdian by means of a special ‘numerative’ ending restricted to this function. An example can be seen in *čاتفār δβar-a* ‘four doors.NUM’. Note that also plural forms occur, e.g., *δwāts δβar-ta* ‘twelve doors’ (Yoshida 2009, p. 313). Finally, also singular marking occurs: *əβt paxarē* ‘seven planets (lit. planet)’ besides *əβt paxarē-t* with the plural morpheme. In any case, non-singularity *can* be marked on nouns when occurring with numerals.
15. Grammaticalized boundedness (CGB): no, singular count nouns are not inherently unbounded and do not require an indefinite marker to signal boundedness. See, for instance, the following clauses:

## (30) Sogdian bounded singular count nouns

- a. ‘Story of the Pearl-Borer’, line 28–29, Manichean version (after Yoshida 2009, p. 331; line 1: transcription, line 2: transliteration of Manichaean Sogdian script)

*art-mī čānō kū xānā sār šīkar*  
 rtmy c’n’kw kw x’n’kh s’r šykr  
 and-me when to house towards lead.IPRF.3SG  
 ‘when he had taken me to (his) house’.

- b. Vessantara-Jātaka, fol. 3,1 (Benveniste 1946)

*(ə)pyār xwaβn-u wēt(u)δāram*  
 ’py’rh xwβnw wytwδ’r’m  
 last.night dream-ACC.SG see.PRET.1SG  
 ‘last night I saw a dream’.

A negative value of CGB neutralizes parameter CCN.

16. Strong article (CGR): no, there is no indefinite article. Indefinite pronouns exist, but no real articles in the sense of this parameter. See the two examples of parameter CGB, directly above.
17. null-N-licensing article (DNN): unclear. This parameter asks whether a determiner identical with the simple article (if present in the language) is used to introduce definite nounless argument nominals (a genitive argument, an adpositional argument or a relative clause). Neither are bare adpositional arguments introduced by means of an article, nor are relative clauses. However, there are so-called fused prepositions (see Yoshida 2009, pp. 293–294 or Sims-Williams 1989, p. 186) which could allow for a positive setting of this parameter (though I am not quite sure, because they appear to be optional, and most adpositional phrases are bare and do not contain adpositions fused with articles). Nounless genitive arguments are difficult to find in the texts (I found no examples).
18. Structured Adjective Phrases (AST): yes, cf. *mana əwī xwārant əpkašya* ‘(into, LOC) my right side’ (Yoshida 2009, p. 305).

19. Feature spread to structured APs (FFS): yes, consider the following example from Yoshida (2009, p. 314): *ō tawa roxšn-u čašm-u* ‘your light eye’ (D–Pro–Adj–N), in which the determiner/article, the adjective and the noun are all inflected for accusative case. Hence the feature of case (and number) is spread from N (in an NP analysis) or D (in a DP analysis) to the structured adjective.
20. Feature spread to predicative APs (FSP): yes, cf. Vessantara-Jātaka, fol 3, 10–11: *⟨xw k’δy wyšy ’β⟩ əxū kādi wəyaš-i (ə)βa* ‘he became very happy’ (with D–Adv–Adj–V) and an adjective inflected for masc. nom. sg.
21. D-controlled infl. on adjectives (ADI): no. Interestingly, there are instances in which structured adjectives, i.e., those occurring between D and N, are inflected or rather agree with their nominal (or determiner, depending on analysis) head. See parameter FFS for an example. However, many other NPs/DPs do not show overt inflectional marking of adjectives surfacing between a determiner and noun (see e.g., the example of parameter AST). Since there is variation, and adjectives still can get marking, the parameter must be set negatively, but the question arises whether the development would eventually have led to a situation comparable to that of Germanic with the weak and strong inflection of adjectives depending on the NP/DP context (overt D or not).
22. DP over relatives (ADR): yes, relative clauses always follow their head, or in other words, the NP can be raised over relative clauses. See Yoshida (2009, p. 318) or the example of parameter DOR above. Since NPs surface before relative clauses, parameter AER is neutralized.
23. Free reduced RelCl (ARR): yes, the criterion of this parameter appears to be fulfilled by means of infinitives and verbal nouns (see Yoshida 2009, pp. 320–321). One example: *čan murtu kārī* ‘after having died’.
24. Adpositional genitive (GAD): no. Genitives are inflectional, either marked

on the noun or on a determiner introducing the genitive (as can be seen in the example given for parameter GFR).

25. Free Gen. (GFR): yes, cf. the following example:

- (31) Sogdian free recursive genitives, Vessantara-Jātaka, fol 3,25 (Benveniste 1946)

<i>ti</i>	<i>ō</i>	<i>anxar</i>	<i>əwī</i>	<i>xwatēn</i>	<i>kaθarē</i>
ZY	ZKw	'nrx	'wyh	xwt'ynh	kδ'r'k
and the.ACC.SG star the.GEN.SG queen belly/stomach					
'and (study) the star of the belly of the queen'.					

Apparently, there are two genitives further specifying the nominal head noun 'star', but one of the two is in a dependent relation to the other ('of the queen's belly' or 'of the belly of the queen').<sup>20</sup> Both receive, however, no case endings and instead are introduced by one common determiner/article in the genitive signaling the genitive function of these two nouns.

26. Uniform Gen. (GUN): yes, there is an inflectional genitive not bound to certain structural positions. Cf. *xu maryārtī xēpθāwand* 'the owner of the pearls' with D–Gen–N order (Yoshida 2009, p. 316) besides the example of GFR above evincing a D–N–Gen–Gen order. Another example:

- (32) Sogdian, embedded free genitives (uniform genitive; example from Yoshida 2009, p. 297)

<i>fritāt</i>	<i>ke</i>	<i>wānī</i>	<i>sāt</i>	<i>šire</i>	<i>əktyə</i>
love	REL	the.GEN-DAT	all	good.GEN-DAT	deed.GEN-DAT
<i>parwēžnē</i>					
nourish.NOM.SG					
'love which is a nourisher of all the good deeds'.					

This presence of a uniform genitive not bound to structural positions neutralizes parameter GFO.

<sup>20</sup> Another interpretation might regard this as an adjectival relation, with 'queen' being understood as an adjective ('royal belly').



27. DP over free Gen. (GPR): yes, see the example of GFR.
28. Gen-feature spread to N (GFS): no. Genitives are dependent NPs with internal phi-feature agreement (if overtly expressed), but they do not agree with their head nouns.
29. D-checking possessives (PDC): no. Possession is expressed via (1) a noun inflected for GEN.-DAT. (or OBL.), frequently together with a non-clitic personal pronoun (as in the example of parameter AST), (2) *δār* ‘to have, hold’ or by means of (3) possessive pronouns which usually combine with the copula (‘mihi est’) (Yoshida 2009, p. 307). In the case of the latter, i.e., if possessive pronouns are used, they will attach to the first stressed item in a clause (Wackernagel position). One example (l.c.): *rti-mī nēst waδu nē zātē nē δuyta* ‘I have no wife, son or daughter’ (lit. ‘and-for.me not.is wife, not son, not daughter’). Hence possessives do not occur in D-position within their NP and the parameter is to be set negatively.
30. Adjectival poss. (APO): no. As just stated, there are clitic possessive pronouns, but no adjectival possessives. This neutralizes parameter AGE.
31. Post-affix poss. (PAP): irrelevant. This parameter asks whether or not a language has a special position for possessives after a suffixed article or not. Since the article in Sogdian is not suffixed, but occurs in the D-area, this parameter becomes irrelevant or is neutralized.
32. Clitic poss. (PCL): yes. This parameter defines whether or not a language allows for possessives as bound morphemes cliticizing on the noun or another stressed item or not. In Sogdian, there are clitic pronouns indicating possession that attach to the first stressed element or phrase in a clause (Wackernagel clitics), i.e., usually (though not always) the sentence connector or conjunction *rti*. See Yoshida (2009, p. 307) or parameter PDC above.

33. N-feature spread to pron. poss. (PHS): yes, a feature like number (sg. or pl.) is expressed by means of the selection of the respective pronoun. See the examples of parameters AST (which had a genitive sg. pronoun) and FFS (also sg.) above. An example with plurality can be seen in *xānd āfrītēt ōtākt ke ʔti-šan sāk ʔti patšmār nēst* ‘those blessed places whose number and counting do not exist’ (lit. ‘**which** do not have...’) (Yoshida 2009, p. 318).
34. N-feature spread to free Gen. (GSP): no, As stated before, genitives constitute NPs on their own with internal agreement, but they do not agree with their head noun.
35. Poss.-checking N (GCN): no, there is no distinctive morphological marking on the noun when occurring with a genitive argument.
36. Strong partial locality (TPL): yes, all demonstratives occur in the D area, though after the actual determiner. Consider *xu xunax ʔrīw* ‘that body’, lit. ‘the that body’ (Yoshida 2009, p. 314).
37. Strong locality (TSL): yes, the system is uniform, all demonstratives are moved to the D area.
38. D-checking demonstratives (TDC): no, demonstratives co-occur with definiteness-compatible articles. See the example of parameter TPL above.
39. N over Demonstratives (NDE): no, this seems to be impossible. Consider e.g., *xānd āfrītēt ōtākt* ‘those blessed places’ (Dem–Adj–N), following Yoshida (2009, p. 318).
40. N over Cardinals (NOC): no, nouns apparently cannot be raised over numerals. See e.g., *ʔwən šē ratne* ‘the three jewels’, with D–Card–N (Yoshida 2009, p. 318).
41. N over Ordinals (NOO): no, this seems to be impossible. Consider e.g., *pīšt čan xēpθ kaštār satu βrēwarmīk βantē* ‘sent from your most trifle

**one millionth servant**’ (lit. ‘hundred ten-thousandth’), following Yoshida (2009, p. 325).

42. N over adjectives (NOA): yes, cf. *ō satu δīnār zern* ‘the hundred gold denars’ with an order D–Card–N–Adj (Yoshida 2009, p. 316).
43. N over GenO (NGO): yes, nouns can be raised over the uniform genitive (which incorporates the notion of an object genitive), cf. the example of parameter GFR.
44. N over external arguments (NOE): yes, see parameters NGO and NOA above.
45. Definiteness on APs (DOA): no, see the example of parameter NOA above.
46. Verbless relative clauses (VRC): no, Sogdian relative clauses apparently always (?) have a predicate, even when built with the copula (e.g., *ke ʔti-šan sāk ʔti patšmār nēst* ‘which do not have number and counting’, the example of parameter PHS).
47. Facultative linker (FLI): no.
48. Generalized linker (GLI): no.

### 3.4.3 Chorasmian

Chorasmian (Cho.; also Choresmian, Khwarezmian vel sim., autonym ⟨zβ'k'y xw'rm⟩) was the language of the inhabitants of Chorasm, the territory to the south of the Aral Sea, alongside the lower Amu Darya. Apart from scanty remains written in a derivative of Imperial Aramaic script (all dated to various points in the 1st millennium CE; for instance, the Isakovka inscriptions, see Livšic 2003), Chorasmian is mostly known from an interlinear translation of a medieval Arabic text written in the year 1135 (the so-called *Muqadimmat al-Adab*, abbreviated Muq.; see Benzing 1968, 1983) and from texts dating to the 13th and 14th centuries CE (*Yatīmat ad-dahr*, abbreviated YD, see Zekī Walidī 1930, as well as the *Qunyat al-Munya*, abbreviated Q, and the *Risāla*, abbreviated R; see MacKenzie 1990), i.e., centuries after the Islamization. These texts are written in Arabo-Persian script, but they cause problems in correct reading and interpretation due to the fact that diacritic dots often are missing. Despite this late attestation, Cho. counts as Middle Iranian, since it closely resembles Sogdian in its morphosyntactic features (presence of articles + case endings, *-ā-* and *m-*augment etc.). Chorasmian has no living descendants (though it shares some features with Sangesari, see Azami and Windfuhr 1972, 15ff., 36).

1. Gramm. Person (FGP): yes, person is grammaticalized in both the verbal domain and demonstratives (distinction between near and far deixis as well as a topical, deictically neutrally pronoun), see Durkin-Meisterernst (2009, 344ff.).
2. Gramm. Number (FGN): yes, singular and plural are systematically distinguished and there are traces of the old dual (inter alia functioning as a so-called 'numerative'). See Humbach (1989, p. 197).
3. Gramm. Gender (FGG): yes, there are two genders (masculine and feminine), see Humbach (1989, p. 196).
4. Feature spread to N (FSN): yes, gender, number and case are marked by

inflectional endings occurring on the noun. See e.g., *ī δasc-ina* ⟨'y δscn⟩ 'the hands' (Muqaddimmat 387.3, following Durkin-Meisterernst 2009).

5. number on N (bounded nouns) (FNN): yes, see the example of FSN.
6. Gramm. partial definiteness (DGP): yes, even though this is somewhat difficult to demonstrate. Parameters DGP and DGR are, in the words of Longobardi and colleagues, 'gemini' parameters. Parameter DGP defines whether or not definite nominal elements must be formally marked in those cases in which they designate an entity already introduced in the previous discourse. Parameter DGR, on the other hand, asks whether definiteness must be marked in *all* contexts and is thus a generalization of DGP. If definiteness is marked irrespective of context, this will imply that it will also be marked in anaphoric contexts, so that a positive setting of parameter DGR (generalized definiteness marking) simultaneously includes the smaller 'anaphoric' definiteness marking defined by DGP. Chorasmian has such a generalized definiteness marking, as will be discussed with parameter DGR immediately below. Therefore also parameter DGP gets a positive value. Note that the extant Cho. texts consist of glosses in an Arabic dictionary and short commentaries in legal documents – a fact that does not facilitate efforts to demonstrate presence (or lack) of anaphoric definiteness marking depending on discourse context.
7. Gramm. definiteness (DGR): yes, as just described, Chorasmian has a definite article (masc. sg. *ī* ⟨'y⟩, fem. sg. *yā* ⟨'y'⟩, dual *yā* ⟨'y'⟩, plural masc. and fem. *ī* ⟨'y'⟩, but note the homonymy and homography). An example of their usage can be seen in the following sentence:

- (33) Chorasmian definite articles (QR 413, following Durkin-Meisterernst 2009, p. 373)

*f-ī    nān šūrat ka    yā    \*βasanjiwān-a ī    juft    δāra wašī*  
*f.y    n'n šwrt k.    y'    βsncw'n    'y jft    δ'r wšy*  
*in.the that case that the \*testatrix    the husband DO    trustee*  
*wāśeda*  
*w'yd*  
*appoint.IPRF.3SG.*  
 'In the case that the \*testatrix appointed (her) husband as trustee'.

8. Strong Person (NSD): (yes). I found no instances of referential material surfacing in the D-area (and the grammatical sketches omit the topic), but since there is a parametric hierarchy between parameters NDS and DNN (the latter depends on the possibility for languages on 'strong person') and DNN has a positive value in Cho. (see below), referential nominal material (e.g., names) can be expected to occur in the D-area. Other positioning (reminiscent of ancient Greek): Q 85 <k's 'y mrc 'y xw'rm ws'c> 'if the Chorasmian man should say' (lit. 'if the man, the Cho. [one]').
9. Free null partitive Q (DPQ): no, there are no examples indicating that this parameter should get a positive value. Count/mass distinctions are signaled by means of an indefinite marker (see parameters CGB and CGR below).
10. Gramm. distal art. (DDA): no, the definite articles do not encode an obligatory distal/proximate distinction. This is a difference in comparison with Sogdian.
11. Def.-checking N (DCN): no, the definiteness marker is no bound morpheme cliticizing on N, it occurs separately in the D position introducing NPs/DPs, see e.g., D-Ord-N in QR 233: <'y ftmyck srδ> 'the first year' (Durkin-Meisterernst 2009, p. 358).

12. Def. spread to N (DSN): no, there are no additional definiteness suffixes on N in addition to free definite morphemes in D.
13. Def. on relatives (DOR): no, Cho. relative pronouns are not marked for definiteness. See QR 368 <'y šyl ny k.m...> 'this business (*ī šuyl nē, ki-mi*) which/that I...'.
14. D-controlled infl. on N (DIN): no, the inflectional form of N does not depend on the occurrence of certain morphemes in D beyond simple phi-feature concord.
15. Plural spread from Cardinals (CPS): (yes), as in Sogdian, there is a separate numerative (< dual) form used with elements following cardinals (with '3' and '4'), see Durkin-Meisterernst (2009, p. 343). Even though not plural spread in the strict sense, it is nevertheless numerical (= non-singular) feature spread and morphological marking on N. Hence the parameter must get a positive value. But genuine plural forms occur as well, e.g., Q. 21 <hz'r pr'cn> 'a thousand divorces'.
16. Grammaticalized boundedness (CGB): yes, an unbounded reading of nouns can be marked by the absence of the definite article and by N being raised over adjectives, see parameters NOA and DOA below for examples. If the bounded reading of a singular count noun shall be marked, *ēw-* <'yw> 'one' can be used (Durkin-Meisterernst 2009, p. 343), cf. M. 50.3 <'yw cm> 'one eye, an eye'.
17. Strong article (CGR): (yes), but since CGB is set positively, this parameter is predictable or irrelevant. Indefinite nominals mark the count reading by means of a determiner-like element homologous with the numeral 'one', see CGB above.
18. Bounded-checking N (CCN): no, the marker for bounded indefinite readings is a separate word (<'yw> 'one') and no enclitic affix.

19. null-N-licensing article (DNN): (yes), this parameter asks whether or not a determiner identical with the simple article is used to introduce definite nounless argument nominals (a genitive argument, an adpositional argument or a relative clause). With adjectives, this is clearly the case, cf. <'y hzwr> 'the good one' or 'what is good' (Humbach 1989, p. 196). Another example (M. 33): <'y frdm fy mrskwnd> 'the best (<frdm>) among men'. However, there appear to be no examples of definite nounless genitives, adpositional arguments or relative clauses.
20. Structured Adjective Phrases (AST): yes. Cf. M. 38.1 *yā cafār(a) pāδ-a spēdec-a* <'y' cf'r p'δ spy<sup>dy</sup>c> 'the four white legs' which attests an order D–Card–N–Adj (Durkin-Meisterernst 2009, p. 347).
21. Feature spread to structured APs (FFS): yes, see the example of AST above, in which the adjective, being NP-final (due to N-raising) is inflected for plural. Another example: M 117 <'ps bz'β'r> 'a thin sheep'. See also Durkin-Meisterernst (2009, p. 358).
22. Feature spread to predicative APs (FSP): yes, cf. QR 69 *ufān-ya ka-mi ḥalāl-a meyta* <'wf'ny k.my ḥl'l m'yṯ> 'when she was lawful (*ḥalāl-a*) to me' (Durkin-Meisterernst 2009, p. 366).
23. D-controlled infl. on adjectives (ADI): no, this appears not to be the case. Even though adjectives frequently do not show overt case endings or rather agreement with their nouns, they sometimes do. The irregular appearance of case endings probably is an artifact of the defective writing system and inflection was regular and obligatory across the NP, affecting all inflecting elements.
24. DP over relatives (ADR): yes, NPs can be raised over relative clauses, i.e., relative clauses can surface to the right of an NP: Cf. e.g.,



- (34) Chorasmian relative clause (QR 368, following Durkin-Meisterernst 2009, p. 365)

*ī šuyl nē(n), ki-mi ...*  
*'y šyl ny, k.my ...*  
 the business this REL-A  
 'This business that I ...'

This neutralizes parameter AER.

25. Free reduced RelCl (ARR): yes. No detailed information in either Durkin-Meisterernst (2009) or Humbach (1989), but there are examples like M. 465.1 <'xb γ'ryk ms'nyd> 'he spent the night keeping watch (γ'ryk)'.
26. Adpositional genitive (GAD): no, genitives are inflectional. Adpositions (postposition *δāra* <δ'r<sup>a</sup>> or preposition *paš* <pš>) can be used together with the inflectional genitive markers, but they do not replace the actual endings.
27. Free Gen. (GFR): unclear, there appear to be no recursive genitives in the Muq., only cases like M. 463.1 <brywβnycn p'rwzd 'y nwk 'y mws'k'n> 'the pin end of the tooth pick frayed'.<sup>21</sup>
28. Uniform Gen. (GUN): (yes). Genitives are inflectional and not bound to fixed structural positions, i.e., they can be separated from their head noun by intervening material. For instance, an order N-Gen is attested in M 188 <y' γwx 'ps'n> 'the sheep's ear' (with a masc. gen.) or M 51 <y' δ'rk ... y'psy> 'the sheep's udder' (with a fem. gen.). Likewise: 354 <bfnykn ... 'y βwmn'n> 'creator of the earths'. A genitive between D and N is found in M 103.6 <c.' mδ'n-y.h k'm> 'from (the inside of) his mouth'. A positive value of this parameter neutralizes GFO. However, a uniform genitive is

<sup>21</sup>The data gap with regard to genitives could not be filled even by an extensive search in the Muqaddimmat (Benzing 1968). Maybe I overlooked instances, or the lack of recursive genitives relates to the nature of the text (glosses). NPs which might, perhaps, be expected to contain recursive genitives often are expressed via only one word or two. Consider e.g., M. 2.2 <pcβ'rd> 'man of the sister of the wife'.

an extension of a free genitive, and strictly speaking, no instances of recursive genitives can be presented. Therefore this parameter should be left unspecified.

29. DP over free Gen. (GPR): unclear due to a lack of data.
30. Gen-feature spread to N (GFS): no. Genitives do not influence their head nouns in terms of gender or number.
31. D-checking possessives (PDC): no, this seems to be impossible. Possession is expressed via clitic pronouns which do not occur in the D-area, e.g., <y' xwb-<sup>a</sup>'-**m**<sup>i</sup> δ<sup>u</sup>γ<sup>o</sup>d<sup>a</sup>> 'my beautiful daughter' (QR 120, following Durkin-Meisterernst 2009, p. 359).
32. Adjectival poss. (APO): no, possessives are clitics on N or adjectives, as seen in the example above (PDC). This neutralizes parameter AGE.
33. Post-affix poss. (PAP): irrelevant, since the definite article is not suffixed to N, but surfaces in the D-area.
34. Clitic poss. (PCL): yes, cf. *mahr-hi* <m<sup>a</sup>hr.h> 'her dowry' with *-hi* (Durkin-Meisterernst 2009, p. 373). Another example would be *yā δuyd-a-mi* <y' δγ<sup>o</sup>d-<sup>a</sup>'-m<sup>i</sup>> 'my daughter' (QR 12, following Durkin-Meisterernst 2009, p. 359).
35. N-feature spread to pron. poss. (PHS): yes, singular or plural as well as person are marked by different sets of clitic possessive pronouns (on which see Durkin-Meisterernst 2009, p. 345).
36. N-feature spread to free Gen. (GSP): (no), non-free genitives do not agree with their head nouns in terms of number or gender. But since no clear examples of free, recursive genitives can be presented, this parameter must be left out of the discussion.

37. Poss.-checking N (GCN): no. There is no ‘construct state’ vel sim. as in Semitic, i.e., distinctive morphological marking on N when occurring with a genitive argument.
38. Strong partial locality (TPL): (yes), one demonstrative pronoun is attracted to the D-area, viz. *nān* ⟨n’(n)⟩ (see NDE below). However, it is a precondition for the setting of this parameter that CGR, ranking higher in the hierarchy, must *not* be positive (and it is positive in Cho.). Therefore this parameter is neutralized or predictable.
39. Strong locality (TSL): (no), not all demonstratives surface in the D-area; in fact, most do not. See NDE below. But again, this is actually a neutralized parameter.
40. D-checking demonstratives (TDC): no, demonstratives in the D-area co-occur with definiteness-compatible articles, see the first NP in example 33.
41. N over Demonstratives (NDE): yes, only *nān* ⟨n’(n)⟩ precedes the noun, all other demonstratives follow (Durkin-Meisterernst 2009, pp. 358–359): QR 363 *ī sukund nē(n) <y sknd ny>* ‘this oath’ (D–N–Dem); same order in M. 65.5 *yā dēn-a nēn-a <y’ δy nyn>* ‘this woman’ (haplology?); QR 281 *ī kt-ina nāw-i <y’ ktn n’w<sup>i</sup>>* ‘these deeds’ (D–N–Dem). In contrast, preposed *nān* in QR 115 *yā nān-ya hunān-ya dāra <y’ n’ny hwn’ny δ’r>* ‘for that slave-girl’ (D–Dem–N–Postposition).<sup>22</sup>
42. N over Cardinals (NOC): yes. In M. 38.1 *yā cafār(a) pāδ-a spēdec-a <y’ cf’r p’δ spy<sup>dy</sup>c>* ‘the four white legs’ an order D–Card–N–Adj is attested (Durkin-Meisterernst 2009, p. 347). Similarly, M. 483.7 *yā-(a)δw(-a) xsin-a <y’.δw xs<sup>y</sup>n>* ‘the two things’. Durkin-Meisterernst (2009, p. 347) cryptically writes that “[s]everal isolated examples appear to diverge from the paradigms which cannot be detailed here,” but does not explain how this

<sup>22</sup>Note that also verbs can be raised to a high position in Chorasmanian, as becomes clear from examples like M 274 ⟨m’sd pry mrc> ‘he came upon the man’.

divergence manifests itself. Humbach (1989, 196ff.) likewise yields no examples of a potential D–N–Card (or D–N–Ord) order. Bearing in mind the N-raising hierarchy of parameters, and given the fact that nouns are regularly ‘raised’ over higher-ranking demonstratives, one could expect nouns to have the possibility to be ‘raised’ also over numerical elements (both cardinals and ordinals). And in fact one such instance may be seen in C2311/200b (following Benzing 1983, p. 129), <’y š’f’r ’yw> ‘one witness, the one witness’.

43. N over Ordinals (NOO): unclear. The normal pattern seems to be Ord–N, as in, for instance, QR 233: <’y ftmyck srδ> ‘the first year’ (Durkin-Meisterernst 2009, p. 358). I checked the entries for all ordinals in Benzing (1983). Perhaps this lack of data is due to the nature of the material – the parametric hierarchy predicts that nouns should have had the ability to be raised over ordinals.
44. N over adjectives (NOA): yes, <’sp ’zr> ‘old horse’ (M. 166.4), or QR 338 <r’c wdncy> ‘old veins’ with N–Adj (both examples from Durkin-Meisterernst 2009, p. 358). If the noun is preposed (or ‘raised’), the NP will get an indefinite reading. See also parameter DOA below for NPs with N–Adj order and definite reading.
45. N over GenO (NGO): yes, genitives very often surface to the right of nouns (see parameter GUN for examples).
46. N over external arguments (NOE): yes, see parameters NOA or NGO.
47. Definiteness on APs (DOA): yes, definiteness of the whole NP is obligatorily expressed with an article introducing postnominal adjectives, cf. M. 196.6 *ī cub-i ī reng-drayāci* <’y cw b ’y rng-dry’cy> ‘discolored water’, lit. ‘the waters (pl. tantum), the discolored (ones)’ (Durkin-Meisterernst 2009, p. 358). Another example (QR 123) demonstrates that also a preposition

must be repeated *f-ī zβāk-a f-ī turkāng-a* <f-y zβ'k f-y trk'nk> 'in the Turkish tongue', lit. 'in the tongue, in the Turkish (one)'.

48. Verbless relative clauses (VRC): no, relative clauses apparently always have a predicate.
49. Facultative linker (FLI): no. The Chorasmian article does not function as a (facultative) linker.
50. Generalized linker (GLI): no.

### 3.4.4 Bactrian

Bactrian (Bactr.) was once spoken in the historical region of Bactria, situated between the Amu Darya and the Hindu Kush. Bactria came under strong cultural influence of the Hellenistic world as a direct result of Alexander's campaign, with the establishment of Greek colonies and accompanying usage of the Greek alphabet. In the middle of the 2nd century BCE, the Yuèzhī, invaders from the north-east (ultimately from Gansu), settled along the middle reaches of the Amu Darya and one of their clans, the Kushans, founded an empire which eventually extended from the north-western part of the Indian subcontinent into the Tarim Basin. The Kushans quickly adopted the Bactrian language of their new homelands as their official language, together with the Greek alphabet to write it; their native language may have been a Middle Iranian language as well.<sup>23</sup> The adoption of the Bactrian language and the Greek alphabet is amply demonstrated by the royal inscriptions of the Kushan era (Surkh Kotal, Rabatak etc.; see Sims-Williams 2004 or Sims-Williams and Falk 2014).

Until the 1990s, Bactrian was only known from these Kushan era inscriptions. That changed with the discovery of hundreds of other documents, mostly letters, but also legal documents, which were written on organic materials (Sims-Williams 2007, 2012). These documents can be dated to various points in the 1st millennium CE (the tradition ceases around 800 CE) and they are written in (cursive) Greek script, evincing the early and long-lasting impact of Greek on Bactrian culture.

In terms of phonological developments and characteristics, Bactrian clearly allies with the eastern Iranian languages Khotanese, Sogdian and Chorasmian. Morphologically, however, Bactrian appears to be more closely affiliated with Middle Persian and Parthian and their attritioned and reduced inventories than with the eastern, conservative systems of Khotanese and Sogdian. Syntactically,

---

<sup>23</sup>See Bonmann et al. (2023). The 'Kushan script' either recorded the native language of the territory between the Amu Darya and the Hisar mountain range, or it preserved the original language of the Yuèzhī.

Bactrian joins Sogdian and Chorasmian in having developed an article system, thus diverging both from the eastern *Randsprachen* Khotanese and Tumshuqese and its western relatives Middle Persian and Parthian.

1. Gramm. Person (FGP): yes, there is verbal person and there are also different sets of demonstrative pronouns (ελλο < \**aġam*, εμμο, μο < \**ima-*, ειδο < \**aġtam*, οο < \**aġam*). It is highly doubtful that Bactrian had no deictic contrasts in its demonstratives, as argued in Gholami (2009). Since Bactrian is a dead language, it is difficult to understand the exact deictic nuances as a modern reader, but I am sure that etymologically different sets of demonstratives imply *some* contrast in terms of semantics and/or deixis. Gholami herself states that (l.c., p. 24) “[t]he first three demonstratives display a proximate deixis. οο is the only demonstrative that can have remote deixis, but the usage of that demonstrative is limited.”
2. Gramm. Number (FGN): yes, there is a distinction between singular and plural. In the Kushan era, two cases were distinguished in singular and plural each (thus leading to a system with four different endings), whereas in the later texts there is only a difference between an unmarked form and a marked plural form (Sims-Williams 2007, p. 40). However, the singular–plural distinction as such is maintained.
3. Gramm. Gender (FGG): Kushan era: yes; later: no. The older documents (inscriptions dating to the Kushan era) still have a two-gender system (masculine and feminine), see Sims-Williams (2007, pp. 40–41). The later language (or rather the Bactrian of documents from the Indo-Sas(s)anian, Kidarite, Hephthalite eras etc.), however, lost this distinction. There are a few remnants of the old gender distinction, e.g., ιμμορρο ‘the meadow’ with a proclitic article marked for feminine gender (ιμ- instead of normal or unmarked ι-). However, as Sims-Williams notes (l.c., p. 41), the feminine article also occurs with an old neuter in ιρρωσο ‘the day’ (or already in

the Kushan-era inscription of Surkh Kotal  $\iota\alpha \nu\iota\beta\alpha\lambda\mu\omicron$  ‘the seat’).<sup>24</sup> The fact that centuries after the (probable) loss of the category of a feminine gender in Bactrian a previously feminine (or neuter) stem still surfaces with the old feminine article appears to be an instance of a nanoparametric setting. Grammaticalized information was stored in the functional head of some lexical items such as  $\mu\alpha\rho\gamma\omicron$  ‘meadow’, viz. that this particular noun (or other, similar nouns) had to occur with a special, distinct form of the article ( $\iota\alpha$ -). The morphosyntactic marking was lexicalized, nanoparametrically. In the later Bactrian language, gender was probably no macroparametric category anymore, but a lexical, individual, nanoparametric one. This implies for a macroparameter like FGG that the Kushan-era Bactrian had a positive value, whereas the later language had a negative value.

4. Feature spread to N (FSN): yes, cf. See O 3’ (abbreviations of documents following Sims-Williams 2007, 2012):  $\beta\alpha\gamma\text{-}\alpha\nu\omicron \beta\alpha\nu\omicron$  ‘king of the gods’ in which the dependent genitive is marked for (gen.) plural. Another example: Rabatak, line 17,  $\beta\alpha\gamma\text{-}\epsilon$  ‘(the) gods’ is marked for direct plural.
5. Number on N (bounded nouns) (FNN): yes, see FSN above.
6. Gramm. partial definiteness (DGP): yes, partial definiteness is expressed by means of the definite articles (there are two sets, based on  $\iota$ - and  $\mu$ -). Consider the following sentence:

---

<sup>24</sup>This situation is reminiscent of Khotanese, Chorasmian and Sogdian, where old neuters likewise sometimes get inflection markers (articles or endings) of masculine or feminine stems/articles. This indicates that neuter forms were not understood anymore as representing a third gender and that old neuter stems were incorporated into one of the two remaining gender systems of morphosyntactic marking, either feminine or masculine.



- (35) Bactrian, anaphoric definiteness marking (jf 6-9), see also Gholami (2014, p. 162)

οδ-ασο ι χαρο πορδανιγο οδο γονζο οδο ροτιγο μαρο  
and-from the donkey saddle(?) and bag and rope hither  
φορτιο

send.IPRF.2SG

‘And (apart) from this, send hither the donkey-saddle and a bag and a rope’.

7. Gramm. definiteness (DGR): no, there are many instances of NPs without article. Hence the article obviously was not obligatory in Bactrian. See e.g.

- (36) Bactrian, article-lacking NP (jh 1-2)

οαρσοχοαγδο λαδοαγαλγο υογαμγο βαγανο βαυ[ο]  
wonderful granter.of.favors renowned god.PL king

‘the wonderful, the granter of favors, the renowned king of the gods’.

Apparently, Bactrian had only partial definiteness marking; it was not generalized as in Sogdian or Chorasmian. A negative value of DGR neutralizes several parameters (DDA, DCN, DSN, DOR, CGR) according to Longobardi et al. (2013).

8. Strong Person (NSD): yes, cf. bg 11: χοαδ[ηο]βανδαγο βαυραβο ‘Xwadew-bandag the Satrap’, or Rabatak, l. 14: κανη|κε βαο ‘King Kaneška’.
9. Free null partitive Q (DPQ): no, the preconditions for a positive value (different cases) are not given in the later language, and Kushan era Bactrian (which still had a two case system) shows no signs of differential case marking.
10. Gramm. distal art. (DDA): (yes). Strictly speaking, this parameter is neutralized due to a lack of generalized definiteness marking (DGR –). However, under the assumption that this parameter is not dependent upon DGR

(as Longobardi et al. propose), but the higher ranking DGP (partial definiteness marking), one could ask several questions related to the definiteness markers of a language, including questions about the parameter DDA. This parameter asks whether or not definite articles encode obligatory distal/proximate distinctions. It is conspicuous that Bactrian has two different sets of articles, viz. one based on  $\iota$ - (< Proto-Iranian  $^{*}(H)ia$ -) and one based on  $\mu$ - (< PIr.  $^{*}ima$ -). Consider, for instance,  $\mu\alpha \lambda\iota\zeta o$  ‘the fortress’ from the inscription of Surkh Kotal (l. 4) in comparison with Rabatak, l. 20:  $\alpha\beta o \iota \alpha\rho\eta\mu\sigma\sigma o \chi\rho o\nu o$  ‘in the third year’ (Prep–D–Ord–N). What exactly was the meaning, the function, the deictic nuance? We do not know. But the mere fact that two sets of articles were in use tells us that there might have been *some* difference. Hence parameter DDA might be set positively, if another parametric hierarchy with a different precondition applied. But according to the concept of Longobardi et al., this is a neutralized parameter.

11. Def.-checking N (DCN): (no), this parameter asks whether the definiteness marker of a language is a bound morpheme cliticizing on the noun or not. Bactrian articles (definiteness markers) occur to the left of N and other elements can intervene between D and N (see e.g., the example of FFS below). Again, in the model of Longobardi et al. (2013), this parameter depends upon a positive value of DGR and should thus be neutralized or be regarded as irrelevant in Bactrian; even though, in my view, a dependence upon DGP might be more appropriate.
12. Def. spread to N (DSN): (no). The parameter defines whether or not a language has, besides free definite morphemes in D, additional definiteness suffixes on N. This is not the case in Bactrian. Again, this parameter is neutralized by a negative DGR.
13. Def. on relatives (DOR): (no). Bactrian relative pronouns signal animacy (animate  $\kappa\iota\delta o$ ,  $\alpha\kappa\iota\delta o$ ,  $\kappa\iota\delta$ -,  $\alpha\kappa\iota\delta$ -,  $\alpha\kappa\iota\delta\alpha$ - ‘who’ : inanimate  $\alpha\sigma\iota\delta o$ ,  $\sigma\iota\delta o$ ,  $\alpha\sigma\iota\delta\alpha$ -,  $\alpha\sigma\iota\delta$ -,  $\sigma\iota\delta$ - ‘what’), but not definiteness. Once again, this is actu-

ally a neutralized parameter due to DGR –.

14. D-controlled infl. on N (DIN): no. The inflectional form of Bactrian nouns does not depend on the occurrence of certain morphemes in D beyond simple phi-feature concord; there is no nunation.
15. Plural spread from Cardinals (CPS): yes, cf. ef 17: *λοι ποσ-αγο* ‘(the) two sheep’ (Card–N-PL). See also Gholami (2014, p. 88).
16. Grammaticalized boundedness (CGB): yes, singular count nouns have an unbounded reading, cf.

- (37) Bactrian, unbounded reading of singular-inflected nouns (cl 8)
- οδ-αλδο γαλο οδο ιασπο ναχωζηιο*  
 and-if thief and the.horse not.ask.OPT.2SG  
 ‘and if you do not search out the thieves and the horses’.

17. Strong article (CGR): (yes), but since CGB is set positively, this parameter is predictable or irrelevant. The numeral ‘one’ marks the indefinite reading of a singular count nominal, cf. e.g.

- (38) Bactrian, bounded reading of singular-inflected nouns by means of numeral ‘one’ (T 6), see also Gholami (2014, p. 164)
- οδο-μο λαδδηι μισο μασκονδο ιωγο ζινο*  
 and-I give.PST.OPT.3SG also person one woman  
 ‘and also I gave it to a certain person, a woman’.

Here it occurs to the left of N, but the reverse order is attested, e.g., in V 7: *αγγαραγο ιωγο* ‘a certain/one property’ (N–Indf).

18. Bounded-checking N (CCN): no, Bactrian does not have a clitic marker for bounded indefinite readings, but uses the word for ‘one’, *ιωγο*.

19. Null-N-licensing article (DNN): yes, the article can be used to introduce definite nounless argument nominals such as a genitive argument. Cf. cp 17 *ισωνοκανιγο* ‘the (representative) of the Sohukan (family)’ (with a ‘fused’ article).
20. Structured Adjective Phrases (AST): yes. Examples may be seen in jh 1: *οαρσοχοανδο λαδοαγαλγο νοναμγο βαγανο* *ḫaυ[o]* ‘the wonderful, the granter of favors, the renowned king of the gods’, cc 14 *κοδο ḫιζγο* ‘good dog’ (with N-Adj). Cf. also FFS below.
21. Feature spread to structured APs (FFS): yes, cf. J 3:
 

(39) Bactrian, feature spread to structured adjectives (J 3; see also Q 8 and A 7)

*πισο μανδαρου-αν-ισο αζαδοβοργ-ανο οιγαλφ-ανο*  
 to the.other-PL-also freeman-PL witness-PL  
 ‘also to the other freemen (and) witnesses’.

in which ‘other’ is embedded between article (D) and N (and gets a plural marker, thus evincing feature spread).

22. Feature spread to predicative APs (FSP): no. In later stages of Bactrian it is difficult to tell whether predicative adjectives show any signs of phi-feature spread, since the only overt signal could be the plural marker occurring on a predicative adjective. Even though the possibility would exist to mark predicative adjectives for the number of their head noun, this does not happen (e.g., xe 18–19 ‘[the people and animals are] all healthy’: Bactr. *οισπο λρογινοδο*, both adjectives in sg. form). In the earlier documents, dating from a time when Bactrian still had a two case system, the marking would have been even more clear, as not only number, but also case could ‘spread’ from N to predicative adjectives. Yet this does not happen and predicative adjectives occur in their basic form (singular; direct case); cf., for example, the following clause from the Rabatak inscription:

- (40) Bactrian, lack of feature spread to predicative adjectives (Rabātak, 17-18; following Sims-Williams 2004, p. 57)

*βαγ-ε (...) αβο βαογ-αγ[ο βαι] αβο κανηβκε*  
 god-DIR.PL (...) DO king-OBL.PL king.OBL.SG DO Kanēshka.OBL.SG  
*κοβανο αβο ιασηδανι ζορριγι λρουγο*  
 Kushan for/to eternity.OBL.SG period/age.OBL.SG healthy.DIR.SG  
*αγγαδδιγγο οανινδογ[ο]*  
 fortunate.DIR.SG victorious.DIR.SG  
 ‘May the gods (...) [keep] the [king] of kings, Kanēshka the Kushan, forever healthy, fortunate (and) victorious!’

23. D-controlled infl. on adjectives (ADI): no. Adjectives do not receive different marking depending on the presence or absence of a determiner.
24. DP over relatives (ADR): yes, relative clauses surface to the right of NPs, so that the NP/DP is ‘raised’. Consider the following example:

- (41) Bactrian, relative clause to the right of N (O 4’-5’)
- οδο πισο μανδαρου-αν-ισο αζαδοβοργανο ακιδο*  
 and with.regard.to the.other-PL-also freeman.PL REL  
*χοησαοοβωστιγο ταβδο*  
 contract.of.undertaking seal.PST.3SG  
 ‘and also with regard to the other freemen who have sealed (this) contract of undertaking’.

Since ADR has a positive value, the dependent parameter AER is neutralized.

25. Free reduced RelCl (ARR): yes (?). There are cases like Surkh Kotal, l. 4 *στο μα λιζο πιδοριγδο* ‘and the fortress became deserted’, in which the last word is a participle (or rather an old verbal adjective). The usual interpretation and translation would operate under the assumption that *πιδοριγδο* does not function as a nominal adjunct, giving additional information (a reduced relative clause), but that it instead functions as the predicate (with

omission of the copula). However, one could also interpret the clause as ‘and the fortress was a deserted one’. Another case in point could be C1:

- (42) Bactrian, free reduced relative clause/participle (C1)  
 εδο μο λαφνοβωστογο μαλο νιβιχτο αβο μο ρωβαγγο ραρο  
 this.here ART deed.of.gift here written at ART of.Rob city  
 αβιο ανδαγο  
 at.this borough  
 ‘This deed of gift was written here in the city of Rob, in the borough...’.

This is the translation of Sims-Williams (2012, p. 32). But would an interpretation as ‘This (is) a deed of gift, written here in the city of Rob, in the borough...’ be too far fetched?

26. Adpositional genitive (GAD): no, genitives are either (a) marked by a separate ending (-ανο, originally the gen. pl., but later largely decoupled from its genitival semantics), (b) the article functioning as a linker (N ι N), (c) morphologically unmarked (in which case only the position within the NP structure reveals the genitival function), or, most frequently, by means of (d) adjectives functioning as genitives (jc 4–5 ι λρουμιναγγο ραυρο ‘the city of the enemy’ with D–Adj<sub>gen</sub>–N). Adpositions are not used to indicate a genitival relationship.
27. Free Gen. (GFR): yes, an example with several recursive genitives is the following one:

- (43) Bactrian, recursive genitives (F 17’-F18’)  
 γαζνο βρηδαγ-ανο ζαροζιδαγο διγαρο πανζασο ν’  
 treasury bredag-GEN struck.gold dinar fifty 50  
 ‘the treasury of the *bredag* of fifty dinars of struck gold’.

Note that only the first genitive receives a genitive ending, whereas the others are not inflected.

28. Uniform Gen. (GUN): yes, cf. Surkh Kotal, line 1: *μο κανηρκο οανινδο βαγολαγγο* ‘the sanctuary of the victorious (one), of Kaniška’, in which *μο ... βαγολαγγο* is the NP core (D–N, ‘the sanctuary’), whereas *κανηρκο οανινδο* are two recursive genitival arguments (though without endings or adpositions) within the structured core of the NP. An order N–Gen is attested, for instance, in C 3: *χοηοι κοσιρδαχμιο* ‘the lord of Kusirdaxm’, Gen–N in dd4: *σαγο πορο* ‘son of Sag’, Gen–Gen–N in eh 4: *τοχοαροστανο οδο γαρσιγοστανο λαδοβαρο* ‘the judge of Tocharistan and Gharsigostan’. Apparently, genitives can surface in various positions inside the NP core or to its right; they are ‘uniform’. This uniform genitive can occur freely or in structured positions such as that of a postadjectival object genitive (example in jh 1–2, see parameter DGR above). A positive value of GUN neutralizes parameter GFO.
29. DP over free Gen. (GPR): yes, see the example of GFR.
30. Gen-feature spread to N (GFS): no. There is no number or gender spread from genitives to their head nouns.
31. D-checking possessives (PDC): (yes), possessives can be used as definite determiners (as in French or German where possessives can occur without any visible article, e.g. *mon livre*, *mein Buch* vs. *\*le mon livre*, *\*das mein Buch*). For an example, cf. C 15: *μανο βραδο* ‘my brother(s)’. However, since parameter DGR is set negatively, this parameter becomes irrelevant. According to Longobardi et al., this parameter depends on full grammaticalization of definiteness. I am not sure whether this assumption is correct, hence I will mark this in the grid with brackets: (+).
32. Adjectival poss. (APO): yes, Bactrian has adjectival possessives, cf. jf 3–4: *ι μαναγγο βαγο* ‘my share’ (D–Poss–N); U 27: *μαναγγο πιδοοαυανο* ‘my request’ (Poss–N) or Y 9: *μαναγγο ζαμμο* ‘my land’ (Poss–N).

33. Post-affix poss. (PAP): irrelevant. Since definite articles are not suffixed to N, this parameter becomes irrelevant.
34. Clitic poss. (PCL): yes, Bactrian also has possessives cliticizing on the first stressed item. This can be, for instance, an article, cf. dd 5: ι-μανο χοβο ‘my property’ (D-Poss-N) or a relative pronoun as in ba 4-5: κοαδο-μo ιαοαρδαοι ‘that my corn’ (RelP-Poss-N).
35. N-feature spread to pron. poss. (PHS): yes, number and person are marked on possessive pronouns, see (besides the examples of PCL) Q 11-12 οδο μαχαγγο βραδο ‘and our brothers’ (Poss-N).
36. N-feature spread to free Gen. (GSP): no. Nominal feature are not ‘spread’ to genitives (see the example of GFR). Adjectival genitives can be used in Bactrian, as can be seen in, e.g., C 3: χοηοι κοσιρδαχμιο ‘the lord of Kusirdaxm’ (-ιγο is an adjective ending inflected for sg.). But this does not permit a positive value for this parameter.
37. Adjectival Gen. (AGE): yes, cf. jc 3-4 ι λρουμιναγγο ιαυρο ‘the city of the enemy’ with D-Adj<sub>Gen</sub>-N (‘Adj<sub>Gen</sub>’ here symbolizes an adjectival genitive).
38. Poss.-checking N (GCN): no, Bactrian shows nothing like the Semitic construct state.
39. Strong partial locality (TPL): (yes). The parameter asks if at least some demonstratives can occur in the D area. This is clearly the case. It is noteworthy that demonstratives even can precede determiners, visible e.g., in Surkh Kotal, line 1: ειδο μα λιζο ‘this fortress’ with Dem-D-N. However, it is a precondition for the setting of this parameter that CGR, ranking higher in the hierarchy, must not be positive (and it is positive in Bactrian). Note that determiners have the possibility (but only μο/μι?) to surface rather low in the Bactrian NP, e.g., Rabatak, l. 2: ασο οισποανο μι βαγανο ‘from all the gods’ (Prep-Adj-D-N).



40. Strong locality (TSL): yes, all demonstratives appear to occur in the D area (Gholami 2009).
41. D-checking demonstratives (TDC): no, demonstratives co-occur with definiteness-compatible articles in the D-area, hence this parameter gets a negative setting. See the first example of TPL above.
42. N over Demonstratives (NDE): no, this seems to be impossible. Demonstratives always precede the noun. See Gholami (2009).
43. N over Cardinals (NOC): yes, apparently. Cf. C 20': διναρο κ' οιστο 'twenty dinars', with N-Card (the number is both written out and abbreviated). Examples of this kind are abundant in the letters, usually with 'dinars' as head.
44. N over Ordinals (NOO): unclear, as Gholami (2014, p. 89) states, "[e]xamples of ordinal numbers are very rare in Bactrian materials", and even though examples with Ord-N appear to be attested (e.g. Rabatak, l. 20), the texts in Sims-Williams (2007, 2012) and the inscriptions did not reveal any examples of the reversed order, N-Ord. The parametric hierarchy predicts that nouns could be raised over ordinals, since they could apparently also be raised over cardinals. Probably the text corpus is too small.
45. N over adjectives (NOA): yes, cf. cc 13-14 κοδο βιζγο 'good dog' or J 10 αγγαργο παροιαβιγο 'disposable property' with N-Adj order each. Similar to Chorasmian, there may be a definiteness-indefiniteness difference depending on the 'raising' of N (with N-Adj signaling indefiniteness of an NP), see Gholami (2014, pp. 92-95). The observation that nouns can be raised over adjectives is beyond doubt.
46. N over GenO (NGO): yes, cf. C 3: χοηοι κοσιρδαχμιγο 'lord of Kusirdaxm' or F 1-2: μολρογο αβησαχοανιγο 'sealed document of renunciation' (N-Gen)

- 47. N over external arguments (NOE): yes, see NGO, NOA, and NOC.
- 48. Definiteness on APs (DOA): no. Definite NPs consistently evince an order D-Adj-N so that adjectives, when occurring postnominally, usually belong to an NP with indefinite reading. Hence this parameter has a negative value (or is irrelevant).
- 49. Verbless relative clauses (VRC): irrelevant due to positive value of FLI.
- 50. Facultative linker (FLI): yes, see Gholami (2011). Examples: xd 6:  $\nu\alpha\tau\omicron \iota \chi\alpha\rho\alpha\gamma\alpha\nu\omicron$  ‘Nat (the) Xaragan’; G 2–3:  $\mu\omicron\zeta\delta\omicron \iota \chi\alpha\rho\alpha\gamma\alpha\nu\omicron \iota \kappa\alpha\nu\delta\omicron\gamma\omicron\lambda\eta\rho\omicron$  ‘Muzd Xaragan, the keeper of the granary’ or L 2:  $\beta\alpha\gamma\omicron \iota \omicron\alpha\chi\beta\omicron$  ‘the god Waxš’ (Gholami 2014, pp. 82–83).
- 51. Generalized linker (GLI): no. The linker construction is facultative and not obligatory or generalized.

### 3.4.5 Parthian

Parthian (Parth.), the official language of the Arsacid dynasty (247 BCE–224 CE), was spoken in present-day northeastern Iran and neighboring parts of Turkmenistan (to the south-east of the Caspian Sea, in other words). The Arsacids themselves were originally part of a federation of tribes living to the north of this territory, between Caspian and Aral Sea, but they quickly adopted the language of the territory they conquered first, viz. Parthian, before gaining power in the rest of Iran. There are some early inscriptions from about 140 BCE onward (and ostraka etc.), but the greatest part of Parthian textual remains dates to the 8th and 9th centuries and is preserved in documents written, preserved (and found) in the Tarim Basin. It is difficult to determine how long Parthian was a living language (see the discussion in Durkin-Meisterernst 2014, 3ff.); it was probably actively spoken from the 2nd century BCE until the 7th century CE, and the later documents only demonstrate that it was in use as a dead, liturgical language of Manichaean communities.

Parthian, like Middle Persian, has a drastically reduced nominal morphology, compared with the Old Iranian languages or Khotanese Saka. The syntactic profile of Parthian, at least concerning noun phrases, can be defined as follows:

1. Gramm. Person (FGP): set to yes. There is a three-way contrast in terms of verbal person and personal pronouns, and a two-way contrast in terms of deictic pronouns, see Skjærvø (2009a, p. 209) and Durkin-Meisterernst (2014, p. 211).
2. Gramm. Number (FGN): yes. There is a distinction between singular and plural. Note, however, that plural marking of nouns is not obligatory (Durkin-Meisterernst 2014, pp. 197–198) in later stages of the language.
3. Gramm. Gender (FGG): no. Parthian has suffixes which can be used to signal feminine gender (e.g., *-čān* in *niyōšāg-čān* ‘female hearer’), but this is more a derivational process than fully grammaticalized gender in the

sense of this parameter. There is no systematic gender marking on nouns and/or their arguments.

4. Feature spread to N (FSN): yes. Number is spread from the D-position to nouns, cf. e.g.,

- (44) Parthian, number marking on nouns, BT 11, 791ff.  
*was-ān*        *dušmen-in*  
 many-OBL.PL enemy-OBL.PL  
 ‘many enemies’.

5. Number on N (bounded nouns) (FNN): yes, cf. the preceding example of (FSN).

6. Gramm. partial definiteness (DGP): set to no. Parthian has no obligatory anaphoric or definiteness marking on nominal arguments. Consider e.g., *harw tawān kām kirbag ispuṛr būd* ‘all your desired good deeds were completed’ (Mir. Man. iii, text m 20–22, following Skjærvø 2009a, p. 223). Perhaps more clearly:

- (45) Parthian, lack of anaphoric definiteness marking, (BT 11, 791–801, following Skjærvø 2009a, p. 238)  
*kē aṣ maḍyān wasān dušmenīn wirēxt ud ... ō dašt*  
 who from middle many.OBL.PL enemy.OBL fled and ... to plain  
*ēw wazurg ud wiyābān yādēndē aḍyān aṣ dūr*  
 a/one large and deserted come.OPT.3SG then from far  
*maḍyān dašt kadag ēw wēnēndē*  
 middle plain house a/one see.OPT.3SG  
 ‘(like a man) who (had) escaped from the middle of many enemies, and were to come **to a large plain** and desert... and then from afar were to see a house in the middle of **the plain**’.

The ‘plain’ in this example must be understood as being definite, yet the noun is not marked for definiteness. Hence, DGP must be set negatively. As before, a negative parametric value of DGP neutralizes several other parameters.

7. Free null partitive Q (DPQ): no. Parthian has no count/mass distinctions by means of differential case marking; in fact, the later language does not even have cases anymore (however, earlier documents, such as royal inscriptions, have a case distinction between direct and oblique, or even between nominative, genitive and ‘prepositional’; see Durkin-Meisterernst 2014, pp. 197–198).
8. D-controlled infl. on N (DIN): no. Parthian has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. the following example:

- (46) Parthian, plural spread from cardinals to nouns, KPT 942-3  
*pañ-enān rōšn-ān*  
 five-PL light-PL  
 ‘five lights’.

10. Grammaticalized boundedness (CGB): yes. Nouns are inherently unbound and require special affixes or markers to get a bounded reading, either plural markers to openly signify plurality (or collectivity) or the numeral *ēw* ‘one’, used enclitically (see example 45). See Durkin-Meisterernst (2014, p. 360):

- (47) Parthian boundedness on nouns
- a. Unbounded reading of nouns not openly marked by suffixes (Durkin-Meisterernst 2014, p. 419)  
*gyāgrōb až man karēnd, kē wirāzēnd mēhan ud mān*  
 broom from me make.3PL REL tidy.up.3PL home and house  
 ‘They make brooms out of me which tidy up home and house.’
  - b. Bounded reading of nouns by means of clitic suffix ‘one’, KNRm 35  
*mard ēw āhāz*  
 man INDF was  
 ‘there was a man’.

11. Bounded-checking N (CCN): yes, the marker is the word for ‘one’, *ēw*.
12. Structured Adjective Phrases (AST): yes. Cf. BT 4, lines 2048–2049 (following Skjærvø 2009a, p. 235) *pad hamag tanbār* ‘in the entire body’. Another example: *ō pawāg-ān ud rāšt-ān dēnāβar-ān* ‘the pure and righteous Dēnāβars’ (object marker/preposition–Adj–Adj–N).<sup>25</sup>
13. Feature spread to structured APs (FFS): yes, see Durkin-Meisterernst (2014, p. 203). The inflection (plural ending) can be seen in the second example of AST, given directly above.
14. Feature spread to predicative APs (FSP): yes, to give an example from the inscription of Šābuhr I at Haijabad, *āgas(ī) būd ahēndē* ‘it would have been **visible** (<’ksy>)’ (Skjærvø 2009a, p. 267), marked for singular.
15. DP over relatives (ADR): yes. Relative clauses can follow their head nouns. See Durkin-Meisterernst (2014, 421ff.): *ud \*ōyādag, \*rabēsēf, senātōr ud hēgemōn, kē...* ‘and the others, the Praetorian prefects, the senators and the hegemons, who...’. This neutralizes parameter AER.
16. Free reduced RelCl (ARR): yes, see Durkin-Meisterernst (2014, p. 390).
17. Adpositional genitive (GAD): no. Genitival relations are either expressed (1.) by means of the (proto-)linker construction (see below) or (2.) by mere juxtaposition, or (3.) by means of a genuine genitive case (older texts). A preposition can be added, but there is no adpositional genitive in the sense of this parameter, see *kē zād aβ bay-ān* ‘who (I am) born from gods’. (Durkin-Meisterernst 2014, p. 403).
18. Free Gen. (GFR): yes, cf. the recursive genitive construction (not morphologically marked) in the Parthian inscription of Šābuhr I at Hajiabad, as per

<sup>25</sup>This example shows that the original ending of the genitive plural had lost its case interpretation and was only used as a plural marker here.

Skjærvø (2009a, p. 266): *puhr mazdēzn bay ardaxšahr* ‘son of the Mazdean god/divine Ardaxšahr’.

19. Uniform Gen. (GUN): no, there is no uniform inflectional and recursive genitive in the sense of this parameter. Compare *pad yazdān čašm* ‘in the eyes of the gods’ (Prep–Gen–N, Durkin-Meisterernst 2014, p. 404) with *zād aš bay-ān* ‘born from gods’ (N–[Prep–]Gen) and the example of GFR above. Note also *abēsūd draxtān* ‘most useless of the trees’ (Durkin-Meisterernst 2014, p. 405).
20. DP over free Gen. (GPR): yes, free, recursive genitives surface to the right of the noun, cf. GFR above.
21. GenO (GFO): yes, see *awāyōn bāmistūn* ‘the column of fame’ (Durkin-Meisterernst 2014, p. 420).
22. Gen-feature spread to N (GFS): no. There is no spread of features from genitives to nouns.
23. Adjectival poss. (APO): no. Parthian expresses possession via clitic possessive pronouns (*gyān-um žām ō wahišt anōšag* ‘guide **my** soul to the immortal paradise’), often together with the copula (*u-š ast puhrān wīst hazār* ‘**he** has 20.000 sons’) (Durkin-Meisterernst 2014, pp. 291–292, 370–371). There are no adjectival possessives, and hence the dependent parameter AGE is neutralized.
24. Clitic poss. (PCL): yes, see Durkin-Meisterernst (2014, p. 273) and APO above.
25. N-feature spread to free Gen. (GSP): no. There is no feature spread of nouns to adjectives.
26. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs

on nouns when occurring with a genitive argument. This is not the case in Parthian.

27. Strong partial locality (TPL): no, demonstratives do not systematically surface in the D-area, introducing NPs. Cf. e.g., *padkar im* ‘this image’ with N–Dem order (Durkin-Meisterernst 2014, p. 279). This parametric value neutralizes TSL.
28. N over Demonstratives (NDE): yes, nouns can be raised over demonstratives. Cf. *padkar im* ‘this image’ with N–Dem order (Durkin-Meisterernst 2014, p. 279).
29. N over Cardinals (NOC): yes, visible, for instance, in the dating formula:

- (48) Parthian, noun raising over numerals, Mir. Man. iii, text d 57-60 (Skjærvø 2009a, p. 260)  
*pad saxt čuhram māh šahrīwar pad dōšambat ud žamān*  
 on passed fourth month Šahrīwar on Monday and hour  
*ēwandas*  
 eleven  
 ‘on the fourth of the month of Šahrīwar, on Monday, at 11 o’clock’.

The ‘hour’, Parth. *žamān*, surfaces before its dependent cardinal, hence N–Card is obviously possible or grammatical in Parthian. See also Durkin-Meisterernst (2014, p. 272).

30. N over Ordinals (NOO): yes, see the example of parameter NOC, directly above (with *saxt* as nominalized and raised element), cf. also Durkin-Meisterernst (2014, p. 272).
31. N over adjectives (NOA): yes, cf. Mir. Man. iii, text k, M104 R line 1ff. (after Skjærvø 2009a, p. 272): *warm ādurēn* ‘fiery cloud’ or *šahrδar aryāw* ‘noble prince’ (both N–Adj). See also Durkin-Meisterernst (2014, p. 264).
32. N over GenO (NGO): yes, cf. *dīdēm šahrδārīft* besides *šahrδārīft dīdēm* ‘diadem of rule’ (Durkin-Meisterernst 2014, p. 265).



33. N over external arguments (NOE): yes, examples under parameters NOO, NGS, NOA and NGO.
34. Verbless relative clauses (VRC): irrelevant due to positive value of FLI.
35. Facultative linker (FLI): yes. Parthian has the option to use an invariable linker to connect dependent arguments (nouns and adjectives) with a head noun in a rightward-branching manner (*čē*, see the discussion in Durkin-Meisterernst 2014, pp. 266–268). The usage of the linker is, however, not obligatory (Boyce 1964, p. 32: “its absence is commoner than its use”). Since the language nevertheless has this option, at least parameter FLI must get a positive value. Examples: *tōhm čē amā* ‘our family’, *pad šang ud srōd čē šādīft* ‘with harp and song of joy’, *šahrestān čē bēlābād* ‘the capital city Belabad’.
36. Generalized linker (GLI): no.

### 3.4.6 Middle Persian

Middle Persian (MiPers.) was the official language of the Sas(s)anian dynasty (224–652 CE), but is, like Parthian, mainly preserved in documents written centuries later. Native to Fars province, but widely used in the Sas(s)anian Empire, Middle Persian is not a direct descendant of Old Persian, but rather a close relative. Literary remains, mostly from the ninth and tenth centuries, and again preserved in the Tarim Basin, can be divided into two main traditions, namely a Manichaean Middle Persian one and a Zoroastrian Middle Persian (or ‘Book Pahlavi’) one. Besides that, there are also Christian texts (Nestorian sect), secular texts and earlier royal inscriptions, but the majority of the material is religious in nature and represents the Zoroastrian or Manichaean tradition. Like Parthian, Middle Persian has a drastically reduced nominal morphology, and its parametric values parallel those of Parthian.

1. Gramm. Person (FGP): set to yes. There is a three-way contrast in terms of verbal person and personal pronouns, and a two-way contrast in terms of deictic pronouns, see Skjærvø (2009a, p. 209) and Durkin-Meisterernst (2014, p. 211).
2. Gramm. Number (FGN): yes. There is a distinction between singular and plural. However, as in Parthian, plural marking of nouns is not obligatory (Durkin-Meisterernst 2014, pp. 197–198).
3. Gramm. Gender (FGG): no. As in Parthian, feminine gender can be marked by derivational affixes (e.g., *-ag* in names, as in *Ĵam* and *Ĵamag*, or, to give another example, *-čān* in *niyōšāg-čān* ‘female hearer’, being identical to the Parthian example).
4. Feature spread to N (FSN): yes. For instance, PT, p. 124 §23 (after Skjærvø 2009a, p. 236) *wišūdāg-ān* ‘abortions’, or *mardōhmān* ‘humans’ (Durkin-Meisterernst 2014, p. 314). Apparently number is not marked on the determiner element, but ‘spread’ to the noun.

5. Number on N (bounded nouns) (FNN): yes, cf. the preceding examples of (FSN).
6. Gramm. partial definiteness (DGP): set to no. There is no obligatory anaphoric or definiteness marking on nominal arguments. See, for instance, the Abnun inscription: *ka hrōmāy ... kū hrōmāy ... kū hrōmāy* ‘(Then, in year three of Šābuhr,) when the Romans (were coming, then I was here. Then, when I heard) that the Romans (were coming, then I implored the gods... Then, when I heard that Šābuhr had smashed) the Romans’ (Skjærvø 2009a, p. 251). As before, this neutralizes several other parameters.
7. Free null partitive Q (DPQ): no. Middle Persian has no count/mass distinctions by means of differential case marking. The earliest texts still have a two-case system, the latter ones have lost it, so that the precondition for this parameter is no longer given.
8. D-controlled infl. on N (DIN): no. Middle Persian has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. the following example:

- (49) Middle Persian, plural spread from cardinals to nouns, M206 II V 4-5, GW § 144 (Skjærvø 2009a, p. 212)
- |   |                 |           |                 |                |
|---|-----------------|-----------|-----------------|----------------|
| <i>pad</i>                                  | <i>hazār-ān</i> | <i>ud</i> | <i>bēwar-ān</i> | <i>wyāg-ān</i> |
| with/in                                     | thousand-PL     | and       | ten.thousand-PL | place-PL       |
| ‘in thousands and ten thousands of places’. |                 |           |                 |                |

Even though the parameter thus has to be set positively given the language’s possibility of ‘spreading’ the feature of plurality/number from cardinals to nouns, this does not mean that MiPers. nouns must always be marked for number when occurring with numerals. Cf. M 98 R, line 20 (after Skjærvø 2009a, p. 223): *čahārdah dar* ‘fourteen doors’ with ‘door’ being unmarked for number (but see immediately below, parameter CGB).

Note again that parametric values are often no either/or decisions and just ask for the neutral possibility per se to allow for a certain syntactical behavior.

10. Grammaticalized boundedness (CGB): yes. Nouns are inherently unbounded and require special affixes or markers to get a bounded reading, either plural markers to openly signify plurality (or collectivity) or the numeral *ēw* ‘one’, used enclitically (in some manuscripts also written as heterogram <-HD> or number <-1>). See Skjærvø (2009a, p. 205):

(50) Middle Persian boundedness on nouns

- a. Unbounded reading of nouns not openly marked by suffixes, Bd 21E.1

*ast ka māhī ayāb wazay ayāb sang wārēd*  
 there.is when fish or frog or stone rains  
 ‘There are times when fish, frogs, or stones rain.’

- b. Bounded reading of nouns by means of clitic suffix ‘one’, KNRm 35

*ud nūn zan-ēw paydā*  
 and now woman-INDF manifest  
 ‘and now a woman (became) manifest (= appeared)’.

11. Bounded-checking N (CCN): yes, the clitic is the word for ‘one’, *ēw*.
12. Structured Adjective Phrases (AST): yes. Cf. Durkin-Meisterernst (2014, p. 361) *zamīg-ēw wuzurg ud istabr* ‘a great and solid earth’ (N-[Idf]-Adj-Adj). Dem-N-Adj(-Gen) is attested in, e.g., *imīn gōwišnān rōšnān ī yamagānī rōzān* ‘these bright homilies of the Yamag-days’ (Durkin-Meisterernst 2014, p. 316).
13. Feature spread to structured APs (FFS): yes, see Durkin-Meisterernst (2014, p. 203) or the second example of AST with the adjective, *rōšn-ān* ‘bright’, being marked for plural.

14. Feature spread to predicative APs (FSP): yes, see *paydāg būd hē* ‘would have been visible’ in the inscription of Šābuhr I at Hajiabad (Skjærvø 2009a, p. 267), marked for singular.
15. DP over relatives (ADR): yes. Relative clauses can follow their head nouns (in fact, this is the normal order). See Durkin-Meisterernst (2014, p. 425). One example: *būd winōhag ō muryān kē...* ‘Confusion arose for the birds who...’. Since NPs can precede relative clauses, parameter AER is neutralized.
16. Free reduced RelCl (ARR): yes, see Durkin-Meisterernst (2014, pp. 390–391). An example: *zan ī ābustan... pad dō mard barišn* ‘(pertaining to) a pregnant woman... to be borne by two men’.
17. Adpositional genitive (GAD): no. Genitives are usually built with the (proto-)linker, alternatively the old ending of the genitive plural sometimes functions as a genitive marker (especially in older texts). Examples: *āwāg ī awēšān mazan guhūdagān* ‘voices of these Mazan-monstrosities’ (Durkin-Meisterernst 2014, p. 418) or *šāh-ān šāh* ‘king of kings’ (e.g., in the inscription of Šābuhr I at Hajiabad). In any case, Middle Persian has no adpositional genitives in the sense of this parameter.
18. Free Gen. (GFR): yes, there is a free, recursive genitive, albeit formed with the linker *ī*, cf. Skjærvø (2009a, p. 222):

(51) Middle Persian, free recursive genitive  
*kārnāmag ī ardašīr ī pābag-ān*  
 book.of.deeds LINK Ardašīr LINK Pābag-son.of  
 ‘the book of deeds of Ardašīr son of Pābag’.

19. Uniform Gen. (GUN): no. Middle Persian does not have a uniform genitive in the sense of this parameter (being inflectional, recursive and not bound to certain structural positions). There are inflectional genitives, but

they are bound to certain structured positions within the DP/NP core, cf. e.g., the famous phrases (irrespective of being archaic or not) with a preposed object genitive *šāh-ān šāh* ‘king of kings’ and *ēr-ān šahr* ‘realm/land/kingdom of the Iranians’ (e.g., in NP<sub>i</sub> 9). Another example is *pad yazd-ān nām* ‘in the gods’ name’ (following Skjærvø 2009a, p. 222), showing the embedding in the midst of the NP (the boundaries of which are delimited by preposition and noun). The normal case is a N–Gen order with (proto-)linker which is attested, for instance, in *ruwān ī mard-ēw* ‘the soul(s) of a man’ and *ēn ruwān ī awēšān druwandān* ‘these (are) the souls of those evil (ones)...’ (both adapted from Durkin-Meisterernst 2014, p. 314). This pertains particularly to free, i.e., recursive genitives (see the example of GFR).

20. DP over free Gen. (GPR): yes, the DP/NP can be raised over the genitive, as in the example of parameter GFR above.
21. GenO (GFO): yes, usually formed with connector: *kēn ī pidar xwāhēnd* ‘they demand vengeance for their father’ (Durkin-Meisterernst 2014, p. 270).
22. Gen-feature spread to N (GFS): no. Genitives do not exert influence on their head nouns in terms of number or gender.
23. Adjectival poss. (APO): no. Middle Persian expresses possession via clitic possessive pronouns (*u-š framān padīrānd* ‘and they will accept **his** orders’), often together with the copula (*ēn zan, kē-š yak pus ast* ‘this woman, who has one son’) (Durkin-Meisterernst 2014, pp. 291–292, 370). There are no adjectival possessives. This negative setting of APO neutralizes parameter AGE.
24. Clitic poss. (PCL): yes, see Durkin-Meisterernst (2014, p. 273) and parameter APO above.
25. N-feature spread to free Gen. (GSP): no. Nouns do not ‘spread’ number or gender to dependent genitives.

26. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Middle Persian. The construction with a connector/linker is different from the construct state of Semitic languages.
27. Strong partial locality (TPL): no, demonstratives do not systematically surface in the D-area, introducing NPs. Cf. e.g., ANRm a (following Skjærvø 2009a, p. 223) *pahikar ēn man* ‘this image (is) of me’ with N–Dem order. This neutralizes parameter TSL.
28. N over Demonstratives (NDE): yes, nouns can be raised over demonstratives, cf. ANRm a (following Skjærvø 2009a, p. 223) *pahikar ēn man* ‘this image (is) of me’ with N–Dem order.
29. N over Cardinals (NOC): yes, see Durkin-Meisterernst (2014, p. 272) and Brunner (1977, p. 47), and cf. the following illustrating example:
 

(52) Middle Persian, N–Card order, M98 R, lines 20–22 (Skjærvø 2009a, p. 223)

*ud čahārdah dar ud mān panz ud gāh sē*

and fourteen door and house five and throne three

‘and fourteen doors and five houses and three thrones’.
30. N over Ordinals (NOO): yes, see Durkin-Meisterernst (2014, p. 272) and Brunner (1977, p. 47). One example: \**mihr yazad ī dudīg* ‘god Mihr, the second’ (Durkin-Meisterernst 2014, p. 433). An example for the more frequent Ord–N order from the Bundahišn is Bd 33.0–4: *nazdist hazārag ... didīgar hazārag ... sidīgar hazārag ... hazārag* ‘first millennium ... second millennium ... third millennium’.
31. N over adjectives (NOA): yes, cf. e.g., *ō ēn xwān yōjdahr* ‘this pure table’ (after Skjærvø 2009a, p. 233), with a marker of the direct object, a demonstrative pronoun, noun and adjective (in this order). See also Durkin-Meisterernst (2014, pp. 264–265), and particularly footnote 7 on p. 264

(“[Boyce] macht darauf aufmerksam (...), daß das Adjektiv im Singular stehen muß [in an N-Adj order, SB].”).

32. N over GenO (NGO): yes, cf. Skjærvø (2009a, p. 222) *kārnāmag ī ardašīr ī pābag-ān* ‘the book of deeds of Ardašīr son of Pābag.
33. N over external arguments (NOE): yes, examples under parameters NOO, NGS, NOA and NGO.
34. Verbless relative clauses (VRC): irrelevant due to positive value of FLI.
35. Facultative linker (FLI): set to yes. Middle Persian has the option to use an invariable linker to connect dependent arguments (nouns, adjectives and clauses) with a head noun in a rightward-branching manner (*ī*, see the discussion in Durkin-Meisterernst 2014, 268ff. or Boyce 1964, pp. 37–47, Haider and Zwanziger 1984 or Yakubovich 2020). The usage of a linker is not obligatory (but very frequent), and since the language has this option, parameter FLI must get a positive value. Note the interesting fact that the alternative possibility to juxtapose a noun and a dependent element could even be combined with or rather integrated into a higher ranking linker construction, as is visible, for instance, in *wāxš ī hwarāsān wimand* ‘spirit of the Khorasan frontier’ with N-EZ-(Gen-N) (Boyce 1975, p. 40).
36. Generalized linker (GLI): no, the linker was not generalized (though very frequent).



### 3.5 Selected ancient Indo-European languages

I will now present the parametric values of three ancient Indo-European languages, namely Vedic, Latin and Ancient Greek. The data presented here shall serve as a first basis of comparative attempts beyond the Iranian level, though a database consisting of several more old Indo-European languages (Hittite, Tocharian A or B, Classical Armenian, Gothic, Old Church Slavonic, Lithuanian, Old Irish etc.) is absolutely essential to plausibly reconstruct the proto-language's parametric features. For the purposes of this pioneering study, however, a comparative section including at least one representative of the Italic and Greek branches as well as Vedic representing the Indo-Aryan sister branch of Iranian may be sufficient, because these three (sub-)branches might be the most indispensable ones for (provisional) reconstructions aiming at the Proto-Indo-European level. Data are easily accessible and the languages are well studied.

#### 3.5.1 Vedic

The *Ṛgveda*, the oldest Vedic text, was probably composed between ca. 1500–1200 BCE in the Punjab, whereas the three other *Samhitās* are dated to ca. 1000–800 BCE (see Jamison 1991, pp. 1–16).<sup>26</sup>

Consider the following sentence from the *Ṛgvedic* language. This sentence has several interconnected noun phrases, but the word order appears to be rather free:

(53) Vedic (RV 2.4.1)

a.	<i>huvé</i>	<i>vah</i>	<i>sudyótman-am</i>
	call.IND.PRS.MED.1SG	you.ACC/DAT/GEN.PL	refulgent-ACC.SG.M
	<i>suvṛkt-ím</i>		
	well_done-ACC.SG.M		

<sup>26</sup>Not all scholars would agree upon this dating – Oberlies (2012, p. 38), e.g., dates the oldest *Ṛgvedic* hymns to 1800 BCE –, but the arguments for a setting around the transition from the Bronze Age to the Iron Age make good sense to me.

- b. *viś-ām*                      *agn-ím*                      *átith-im*  
 clan/family-GEN.PL.F agní-ACC.SG.M guest-ACC.SG.GM  
*suprayás-am*  
 rich\_in\_oblations-ACC.SG.M
- c. *mitr-áh*                      *iva yáh*                      *didhiṣṣāyyah*  
 friend-NOM.SG.M like REL.NOM.SG.M win.DES.NOM.SG.M  
*bhū́t*  
 become/be.INJ.AOR.ACT.3SG
- d. *dev-áh*                      *ādev-e*                      *ján-e*                      *jātáved-āh*  
 god-NOM.SG.M godly-LOC.SG.M people-LOC.SG.M jātávedas-.NOM.SG.M

“For you I call the glorious refulgent Agni, the guest of men, rich in oblations whom all must strive to win even as a lover, God among godly people, Jatavedas.”<sup>27</sup>

The language of the *Ṛgveda* has strikingly parallel parametric settings to Old Avestan:

1. Gramm. Person (FGP): set to yes. Cf. the distinction between 1st, 2nd and 3rd person pronouns (*sá-/tá-*) and the contrast in terms of deictic demonstratives between proximate *eṣá-/etá-* or *ayám/iyám/idám* and distal *amú-(asáu/adás)* as well as *syá-/tyá-* ‘that’.
2. Gramm. Number (FGN): yes. There is a difference between singular, dual and plural. Consider e.g., the NPs marked for singular and plural in RV 1.1.2c: *sá deván ā ihá vakṣati* ‘he (*sá*) may bring hither the Gods (*deván*)’.
3. Gramm. Gender (FGG): yes. Vedic distinguishes between masculine, feminine and neuter gender. Consider RV 6.31.1:

<sup>27</sup>Translation following Griffith (1890). Geldner (1951) has: “Ich rufe den schönleuchtenden Agni, euren Preis, den gutbeköstigten Gast der Clane, der wie ein Freund gewonnen werden muß, der Gott unter dem götterfreundlichen Volke, der Jatavedas.”

- (54) Vedic, gender distinctions in nouns/NPs (RV 6.31.1)

*ví toké apsú tánaye ca*  
 against progeny.LOC.SG.N water.LOC.PL.F continuous.LOC.SG.N and  
*sūre ávocanta carṣaṇáyah*  
 sun.LOC.SG.M invoke.IND.MED.IPRF.3PL human.NOM.PL.F  
*vívācaḥ*  
 contender.NOM.PL.F

‘The humans (lit. the nomads/wandering ones), contenders (against each other with their voices), invoked (thee) for lasting progeny, for water(s) and for the sun.’

4. Feature spread to N (FSN): yes. Cf. RV 1.3.9 *viśve devāsaḥ (...)* ‘all gods’ with both the adjective and N being marked for plural.
5. Number on N (bounded nouns) (FNN): yes, cf. the preceding example of (FSN).
6. Gramm. partial definiteness (DGP): set to no. There is no obligatory anaphoric marking of nouns mentioned previously in the discourse. Vedic does not have either specialized articles or definiteness affixes on nouns (or nominal arguments); definiteness is in fact no category relevant or in any way marked. Consider e.g., RV 6.27.4b *yéna ávadhīḥ varásikhasya śéṣaḥ ...* ‘(Indra-power) with which you killed Varasikha’s offspring...’ and RV 6.27.5a *vádhīt índraḥ varásikhasya śéṣaḥ* ‘Indra killed Varasikha’s offspring’. The NP *varásikhasya śéṣaḥ* ‘Varasikha’s offspring’ is not marked for definiteness by means of an anaphoric pronoun or otherwise, despite a clear definite reading in the second stanza. Another example may be seen in the frequent epithets of deities which have a definite meaning, such as 6.49.7: *pāvīravī kanyā citráyuḥ sárasvatī vīrápatnī* ‘Sarasvati, the lightning’s child, the virgin of excellent vitality, the hero’s consort...’. Once again, negative setting of DGP neutralizes 14 parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA).

7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking. There is a partitive genitive, but this does not permit a positive setting of this parameter. If entities are countable, they will not receive any special case marking and will behave according to their normal function in the higher clausal setting, e.g. as a direct object (acc.) etc. Cf. 6.47.2 *ví navatím náva ca dehyàḥ hán* '(Indra) destroyed the ninety and nine ramparts (*dehyàḥ* acc.pl) (of Śambara)'.
8. D-controlled infl. on N (DIN): no. Vedic has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. RV 10.161.4 *śatám (...) śarā-daḥ (...) śatám hemantān (...) śatám (...) vasantān* '(Live, waxing in thy strength,) a hundred autumns, a hundred winters, a hundred springs' with 'hundred' + nouns inflected for ACC.PL. See also Debrunner and Wackernagel (1975, p. 338).
10. Grammaticalized boundedness (CGB): no. Singular count nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. Consider, for instance, RV 4.32.16: *vad-hūyúḥ iva yóṣaṇām* '(rejoice) like a lover (lit. one demanding a bride) in a young woman (*yóṣaṇām*)'. Or compare RV 8.67.14 *té naḥ āsnāḥ vṛkāṇām ādityāsaḥ mumócata* 'O ye Adityas, rescue us from the mouth of wolves!' with RV 2.29.6c *trādhvam naḥ devāḥ nijúraḥ vṛkasya* 'protect us, o gods, from being devoured by a wolf (lit. 'from the devouring of a wolf')'. This neutralizes parameter CCN.
11. Structured Adjective Phrases (AST): yes. Cf. RV 3.15.7 *bṛhadúkṣaḥ marú-taḥ viśvávedasaḥ* 'the dripping/raining Maruts, the all-knowing ones' (Adj–N–Adj, all inflected for nom.pl). Or RV 8.41.4e *sá hí gopāḥ iva íryaḥ* 'for he (is) like an active herdsman' with N–Adj.

12. Feature spread to structured APs (FFS): yes, see AST directly above. Another illustrating example is RV 3.2.15, or RV 9.1.1 *svādiṣṭhayā mādiṣṭhayā ... dhārayā* ‘in sweetest and most inebriating stream...’ (all inflected for INS.SG.F).

13. Feature spread to predicative APs (FSP): yes, cf.

- (55) Vedic, feature spread to predicative adjectives (RV 7.56.5)  
*sā*                      *vīt*                      *suvīrā*                      *marúdbhiḥ*  
 this.NOM.SG.F tribe.NOM.SG.F heroic.NOM.SG.F Marut.INST.PL.  
*astu*                      *sanāt sáhantī*  
 be.IMP.PRES.ACT.3SG ever victorious.NOM.SG.F  
*púṣyanti*                      *nṛmṇám*  
 prospering.NOM.SG.F manliness.ACC.SG  
 ‘Let this tribe be heroic through the Maruts, ever victorious, prospering in manliness.’

14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. Postposed relative clauses prevail in Vedic (see Holland 1986); see also Hettrich (1988) for examples. This neutralizes parameter AER. One example:

- (56) Vedic, NP raising over relative clauses (RV 9.2.7)  
*gírah*                      ... *yābhiḥ*                      *mádāya*  
 song.NOM.PL ... REL.INS.PL intoxication.DAT.SG  
*śúmbhase*  
 adorn.IND.PRS.MED.2SG  
 ‘Songs... with which you adorn (yourself) for the intoxication.’

15. Free reduced RelCl (ARR): set to yes. Vedic, like other ancient Indo-European languages, has free adjectives in form of verbal adjectives and participles functioning as reduced relative clauses, the latter of these also in so-called

absolute constructions. See Keydana (1997) for examples. The frequent usage as an adjunct can be illustrated by RV 6.47.18 *yuktāḥ hí asya hárayaḥ śatā dáśa* ‘for yoked (are) his ten hundred fire-colored (horses)’; the first word here is a verbal adjective (‘yoked’).

16. Adpositional genitive (GAD): no, genitives are inflectional.

17. Free Gen. (GFR): yes, there are free, recursive genitives, e.g.,

(57) Vedic, free recursive genitives (RV 3.15.1)

*suśármaṇaḥ                      bṛhatáḥ                      śármaṇi*  
 good.protector.GEN.SG high.GEN.SG protection/shelter.LOC.SG  
*syām                      agnéḥ                      ahám                      suhávasya*  
 be.OPT.PRS.ACT.1SG Agni.GEN.SG I.NOM.SG easily.invoked.GEN.SG  
*práñītau*  
 lead.LOC.SG

‘I want to be under the protection and lead of the good protector,  
 of the high one, of Agni, of the easily invoked one.’

Another instance may be seen in RV 9.58.2:

(58) Vedic, free recursive genitives (RV 9.58.2a-b)

*usrá                      veda                      vásūnām*  
 usrá.NOM.SG.F know.IND.PRF.ACT.3SG good.GEN.PL.N  
*mártasya                      devī                      ávasaḥ*  
 mortal.GEN.SG.M goddess.NOM.SG.F support.GEN.SG.N

‘Dawn, the goddess, knows the good (things) of the support of (= by) the mortal one/man.’

Griffith translates this as “The Morning knows all precious things, the Goddess knows her grace to man”, Geldner as „Die Göttin Morgenröte kennt die Schätze, die Gunst des Sterblichen“, Renou finally gives »L’Aurore connaît les biens-matériels du mortel, la déesse (connaît comment lui octroyer

sa) faveur». <sup>28</sup> On formal grounds, my translation is perfectly possible, interpreting all words of lines a and b as forming one single NP instead of two (as Griffith and Renou seem to assume). Either *usrā* or *devī* would then be an appositive to the other, and all three genitives were dependent upon each other (Geldner apparently reckons with two genitival arguments, namely with a two-tiered genitive and a separate, basic one). In any case, Vedic had the possibility to recursively link genitives to other genitives.

18. Uniform Gen. (GUN): yes, there is a uniform genitive being inflectional and not bound to certain structural positions. Cf., besides the examples of GFR above, the following instances:

(59) Vedic, uniform inflectional genitive occupying different structural positions

- a. Uniform genitive, postnominal (RV 1.3.11)

*codayitrī*                      *sūnṛt-ān-ām*                      *cétantī*  
 inciter.NOM.SG.F pleasant.song-GEN.PL inspiring.NOM.SG.F  
*sumat-īnām*  
 favor-GEN.PL  
 ‘Inciter of hymns/pleasant songs, inspirer of favors.’

- b. Uniform genitive, prenominal (RV 1.4.2)

*sómasya*                      *somapāh*  
 Soma.GEN.SG drinker.of.Soma.VOC.SG  
 ‘(Come,) (Soma)drinker of Soma.’

- c. Uniform genitive, recursive, postnominal (RV 3.1.8)

*ścótanti*                                      *dhārāḥ*                                      *mādhunāḥ*  
 trickle.IND.PRS.ACT.3PL stream.NOM.PL sweet.juice.GEN.SG  
*ghṛtāsya*  
 cream/fat.GEN.SG  
 ‘Streams of sweet juice, of fat/cream trickle.’

<sup>28</sup>Following VedaWeb (Casaretto et al. 2023, accessed 8–14–2023).

Note, by the way, the verb-raising in example c, showing that not only N can be raised (see the N-raising parameters below), but also the verb. Many more examples showing freedom of placement could be brought forward (e.g., RV 3.1.7.b, etc.), but this should suffice to demonstrate the presence of a genitive not bound to any structural limitations. This presence of a uniform genitive neutralizes parameter GFO.

19. DP over free Gen. (GPR): yes, (free) genitives can occur both before and after the rest of the NP/DP, so in generative terminology, DPs/NPs can be raised over free genitives. See the examples of GFR above.
20. Gen-feature spread to N (GFS): no. Genitives form dependent NPs with internal number and gender agreement (if overt and possible due to the presence of adjectives or demonstratives), but they do not influence their head nouns. Consider example 58.
21. Adjectival poss. (APO): yes, see Debrunner and Wackernagel (1975, pp. 492–494). Vedic can express possession *inter alia* via adjectival possessives (those in plain thematic *-a-* rarely used), cf. RV 2.20.2 *tvābhiḥ ūtī* ‘with your protection(s) (INS.PL)’, others built with a *-ka-* suffix more often (e.g., in RV 1.31.11 *māmakasya* ‘my.GEN.SG’).
22. Clitic poss. (PCL): yes, Ved. has the possibility to express possession by means of enclitic pronouns, see Debrunner and Wackernagel (1975, pp. 492–493). Two examples: RV 1.80.8 *bāhvós te* ‘in your arms’, or RV 1.55.7 *yāmiṣṭhāsaḥ sārathayaḥ yé indra te* ‘the best steering charioteers which are yours’ (or ‘your charioteers which are the best in steering’ vel sim.).
23. N-feature spread to free Gen. (GSP): no. Genitives do not agree in gender or number with their head nouns. See the examples of GFR or GUN above.
24. Adjectival Gen. (AGE): no. Vedic does not have adjectival genitives in the sense of this parameter.



25. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Vedic.
  
26. Strong partial locality (TPL): no. TPL asks whether or not a subset of demonstrative pronouns, viz. deictic ones, systematically surface in the D-area. This is not the case. Even though demonstratives can introduce NPs, there are many cases in which demonstratives follow their noun (see NDE immediately below). Thus this parameter must receive a negative value. This neutralizes parameter TSL.
  
27. N over Demonstratives (NDE): yes. Cf., for instance, RV 1.3.4: *sutāḥ imé* ‘these Soma-juices/libations’ with N–D, or RV 10.97.6c: *vípraḥ sáḥ* (*ucyate bhiṣák*) ‘this poet (is called healer)’ with N–D (*sáḥ* here functions as a demonstrative and not as a 3rd person pronoun). Illustrative also RV 6.47.1 *svādúḥ kíla ayám mádhumān utá ayám tīvráḥ kíla ayám rásavān utá ayám* ‘Truly this one (is) tasteful and full of sweetness, spicy indeed and strong (is) this’ (four times N–D).
  
28. N over Cardinals (NOC): yes, cf. RV 9.103.3: *vāṇīḥ ṛṣīṇām sapṭá* ‘the seven voices of the Ṛṣis’ (with N–Gen–Card), or RV 4.26.3: *púraḥ (...) náva (...)* *navatīḥ* ‘ninety-nine forts’ (with N–Card–Card), or RV 8.21.10 etc. Also RV 6.47.18 *yuktāḥ hí asya hárayaḥ śatá dáśa* ‘for yoked (are) his ten hundred fire-colored (horses)’. Debrunner and Wackernagel (1975, p. 338) give another example (Sanskrit, not Vedic) from the Mahābhārata (13.101.5), *narakam trimśatam* ‘thirty hells’ (N–Card, with N being inflected for sg.). The other order, Card–N, can be seen in e.g., RV 6.47.23–24.
  
29. N over Ordinals (NOO): unclear. Debrunner and Wackernagel (1975, 400ff.) present no clear examples of the expected N–Ord order, only ordinals used predicatively (for instance, RV 10.85.45 *pátim ekādaśám kṛdhi* ‘make the

husband the eleventh.’ The order Ord–N, being in accordance with the concept of a ‘base-generated’ order, is attested, but uninteresting for our purposes, as we are interested in the amount of N-raising. The hierarchy predicts that NPs with N–Ord should be possible; quite surprisingly, there seems to be no instance of such an order.

30. N over adjectives (NOA): yes, cf. e.g., RV 1.82.5 *jāyā́m (...) priyā́m* ‘beloved wife’ or RV 3.25.3: *devī́ amṛ́te* ‘(two) immortal goddesses (ACC.DU.F)’, with N–Adj; cf. also RV 3.2.5 etc.
31. N over GenO (NGO): yes, cf. RV 3.10.11: *tvā́m (...) samrājā́m carṣaṇī́nām* ‘thee (, O Agni,) the overlord/supreme ruler (*samrājā́m*, ACC.SG) of all living beings (*carṣaṇī́nām*, GEN.PL)’, thus N–GenO.
32. N over external arguments (NOE): yes, examples under parameters NDE, NOO, NGS, NOA and NGO.
33. Verbless relative clauses (VRC): yes, though apparently in the RV only with nominatives, cf. RV 5.13.3b: *hótā yáḥ mā́nuṣeṣu á* ‘the priest among mankind’ or RV 7.34.24: *vísve marútaḥ yé sahā́saḥ* ‘all the Maruts who (are) powerful’. The Atharva-Veda evinces also non-nominative case agreement (see Caland 1897), cf. AV 19.20.1 *paúruṣeyaṁ vadháṁ yám indrā́gnī́ dhātā́ savitā́ bṛ́haspátīḥ* ‘man-made/caused death/murder which (was appointed by) Indra-Agni, Dhātar, Savitar and Brhaspati.’ Even though these are isolated instances, Vedic nonetheless has the ability to specify NPs by means of verbless relative clauses with *attractio relativi*.
34. Facultative linker (FLI): no. There is no linker construction, only verbless relative clauses.
35. Generalized linker (GLI): no.

The most interesting observation relates to the fact that heads (nominal or verbal ones) can be raised over all ‘base-generated’ preceding arguments. Even

though the impression one gets when reading the *Ṛgveda* is that of Vedic being a head-final language in the nominal domain, this is not true in parametric terms. Vedic has the possibility to raise nouns over other elements. This option is more often used with hierarchically lower-ranking genitives and adjectives and rather infrequently with demonstratives and (cardinal) numerals. Yet the mere possibility is given, and this stands in a sharp contrast with modern Indo-Aryan languages like Hindi or Marathi which are rigidly head-final (without giving here a full account of their parametric values). Parametrically, Vedic and its modern descendants are at the opposite ends of the theoretically possible spectrum, at least with regard to N-raising parameters.

### 3.5.2 (Archaic) Greek

The Greek branch of Indo-European is attested from the second millennium BCE onward. The Mycenaean material mostly consists of lists and inventories with a very stereotypical character, and not all parametric settings can be securely determined. Younger stages of Greek with more varied texts instead allow for a full parametric characterization of Greek. In the following pages, I will mainly give examples from the language of epic poetry, i.e., the Ionic-Aeolic dialect of the *Iliad* and *Odyssey* (“Homeric Greek”) or the Ionic one of Hesiod, but if feasible I will also present examples from Mycenaean Greek.

1. Gramm. Person (FGP): set to yes. Cf. the distinction between 1st, 2nd and 3rd person pronouns and the contrast in terms of deictic demonstratives between proximate ὅδε, distant κεῖνος (classical ἐκεῖνος) and intermediate οὗτος.
2. Gramm. Number (FGN): yes. There is a difference between singular, dual and plural. The dual appears to be regularly used in Mycenaean Greek (see Bartoněk 2003, pp. 157–158), whereas in Homeric Greek it is no longer obligatory and even NPs built with the numeral ‘two’ can be formed with nouns marked for plural (e.g., *Od.* 10.515 δῶ ποταμῶν ἐριδούπων ‘two thundering rivers’). In any case, Greek has number marking on nouns and NPs.
3. Gramm. Gender (FGG): yes. (Archaic) Greek distinguishes between masculine, feminine and neuter gender.
4. Feature spread to N (FSN): yes. Cf. e.g., *Il.* 6.1 χαλκήρεα δοῦρα ‘iron spears’, inflected for plural number. Also *Il.* 20.233 κάλλιστος (...) θνητῶν ἀνθρώπων ‘the fairest of mortal men’; or *Il.* 2.1 ἄλλοι μὲν ῥά θεοί τε καὶ ἄνδρες ἵπποκορυσταὶ ‘now all the other chariot-equipped gods and men’.
5. Number on N (bounded nouns) (FNN): yes, cf. the preceding examples of (FSN).

6. Gramm. partial definiteness (DGP): no. Classical Greek has an article, but the earliest stages of Greek do not. Mycenaean has no article or obligatory anaphoric definiteness marking, cf. PY Na 568: to-sa-de na-u-do-mo o-u-di-do-si *tos(s)ade naudomoi ou didonsi* ‘the shipwrights do not deliver that much’, or PY Vn 10 o-di-do-si du-ru-to-mo (...) (*h*)ōs *didonsi drutomoi* ‘as/so the loggers deliver...’. Homeric Greek likewise does not have obligatory anaphoric definiteness marking. ὁ, ἡ, τό (which were to become the definite articles of classical Greek) are frequently used, but they still have a demonstrative meaning in many instances (see Jacquinod 2017, p. 685). There are many definite nouns or NPs which are not formally marked in any way for (anaphoric or general, implicit) definiteness; cf. e.g., Il. 2.546–47 Ἀθήνας (...) ἐϋκτίμενον πτολίεθρον ‘Athens, the well-built polis’, or Il. 10.1: ἄλλοι μὲν παρὰ νηυσὶν ἀριστῆες Παναχαιῶν εὔδον παννύχιοι ‘Now beside their ships the other chieftains of all the Achaeans slept the whole night through’ (neither the ships – νηυσὶν – nor the chieftains – ἀριστῆες – are marked for definiteness). Or cf. Il. 13.1: Ζεὺς δ’ ἐπεὶ οὖν Τρῳᾶς τε καὶ Ἑκτορα νηυσὶ πέλασσε ‘After Zeus had brought the Trojans and Hector to the ships...’. Another example, showing a definite NP not receiving any definiteness marker (and the demonstrative use of the prospective article), is Il. 13.404: ἀλλ’ ὃ μὲν ἄντα ἰδὼν ἠλεύατο χάλκεον ἔγχος Ἰδομενεύς ‘he (or: this one), looking steadily at him, avoided the spear of bronze, Idomeneus’. Also Il. 5. 365: καὶ ἡνία λάζετο χερσὶ ‘and she took the reins in (her) hands’; etc. (examples abound; many more could be given). Again, this neutralizes 14 other parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA).
7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking. As in the oldest Indo-Iranian languages, there is a partitive genitive, but this does not permit a positive setting of this parameter.

8. D-controlled infl. on N (DIN): no. Ancient Greek has nothing resembling the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes. Cf. Il. 20.221: τρισχίλια ἵπποι ‘three thousand steeds’, or Il. 20.231: τρεῖς παῖδες ἀμύμονες ‘three blameless boys’.
10. Grammaticalized boundedness (CGB): no. Nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. This can be demonstrated with Mycenaean material, cf. PY Ta 722: ta-ra-nu a-ja-me-no e-re-pa-te-jo a-to-ro-qo i-qo-qe po-ru-po-de-qe po-ni-ke-qe FOOTSTOOL I ‘one footstool (*t<sup>h</sup>rānus*) inlaid with an ivory (picture of) a man, a horse, an octopus, and a palm tree (?)’. There are several singular count nouns, all inflected for dative singular (a-to-ro-qo *ant<sup>h</sup>rōk<sup>u</sup>ōi* ‘a man’; i-qo-qe *hikk<sup>u</sup>ōi-k<sup>u</sup>e* ‘and a horse’; po-ru-po-de-qe *polupodei-k<sup>u</sup>e* ‘and an octopus (lit. many-footed)’; po-ni-ke-qe *p<sup>h</sup>oinikei-k<sup>u</sup>e* ‘and a palm-tree [or phoenix; or rather an adjective crimson/purple?]’), but none of them requires a special marker to signal boundedness. The same is true with regard to Epic Greek, cf. Hesiod, Works and Days, 71: παρθένω αἰδοίη ἵκελον ‘similar to a modest maiden/virgin’. The noun παρθένω is not marked for ‘boundedness’ by means of an indefinite article or another overt marker. Cf. also Il. 3.33: ὥς δ’ ὅτε τίς τε δράκοντα ἰδὼν ‘and even like a man (or: someone) when he sees a snake’. This neutralizes parameter CCN.
11. Structured Adjective Phrases (AST): yes. Cf. Il. 13.423: νῆας ἔπι γλαφυράς ‘to the hollow ships (acc. pl)’, Il. 5.550: μελαινάων ἐπὶ νηῶν ‘on the black ships’ (dat. pl.), or Il. 5.238: ὀξεῖ δουρί ‘with the sharp spear’ (dat. sg.).
12. Feature spread to structured APs (FFS): yes, see AST directly above.
13. Feature spread to predicative APs (FSP): yes, cf.

- (60) Ancient Greek, feature spread to predicative adjectives (Od. 4.437)  
 πάντα δ' ἔσαν νεόδαρτα  
 all.NOM.PL.N but be.IND.IPRF.ACT.3PL newly.flayed.NOM.PL.N  
 '(but) all were newly flayed.'

14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. One example from the Iliad (5.761): ἄφρονά τοῦτον (...), ὃς οὐ τίνα οἶδε θέμιστα 'this madman who does not know (obey) any law'. This neutralizes parameter AER.
15. Free reduced RelCl (ARR): set to yes. Archaic Greek, like other ancient IE languages, has free adjectives in form of participles or verbal adjectives functioning as reduced relative clauses. Cf. Mycenaean (MY Oe 127) a-re-pa-te ze-so-me-no *aleip<sup>h</sup>atei dzes(s)omenōi* 'for the ointment to be cooked'. One example from the Iliad (5.773): ποταμὸν τε ρέοντε 'and the two flowing rivers', or 'and the two rivers (which were) flowing'.
16. Adpositional genitive (GAD): no, genitives are always inflectional and receive a case marker/ending.
17. Free Gen. (GFR): yes, there are free, recursive genitives, e.g.,

- (61) Archaic Greek, free recursive genitives
- a. Mycenaean (Eo 247, following Bartoněk 2003, p. 443)
- |   |   |                     |
|---|---|---------------------|
| <i>ek<sup>h</sup>ei-k<sup>u</sup>e</i>  | <i>onāton</i>                           | <i>ktimenās</i>     |
| e-ke-qe                                 | o-na-to                                 | ki-ti-me-na         |
| have.IND.PRS.3SG-and leased.land.ACC.SG |   | cultivated.GEN.SG.F |
| <i>ktōinās</i>                          | <i>ait<sup>h</sup>iok<sup>u</sup>os</i> |                     |
| ko-to-na                                | a <sub>3</sub> -ti-jo-qo                |                     |
| land.parcel.GEN.SG.F                    | PN.GEN.SG.M                             |                     |
- 'and has leased land of the cultivated land parcel of Ait<sup>h</sup>iok<sup>u</sup>os.'

## b. Homeric Greek (Il. 5.796)

ιδρώς                      γάρ μιν ἔτειρεν                      ὑπὸ  
 sweat.NOM.SG.M for him distress.IPRF.ACT.3SG under  
 πλατέος                      τελαμῶνος                      ἀσπίδος  
 broad.GEN.SG.M baldric.GEN.SG.M shield.GEN.SG.F  
 εὐκύκλου  
 well.rounded.GEN.SG.F  
 'For the sweat vexed him beneath the broad baldric of his well-  
 rounded shield.'

18. Uniform Gen. (GUN): yes, there is a uniform genitive that is inflectional and not bound to certain structural positions. Cf. the following instances:

(62) Ancient Greek, uniform inflectional genitive occupying different structural positions

## a. Uniform genitive, postnominal (Il. 5.827f.)

τιν'                      ἄλλον                      ἀθανάτων  
 someone.ACC.SG other.ACC.SG immortal.GEN.PL  
 '(do not fear) any other of the immortals'.

## b. Uniform genitive, prenominal (Il. 21.184-185)

χαλεπόν                      τοι ἔρισθενέος Κρονίωνος  
 difficult.ACC.SG.N surely huge.GEN.SG son.of.Kronos.GEN.SG  
 παισὶν                      ἐριζέμεναι  
 child.DAT.PL fight.INF  
 'It is (surely) difficult to fight with the children of the mighty  
 son of Kronos.'

## c. Uniform genitive, prenominal and dislocation (Il. 5.790)

κείνου                      γὰρ ἐδείδισαν                      ὄβριμον                      ἔγχος  
 that.GEN.SG for fear.ACT.PLPF.3PL mighty.ACC.SG spear.ACC.SG  
 'for they had feared the mighty spear of that one.'

The presence of a uniform genitive neutralizes parameter GFO.

19. DP over free Gen. (GPR): yes, (free) genitives can occur both before and after the rest of the NP/DP, so in generative terminology, DPs/NPs can be raised over free genitives. See the example of GFR above.



20. Gen-feature spread to N (GFS): no. Genitives do not ‘spread’ their features to their head nouns; both agree with specifying adjectives or demonstratives, respectively, but a genitival argument is an NP dependent on another, higher ranking NP.
21. Adjectival poss. (APO): yes, Greek has adjectival possessives. See, for instance, Il. 2.374 *χερσὶν ὑφ’ ἡμετέρησιν* ‘beneath our hands’ or Il. 5.234 *τεὸν φθόγγον* ‘your voice’.
22. Clitic poss. (PCL): yes, Ancient Greek has the possibility to express possession by means of enclitic pronouns. Cf. e.g., Il. 7.52 *οὐ γάρ πώ τοι μοῖρα θανεῖν* ‘For it is not yet your (τοι) fate to die.’
23. N-feature spread to free Gen. (GSP): no. Genitives show agreement with adjectives or demonstratives further specifying them (e.g., Il. 21.184–185 *ἐρισθενέος Κρονίωνος παῖσιν* ‘with the children of the mighty son of Kronos’), but not with their head noun.
24. Adjectival Gen. (AGE): yes, cf. Il. 10.326 *νῆ’ Ἀγαμεμνομένην* ‘Agamemnon’s ship’ or Il. 2.658 (etc.) *βίῃ Ἡρακλείῃ* ‘for Herakles’s force’.
25. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Archaic Greek.
26. Strong partial locality (TPL): no. TPL asks whether or not a subset of demonstrative pronouns, viz. deictic ones, systematically surface in the D-area. Demonstratives frequently, but not systematically surface there; they can also surface after nouns and adjectives (N-raising). See NDE below; as before, negative TPL neutralizes TSL.
27. N over Demonstratives (NDE): yes. See Il. 2.37 *ἥματι κείνῳ* ‘on that day’ (also Hes. *Theogony*, 667, 836), Il. 9.439 *ἥματι τῷ* ‘on this day’, or Il. 6.38

ἵππῳ γάρ οἱ ἀτυζομένῳ ‘for these two horses, being in terror...’.<sup>29</sup>

28. N over Cardinals (NOC): yes, cf. Od. 4.408–409 ἑταίρους τρεῖς ‘three companions’ or Il. 5.554: τῷ γε λέοντε δύο ‘two lions’.
29. N over Ordinals (NOO): yes, cf. Il. 21.(155–)156 (ἦδε δέ μοι νῦν) ἡὼς ἑνδεκάτη ‘(this is now my) eleventh dawn/morning...’. It may be remarked that the great majority of NPs built with ordinals has an Ord–N order in Epic Greek. However, this is entirely irrelevant. All that matters is the fact that speakers of Archaic Greek apparently had the possibility to raise nouns over ordinals. How frequently this option is or was chosen, has no effect on the positive parametric value, as syntax is not about quantitative evidence, but about syntactic options.
30. N over adjectives (NOA): yes, cf. e.g., the first verse of the Od. (1.1): ἄνδρα (μοι ἔννεπε, μοῦσα,) πολύτροπον ‘(tell me, Muse, of) the man of many devices’.
31. N over GenO (NGO): yes, cf. Il. 9.416 τέλος θανάτοιο ‘the end of death’ (but note the reverse order in 9.411).
32. N over external arguments (NOE): yes, see the examples from the parameters NGO, NOA, NOO, NOC and NDE above.
33. Verbless relative clauses (VRC): yes. In general, verbless relative clauses are restricted to early forms of Greek (Probert 2015, p. 413); to give but one example from the Iliad (15.91): ἦ μάλα δὴ σ’ ἐφόβησε Κρόνου παῖς, ὅς τοι ἀκοίτης ‘Truly, the son of Kronos, who [is] your companion, has put you to flight’. A related phenomenon involves the so-called *attractio relativi* in which a relative pronoun agrees with its nominal head instead of being inflected in the nominative case; *attractio relativi* is very rare in

---

<sup>29</sup>Note that the demonstrative/proto-article here is a plural and no dual form.

verbal clauses in Homeric Greek and more common in nominal (i.e., verbless) relative clauses (see Chantraine 1953, p. 237 or Viti 2013, p. 95). Consider *Od.* 19.232–233 τὸν δὲ χιτῶν' ἐνόησα περὶ χροῖ σιγαλόεντα, / οἷόν τε κρομύοιο λοπὸν κάτα ἰσχαλέοιο 'I noticed the tunic, gleaming on his body like the skin of a dried onion', instead of (...) οἷός τε κρομύοιο λοπὸς κάτα ἰσχαλέοιο. Other instances are *Il.* 1.262, *Od.* 9.321–322, 9.325, 10.167, 10.517, 11.25.

34. Facultative linker (FLI): no. There is no linker in Greek, only verbless relative clauses.
35. Generalized linker (GLI): no.

### 3.5.3 (Old) Latin

Latin shall serve as a representative of the Italic branch of Indo-European. It is the best attested member and provides enough material for a secure parametric characterization. Two other Italic languages can be expected to yield enough material as well if investigated (namely Oscan and Umbrian), whereas the Venetic or Faliscan corpora probably are too small. I will focus mostly on Old Latin, i.e., on the Latin of early inscriptions (Duenos, Garigliano bowl etc.) and the works of Plautus and Terence, though occasionally supplemented by examples from classical Latin.

1. Gramm. Person (FGP): set to yes. Latin has verbal person and a tripartite system of deictic demonstratives *hic, iste, ille*.
2. Gramm. Number (FGN): yes. There is a morphosyntactic difference between singular and plural. Cf. Plaut. Trin.<sup>30</sup> 833: *distraxissent disque tulissent satellites tui me miserum foede* ‘Your accomplices would have horribly torn apart and dispersed poor me’, with *satellites* marked for plural and *me miserum* for singular.
3. Gramm. Gender (FGG): yes. Latin distinguishes between masculine, feminine and neuter gender. Visible, e.g., in the ‘altar to the unknown deity’: SEI·DEO·SEI·DEIVAE·SAC ‘whether sacred to a god or goddess...’.
4. Feature spread to N (FSN): yes. Cf. e.g., Plaut. Poen. 1062 *ecquid meministi tuom parentum nomina* ‘Do you remember at all the names of your parents?’, with *nomina* being marked for plural, and the dependent genitive phrase *tuom parentum* as well. Or consider the Duenos inscription, line 1: IOVESATDEIVOS ‘he prays to the gods’, with *deiuos* ‘gods’ being marked for (acc.) plural.
5. Number on N (bounded nouns) (FNN): yes, cf. the preceding example of (FSN).

---

<sup>30</sup> Abbreviations of authors and works according to the Oxford Latin Dictionary.

6. Gramm. partial definiteness (DGP): no. There are cases like Pl. Trin. 493–494 *aequo mendicus atque ille opulentissimus censetur censu ad Acheruntem mortuos* ‘the beggar is held of equal value at Acheron with the most wealthy man when dead’ in which *ille* appears to have the function to mark *opulentissimus* ‘most wealthy(one/man)’ for definiteness. But NPs without demonstrative or rather anaphoric pronoun are not rare; often both exist side by side. Consider e.g., Plautus (Merc. 236–239) *ait sese illius opera atque adventu caprae flagitium et damnum fecisse haud mediocriter; dicit capram, quam dederam servandam sibi, suae uxoris dotem ambedisse oppido*. ‘he said that by her means and through the arrival of the she-goat he had suffered injury and loss in no slight degree; he said that the she-goat, which I had entrusted to him to keep, had gnawed away the marriage-portion of his wife.’ *capra*- ‘she-goat’ is at first specified by a demonstrative (*illius ... caprae*), whereas later on it does not get a definite marker (just *capram*). This is a common pattern; definiteness can, but need not be marked. Classical Latin likewise does not have obligatory definiteness marking of nouns or NPs, cf. e.g., Cic. Nat. D. 2.103 *luna autem quae est...* ‘the moon, however, which is...’.<sup>31</sup> The conclusion is that there is (or was) no systematic and grammaticalized anaphoric definiteness marking in Latin (neither in Old nor classical Latin). Once again, this neutralizes several other parameters.
7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking. There is a partitive genitive, but this does not permit a positive setting of this parameter.
8. D-controlled infl. on N (DIN): no. Latin has nothing resembling the nunation of Arabic.

---

<sup>31</sup>Roughly 1650 years before Galilei discovered, for the first time in human history, moons around another celestial body, there was only one ‘moon’ known to mankind; *luna* was inherently definite.

9. Plural spread from Cardinals (CPS): yes. Cf. the Garigliano bowl (Vine 1998): ESOMKOMMEOISSOKIOISTRIFOI (~ *sum cum meīs sociīs tribus*) 'I am with my three companions', in which 'three' triggers feature spread of overt morphological plural marking to the noun (and the poss. adjective).
10. Grammaticalized boundedness (CGB): no. Singular count nouns are not inherently unbounded and thus do not require special affixes or markers to get a bounded reading. Consider the 'altar to the unknown deity', SEI·DEO·SEI·DEIVAE·SAC 'whether sacred to god or goddess...'. Another example (again the 'she-goat' of parameter DGP; Plaut. Merc. 229): *mercari visus mihi sum formosam capram* 'I seemed to have purchased for myself a beautiful she-goat'; *capram* is not marked for 'boundedness'. This neutralizes parameter CCN.
11. Structured Adjective Phrases (AST): yes. Cf. the just mentioned example (Plaut. Merc. 229): *formosam capram* 'a beautiful she-goat' (Adj–N), or Plaut. Merc. 260: *forma eximia (mulierem)* '(a woman) of extraordinary beauty' (N–Adj). Or see l.c., 398: *cottidianum familiae coctum cibum* 'everyday's food cooked for the household', Adj–Gen–Adj–N, with the two adjectives and the noun agreeing in terms of number (sg.), gender (m.) and case (acc.).
12. Feature spread to structured APs (FFS): yes, see AST directly above.
13. Feature spread to predicative APs (FSP): yes, predicative adjectives must agree with their head nouns in terms of gender, number and case, cf.

- (63) Latin, feature spread to predicative adjectives (Duenos inscription, line 1)

<i>nei tēd</i>	<i>endō</i>	<i>cosmis</i>	<i>virco</i>
NEI TED	ENDO	COSMIS	VIRCO

lest you.ACC.SG towards kind.NOM.SG.F virgin.NOM.SG.F

*siēd*

SIED

be.OPT.PRS.3SG

‘lest the girl be not kind towards thee.’

14. DP over relatives (ADR): yes. Relative clauses can precede or follow their head nouns, and since they can occur/surface after the NP core (or, from another point of view, the NP can be raised over the relative clause), this parameter has a positive value. Cf. Plautus (Merc. 238) *capram, quam dederam* ‘the she-goat which I had given...’. An example from the *Senatus consultum de Bacchanalibus*, lines 2–3: DE BACANALIBVS QVEI FOIDERATEI ESENT ITA EXDEICENDVM CENSVERE ‘Regarding the Bacchanalia, they decided to give the following directions to those who are in alliance with us’ (cf. also l. 24). Consider also the Duenos inscription, line 1: IOVESATDEIVOSQOIMEDMITAT (transcription with macrons and word breaks: *iouesāt deiuos qoi mēd mitāt*) ‘He who gives me prays to the gods’ – note that the head noun/NP is not overt, but expressed in form of a verbal ending. This sentence also illustrates verb raising in Old Latin, by the way. Another interesting fact to note: the *Senatus consultum de Bacchanalibus*, lines 3–4, has a dependent relative clause surfacing after its head relative clause (free, recursive relative clause, so to say): SEI QVES ESENT QVEI SIBEI DEICERENT (~ *Sī quī essent, quī sibi dicerent*) ‘if there were any who say that they...’. Positive ADR neutralizes parameter AER.
15. Free reduced RelCl (ARR): set to yes. Latin, like other ancient Indo-European languages, has free adjectives in form of participles or verbal adjectives functioning as reduced relative clauses. Cf. Pl. Trin. 493–494 *aequo mendicus atque ille opulentissimus censetur censu ad Acherun-*

*tem mortuos* ‘the beggar is held of equal value at Acheron with the most wealthy man **when dead**’.

16. Adpositional genitive (GAD): no, genitives are always inflectional.

17. Free Gen. (GFR): yes, there are free, recursive genitives, e.g.,

(64) Latin, free recursive genitives (Cic. Tusc 2.35)

*functio*                      *quaedam*                      *vel animi*                      *vel*  
 performance.NOM.SG.F a.certain.NOM.SG.F or mind.GEN.SG or  
*corporis*                      *gravioris*                      *operis*  
 body.GEN.SG heavier.GEN.SG work.GEN.SG  
 ‘A certain performance of (by) the mind or of some significant work  
 of (= performed by) the body.’

18. Uniform Gen. (GUN): yes, there is a uniform genitive being inflectional and not bound to certain structural positions. Cf. the following instances:

(65) Latin, uniform inflectional genitive occupying different structural positions

a. Uniform genitive, occupying functions of GenS and GenO (Caesar, BG 1.30)

*prō veteribus*                      *helvētiōrum*                      *iniūriīs*  
 for old.ABL.PL Helvetian.GEN.PL injustice.ABL.PL  
*populī*                      *rōmānī*  
 people.GEN.SG Roman.GEN.SG  
 ‘for previous injuries inflicted by the Helvetii on the Roman people’ (Devine and Stephens 2006, p. 316).

b. Uniform genitive, postnominal (Caes. BC 1.5)

*in dēspērātiōne*                      *omnium*                      *salūtis*  
 in desperation.ABL.SG all.GEN.PL salvation.GEN.SG  
 ‘Under circumstances in which everyone despaired of salvation’ (Devine and Stephens 2006, p. 316).



- c. Uniform genitive, prenominal ('altar to the unknown deity')

*dē senātī sententiā*  
 DE SENATI SENTENTIA  
 from senate.GEN.SG vote.ABL.SG  
 'by means of/on a vote of the Senate.'

The presence of a uniform inflectional genitive neutralizes parameter GFO.

19. DP over free Gen. (GPR): yes, (free) genitives can occur both before and after the remaining parts of the NP/DP, so in generative terminology, DPs/NPs can be raised over free genitives. See the example of GFR above.
20. Gen-feature spread to N (GFS): no. Genitives do not spread their features to their head nouns.
21. Adjectival poss. (APO): yes, Latin has adjectival possessives, see e.g., the Garigliano bowl (Vine 1998): ESOMKOMMEOISSOKIOISTRIFOI (~ *sum cum meīs sociīs tribus*) 'I am with my three companions', MEOIS is the adjectival possessive. This sentence, by the way, also shows high verb raising. Another instance from Plautus (Most. 162): *haec illa est tempestas mea* 'this is the storm (I mentioned)'.
22. Clitic poss. (PCL): yes, Latin has the possibility to express possession by means of enclitic pronouns (attaching to the first stressed item in a clause). Cf. the difference between clitic *me* and orthotone *meum* (both acc.sg. 'me') in Plaut. Merc. 961: *Sequere me. sed exeuntem filium video meum* 'Follow me – but I see my son coming out'.
23. N-feature spread to free Gen. (GSP): no. Genitives show internal agreement between elements, but there is no 'spread' of features from the noun to genitives.
24. Adjectival Gen. (AGE): yes, cf. Ter. Eun. 289 *erilem filium* 'the master's son', Plaut. Mil. 458 *erilis concubīna* 'the master's concubine', or NOUTRIX

PAPERIA ‘nurse of Paperius’ (ILLRP 81, Nemi). Note that genitival adjectives and regular genitives had certain functional differences, as noted by Weiss (2020, pp. 476–477).

25. Poss.-checking N (GCN): no. This parameter asks if distinctive morphological marking (head marking, such as the Semitic construct state) occurs on nouns when occurring with a genitive argument. This is not the case in Latin.
26. Strong partial locality (TPL): no. TPL asks whether or not a subset of demonstrative pronouns, viz. deictic ones, systematically surface in the D-area. Demonstratives frequently, but not systematically (in the sense of this parameter) surface in the D-area; there are many NPs with N–Dem order (see NDE below). This neutralizes parameter TSL.
27. N over Demonstratives (NDE): yes. cf. Plaut. Merc. 274: *(Nec) omen illud (mihi nec auspicium placet)* ‘(Neither) this omen (nor this augury pleases me)’. Another instance (Pl. Merc. 349–350): *nec pater potis videtur induci ut putet matri ancillam emptam esse illam* ‘it doesn’t seem possible for my father to be induced to think that she was bought as a maid-servant for my mother’, with the head noun *ancillam* and a postposed demonstrative *illam*. Likewise classical Latin, despite showing much more frequently a Dem–N order, had the possibility to raise nouns over demonstratives, cf. Cicero, Top. 53 (as well as Orat. 162 and De Inv. 1.57) *locus hic* ‘this place’, similarly *status hic* ‘this state’ (De Fin. 2.32, Ad Fam. 1.7.10, Ad Att. 2.22.1). See Devine and Stephens (2006, 512ff.) for several other examples from classical Latin.<sup>32</sup>

<sup>32</sup>Maybe a N–Dem order is also attested in line 3 of the Duenos inscription. Deviating from common interpretations (verbal abstract noun, EINOM ‘going’, or (M)EINOM building a figura etymologica with MITAT of line 1), one could analyze the direct object in the first part of line 3, (DVENOSMEDFECEDENMANOMEINOMDVENOI), namely ENMANOMEINOM, as *en mānom eīnom* ‘in this good (vessel)’. The reasoning is as follows: Old Latin evinces a stem *eī-* in three of five case forms of the demonstrative pronoun which was to become classical *is*, namely (all singular and masculine) nom. EIS (CIL 583), gen. *eīius* (Pl. Am. 108), dat. EIE (CIL 583,9),

28. N over Cardinals (NOC): yes, cf. at least the first of the two NPs in the Garigliano bowl: ESOMKOMMEOISSOKIOISTRIFOIAUDEOMDUO[M] (~ *sum cum meīs sociīs tribus...*) ‘I am with my three companions’. Following Vine (1998), also the second NP shows a numeral (the NP being inflected for genitive case, ~ *audiōrum duōrum*), so the full translation would be, as per Vine l.c., ‘I am, together with my three companions, [the bowl/possession/votive offering] of the two Audii’. Two NPs, and both evince an order N–Num. Another exaple from the *Sen. Bacch.*, l. 20(–21): (NEVE INTER IBEI) VIREI PLOVS DVOBVS MVLIERIBVS PLOVS TRIBVS (ARFVISE VELENT) ‘(nor shall there be present among them) more than two men and three women’. Or consider Plaut. *Merc.* 542: *hunc (me) diem unum* ‘this one day’.
29. N over Ordinals (NOO): yes, cf. *bello Punici primo* (Cic. *De Div.* 2.20) ‘in the first Punic war’ or *bello Punico tertio* (Cic. *Verr.* 2.4.73) ‘in the third Punic war’.
30. N over adjectives (NOA): yes, cf. e.g., Plaut. *Merc.* 537: *Di immortales* ‘immortal gods’.
31. N over GenO (NGO): yes, cf. e.g., *magister equitum*, *pater familiās* etc.

---

compared with abl. EOD (CIL 366) and acc. *im* (Paul. *Fest.* 41,7), *em* (id. 67,23), see Meiser (2010, p. 160). Paradigmatic leveling is a well-known phenomenon, and an accusative sg. masculine \**ej-om* is not a priori impossible (note that a form IUM is attested in CIL 401). However, besides that one may point out the nearly perfect match (apart from the vocalism of the stem) of a hypothetical old Latin accusative (\*)*ej-n-om* with Gothic *ina* ‘he.ACC.SG’, taking both as archaisms. The fact that Vedic has an acc. sg. m. *imām* of this pronominal stem is no fatal counterargument, given the fact that Vedic likewise was subject to an assimilation between two nasals in the genitive singular of the 1st person pronoun, leading to Ved. *māma*. This becomes clear from the Iranian (e.g., Young Avestan *mana*, Old Persian *manā* or many modern Iranian languages continuing this form in a new function) and Slavic (Old Church Slavonic *mene*) cognates. Proto-Indo-European acc. sg. m. \**inom* is thus a possibility to be reckoned with, and taking into account paradigmatic leveling, Old Latin *ej-n-om* as well. If correct, it may be preserved, of all things, in an NP with N–Dem order. Bearing in mind the danger of circular reasoning in my line of argumentation, *en mānom ejnom* might turn out to be both a remarkable archaism concerning internal syntactic behavior and the stem of the demonstrative.

- 32. N over external arguments (NOE): yes, examples under parameters NDE, NOC, NOO, NOA and NGO.
- 33. Verbless relative clauses (VRC): yes. Cf. Plaut. Stichus 649: *salvete, Athenae, qui nutrices Graeciae* 'Greetings, Athens, who (is) the nourisher of Greece'.
- 34. Facultative linker (FLI): no. There is no linker construction, only (rare) verbless relative clauses).
- 35. Generalized linker (GLI): no.



## Chapter 4

# Parametric Syntactic Reconstruction

The idea of parametric syntactic reconstruction will now be tested using the data that were collected in Chapter 3. Since parameters are structurally similar to phonemes – discrete, finite in number, synchronically arbitrarily set in every language and common or rather applicable to all human languages –, they may be an integral part of smallest meaningful syntactic units, syntaxemes. Syntaxemes should be able to yield minimal pairs and they should be comparable and reconstructible in an abstract (and discrete) manner, of the type: Did a particular proto-language have the syntaxeme ( $\sim$  phoneme) X?

A comparison of parameters is reminiscent of a comparison of phonological features, and this implies that parameters can be used for comparative purposes on an abstract, quasi phonological level (as if one were to compare the phonological inventories of languages, only with a syntactic module: NPs). Note that the parameters to be compared in section 4.2 do not constitute cognates in the sense of syntactic equivalents to words or morphemes, because real cognates consist of several meaningfully arranged basic units. Just as words or morphemes consist of several phonemes arranged in a particular way or order, syntactic cognates consists of phrases (or clauses) with specific configurations

of syntaxemes. What we will first reconstruct is, in other words, a syntactic ‘phoneme’ inventory. The first step in the process of reconstruction is always abstract in nature, aiming at an elaboration of the exact settings and values of the relevant basic discrete units. The concrete reconstructions can only follow in a second step, after the phonemes or syntaxemes as well as general or context-induced developments (sound changes  $\sim$  syntactic changes) have been worked out.

## 4.1 A note on Indo-Iranian ethnogenesis and language contact

In this section I will describe the paralinguistic (archaeological, prehistoric, cultural) background of the language(s) to be reconstructed. The reason is simple: The ethnogenesis of the Indo-Iranians implies that language contact with non-IE speech communities may have been a factor with potentially distorting effects on the P(I)Ir. parametric values. I will now explain what is meant by this.

The oldest Ir. languages we know of were spoken somewhere in southern Central Asia and in the Punjab, far away from the IE Urheimat in the Pontic-Caspian steppe.<sup>1</sup> If archaeology and archaeogenetics are to be trusted, the ultimate origin of those steppe herders that were to migrate to southern Central Asia, the Iranian plateau and the Punjab during the course of the second millennium BCE must be two-fold. The later Aryans<sup>2</sup> were partly descendants of a

<sup>1</sup>For the location of the PIE Urheimat in the Pontic-Caspian region, see most recently Olander (2019), Anthony (2013) or Anthony and Ringe (2015) with references, but also Mallory and Mair (2000) or Day (2001) for the anthropological facts and Anthony (2007) as well as Mallory (1989) for the archaeological background. Recent genetic data and their implications are discussed and reviewed by Anthony (2019).

<sup>2</sup>Since many old Indo-Iranian tribes and communities had a similar autonym for their language, religion or culture as a whole (cf. Vedic *ārya*- ‘Aryan’, *aryá*- ‘lord’, Young Avestan *airiia*- ‘Aryan’, Old Persian *ariya*-, or the name of the *Alans* < *\*ariāna*- etc.), it is reasonable to conclude that this is an inheritance dating back to common Indo-Iranian times and that already the Proto-Indo-Iranians called themselves or their language/culture/religion ‘Aryan’. See EWAia (Mayrhofer 1992–2001), I, p. 174–175 s.v. *ārya*- for conventional etymologies. The most recent

successor culture to the famous Yamnaya culture in the Pontic steppe, viz. the Catacomb culture, and partly of the more northerly Corded Ware complex; here in form of the so-called Abashevo culture. Livestock herders associated with both the Poltavka and Abashevo cultures began to explore and colonize new territories along and beyond the southern part of the Ural mountain range around 2500 BCE, and eventually they mingled into a new cultural complex stretching around the southern Urals.<sup>3</sup> Ecological stress (increasing droughts) and competition for dwindling resources may have been the ultimate impetus for this eastward migration. The result was a new, highly distinctive culture: the so-called Sintashta culture of the Trans-Urals which flourished between 2100–1800 BCE.<sup>4</sup>

Whereas some Poltavka and Abashevo groups crossed the Urals and formed the Sintashta culture, a large part of the population remained in their original eastern European homelands between the upper Don, the middle Volga and the southern tip of the Ural range. Abashevo groups persisted in the northern forest-steppe until 1900/1800 BCE, while to the south the Poltavka-Potapovka cultures (besides other, smaller cultural groups like Filatovka etc.) flourished in the centuries around 2000 BCE, thus forming a kind of cultural triad – with Sintashta in the north-east, late Abashevo in the northern and north-western forest-steppe zone and late Poltavka/Potapovka in the Pontic-Caspian steppe.

All these cultures probably belonged to the same higher-level Aryan cultural complex, to judge by the fact that the material culture in the eastern European/southern Uralic area was very similar in the centuries around 2000 BCE. New military tactics and technologies were developed in this area, among them the hallmark feature of the Sintashta culture: the chariot.<sup>5</sup> Horses were held in high

---

publication on the topic (to the best of my knowledge) is Benedetti (2023) which lacks, however, scientific rigor in its etymological speculations.

<sup>3</sup>For Aryan origins in the southern Urals, see Anthony (2007, pp. 371–411), Kuz'mina (2007) or Kroonen et al. (2018).

<sup>4</sup>Recent genetic studies have confirmed the archaeological findings, as both Yamnaya/Catacomb- and Corded Ware-related ancestry has been detected in the populations associated with Arkaim, Sintashta and related sites (Allentoft et al. 2015; Barros Damgaard et al. 2018; Narasimhan et al. 2019).

<sup>5</sup>See Anthony (1995, 2007), Kuz'mina (2007), and Kuznetsov (2006) and, specifically for the



esteem by the Sintashta people, and burials including chariots and richly decorated horses perfectly match the ideology and worldview of the later Vedic and Old Iranian tribes. It is hard to deny the conclusion that Sintashta was a (or rather *the*) prototypical Aryan culture, and that its formation depended on both internal (frequent war) and external (trade, climate) factors.<sup>6</sup>

After the invention of chariots they were quickly adopted by all groups (Sintashta, late Abashevo and late Poltavka/Potapovka), as were other technological innovations related to the increase in copper and bronze production and processing (e.g., javelins). Both to the east and to the west of the Urals the majority of the people had remained semi-nomads who valued horses and cattle, had little appreciation for farming, quickly adopted new technologies, were always ready for war and seem to have had very similar or almost identical religious beliefs. Funeral rituals and animal sacrifices in both Sintashta and Potapovka sites (both horses and, interestingly, dogs)<sup>7</sup> perfectly match later Vedic descriptions and are almost indistinguishable from ones in Sintashta-related sites.

Between 1900–1750 BCE (Anthony 2007, p. 397), an eastward outgrowth of the Sintashta culture, the so-called Petrovka culture, began to reach out on the steppe in what is today northern Kazakhstan. Contact between Sintashta-related groups, which exported copper and bronze in a large scale to the south, and the towns of southern Central Asia may have begun via Sarazm in the Zeravshan valley. By then, Sarazm existed for more than 1500 years (at least between 3500–

Ilr. chariot-related terminology, Malandra (1991).

<sup>6</sup>See Anthony (2007), especially p. 393: “To succeed in war, chiefs needed wealth to fund alliance-building ceremonies before the conflict and to reward allies afterward. Similarly, during the climatic crisis of the late MBA [Middle Bronze Age] in the steppes, competing steppe chiefs searching for new sources of prestige valuables probably discovered the merchants of Sarazm in the Zeravshan valley, the northernmost outpost of Central Asian civilization. Although the connection with Central Asia began as an extension of old competitions between tribal chiefs, it created a relationship that fundamentally altered warfare, metal production, and ritual competition among the steppe cultures”. See also Kroonen et al. (2018), Kuz'mina (2007), and Parpola (2002) for the emergence of the Sintashta culture and its identification with early Indo-Iranians.

<sup>7</sup>Concerning the latter, see Brown and Anthony (2019) for the archaeological remnants of such dog sacrifices and Oberlies (2012, 22–31, also p. 340 with fn. 32) for the Vedic counterpart, i.e., the *śvaghnín*- ‘winner of a (ritual) dice game’ (lit. ‘dog-killer’).

2000 BCE, if not founded even earlier) and it must have been a major hub for all inner-Asian trade of metals (copper, tin, gold) and other valuables (turquoise, lapis lazuli etc.) in the early and middle Bronze Ages (see Anthony 2007, pp. 418–421).

Sarazm itself was a northern colony of “early Iranian farmers” of the Kopet Dagħ piedmont (Anau, Namazga-Tepe, Altyn-Tepe etc.) and as such was populated by humans with a different genetic profile and heritage than the Yamnaya- and Corded Ware-descended Sintashta-Petrovka peoples (Narasimhan et al. 2019).<sup>8</sup> Around 2000 BCE, Sarazm, as well as the villages of the Zaman Baba culture along the lower Zerafshan, were abandoned. Soon after that, Petrovka immigrants from the north established colonies in the Zerafshan valley and exploited the copper and tin mines of the area. This was the first time Aryans directly moved into the territory of the early Central Asian towns.

Further to the south, peoples of the same Near Eastern cultural complex that had established the northern outpost of Sarazm 1500 years earlier – i.e., “early Iranian farmers” of the Kopet Dagħ/southern Caspian shore – colonized the desert oases of the Murgab river delta and settled along the (upper) Oxus River/Amu Darya (Masson 1986). This eastward colonization was a constant process which started in the second half of the 3rd millennium BCE and intensified between 2100–1800 BCE.

Settlements in the regions later to be known as Margiana and Bactria were

---

<sup>8</sup>Besides the Yamnaya- and Corded-Ware-derived Sintashta-Petrovka steppe pastoralists in the north and the early Iranian farmer-descended colonists in Sarazm there was also a third component present in Central Asia, namely the autochthonous hunter-gatherers and fishermen of the Kelteminar culture. Whatever language they may have spoken, it appears to have left no traces in Indo-Iranian. Vice versa, things could be different, viz. if the Kelteminar people spoke, perhaps, languages ancestral to the modern Yeniseian languages. Yeniseian languages are nowadays nearly extinct, but appear to have been spoken over a comparatively large territory in southern Siberia in prehistoric times, to judge by preserved hydronyms (see Dul’zon 1959a,b; Maloletko 1992; Vajda 2019 or Janhunen 2020, p. 167). According to Witzel (2003), very old loanwords from either Proto-Indo-Iranian or an archaic Indo-Aryan or Iranian language are reflected in Kott *art’a* ‘true, veritable’ (← PIIr. \**rta*- > cf. Ved. *rta*-, YAv. \**ərta*- etc. [Witzel probably meant Kott *art<sup>h</sup>a*]) or Kott *čāk* ‘force’, *čaga* ‘strong’ (← PIIr. \**čak*- > Ved. *śak*- ‘to be able’, *śakti*- ‘force’ etc. [actually, Kott *čāk*, *čâx*, *čâg*]). Certainly, this needs a more detailed investigation.

either newly founded or saw a great increase in size and wealth at that time, eventually leading to real towns which dominated their surrounding areas. Between 2000–1800 BCE, the so-called Oxus culture or Bactria-Margiana Archaeological Complex (BMAC) passed through its mature phase,<sup>9</sup> and was on the cusp of becoming a genuine civilization.<sup>10</sup> Its inhabitants produced bronze artifacts of finest craftsmanship and interacted both with the Indus Valley Civilization (IVC) in the south-east as well as with cultures of the Iranian plateau to the south, the Caucasus in the far west and the steppe pastoralists to the north (see Luneau’s chapter in Wemhoff and Koelbl 2018).

Contact with the latter may have proved fatal, since an “explanatory phase of contact and trade between the northern steppes and the southern urban civilizations about 2100–2000 BCE” (Anthony 2007, p. 431) was supplanted by a second phase, visible in the “establishment of the Petrovka metal-working colony at Tugai, probably around 1900 BCE” which signaled “the actual migration of chariot-driving tribes from the north into Central Asia” (Anthony 2007, p. 433). Thus, “[a]fter 1900 BCE a contact zone developed in the Zerafshan valley and extended southward to include the central citadels in the BMAC towns. In the Zeravshan, migrants from the northern steppes mixed with late Kelteminar and BMAC-derived populations” (op. cit., p. 435).

The influx from the steppes permanently increased, and after 1800 BCE, many

---

<sup>9</sup>Sometimes the BMAC as a whole is dated to 2400–1600 BCE, but it is perhaps better to distinguish between (a) BMAC precursors in the 4th and 3rd millennia BCE, (b) a premature or colonization phase (2500/2400–2100/2000 BCE), (c) a classical phase (2100/2000–1800 BCE) and (d) a period of decay (1800–1600 BCE), so that we should define the BMAC in the narrow sense as flourishing between 2100–1600 BCE. For the BMAC in general, see Hiebert (1994, 2002a) or the different articles in Wemhoff and Koelbl (2018); cf. further Francfort (2001), Anthony (2007, pp. 421–427) and, of course, the various writings of the discoverer of this civilization, Victor I. Sarianidi.

<sup>10</sup>It is not clear whether the classical criterion for the evaluation of a culture as a ‘true’ or ‘genuine’ civilization/Hochkultur – i.e., literacy – was met in the case of the BMAC. The verdict depends on one’s interpretation of the so-called Anau seal and similar tokens (the Niyā seal in Xinjiang etc.). Do these seals represent some kind of logo-syllabic writing system used in the towns of Bronze Age Central Asia (see Colarusso 2002; Hiebert 2002b)? Or are they merely pictographic?

BMAC centres/towns sharply decreased in size. At this point, the Sintashta-Petrovka cultures to the east of the Urals had been superseded by several local cultures (the most prominent being the Alakul and Federovo ones, besides others) which together are classified as belonging to the larger Andronovo horizon. Andronovo as well as its sister, the Srubnaya horizon west of the Urals (which emerged out of late Abashevo and Potapovka, see Anthony 2007, pp. 435–441), flourished between 1800–1200 BCE. Both Andronovo and Srubnaya bear testimony to the great mobility and dynamics of early Indo-Iranian tribes in Central Asia and on the Pontic-Caspian steppe.

Andronovo groups continued the copper and tin mining endeavors of their Petrovka antecedents in the Zerafshan valley, but they also expanded further south towards Margiana and Bactria proper. The contact, which originally seems to have started as a trade relation for the benefit of both sides, apparently was not always a peaceful coexistence. If the Ṛgvedic hymns have a kernel of historical truth in them, at least those Andronovans who were the ancestors of the later Indo-Aryans must have had a rather devastating effect on anybody coming in between them and their path. Around 1600 BCE, the last BMAC towns were abandoned and semi-nomadic Andronovans – as well as some Andronovans who had begun to start a life as settled farmers along the lower course of the Amu Darya – controlled southern Central Asia.

Indo-Iranians must have encountered speakers of *some* other language(s) on their great trek towards their new homes and consequently language contact must have been a factor to reckon with. And in fact, we know of several dozen non-allocatable loanwords found in both Ir. and IA. and which therefore must date back to common PIIr. times (the so-called ‘Indo-Iranian substrate’).<sup>11</sup> Previous research on language contact situations has shown that a limited number of loanwords, as observable in PIIr., points to a substratal or adstratal contact situation, and that such situations typically result in a simultaneous structural

---

<sup>11</sup>See Lubotsky (2001) as well as Schmitt (2018, pp. 1948–1950), Witzel (1999, pp. 54–56) or Witzel (2003).

convergence on the phonological and syntactical level of the languages involved (see Thomason and Kaufman 1988, Curnow 2001, Thomason 2001, Aikhenvald 2006 or Hajnal 2018, as well as Table 4.1).

	<b>Borrowing scenario</b>	<b>Substratum or adstratum scenario</b>
<i>Lexicon</i>	numerous borrowings	few borrowings or isolated loanwords
<i>Phonology</i>	no interferences	numerous interferences
<i>Morphology</i>	possibly morphological borrowings	very few morphological borrowings (if at all)
<i>Syntax</i>	no interferences	numerous interferences

Table 4.1: Language contact scenarios following Hajnal (2018).

Since the IIr. loanwords show a characteristic terminology pointing to a sedentary, agricultural lifestyle in oases (words for ‘brick’, ‘permanent building’, ‘pit, well’, ‘canal’, ‘donkey’, ‘tortoise’, ‘bread’ etc.), it has been surmised that the language reflected in the substrate was spoken by the population of the Bactria-Margiana Archaeological Complex.<sup>12</sup> The unanimous scholarly opinion is that this substratal language, enigmatic and elusive as it is, did in any case *not* belong to the IE family.

Both cultural groups had roughly five centuries to interact between the first trade contacts and the ultimate decline of the BMAC; and indeed the cultural impact of the BMAC on early Indo-Iranians appears to have been significant. Several distinctive traits of early IIr. groups – not only material aspects (such as the type of pottery found in later abodes of IIr. peoples),<sup>13</sup> but also religious, cul-

<sup>12</sup>See Lubotsky (2001, 306 ff.), Witzel (2003, p. 30) or Pinault (2006).

<sup>13</sup>Mallory and D. Q. Adams (1997, p. 73): “It has become increasingly clear that if one wishes to argue for Indo-Iranian migrations from the steppe lands south into the historical seats of the Iranians and Indo-Aryans that these steppe cultures were transformed as they passed through a membrane of Central Asian urbanism. The fact that typical steppe wares are found on BMAC sites and that intrusive BMAC material is subsequently found further to the south in Iran, Afghanistan and Pakistan, may suggest then the subsequent movement of Indo-Iranian-speakers after they had adopted the culture of the BMAC.”

tic and mythological idiosyncrasies like the \**sauma*- and fire cults<sup>14</sup> – lack good parallels elsewhere in the ancient IE *Kulturkreis*. It has been argued that many, if not all, of these characteristics owe their existence to a period of intensive contact between early Indo-Iranians and the BMAC and that the Indo-Iranians were largely the receiving part in this interaction. The whole process has been compared by J. P. Mallory with a projectile (= the Aryans) losing its shell of outer metal while passing through the area (= body) of the BMAC (the so-called “Kulturkugel” model).

The interesting question, now, is very simple: If substratal contact relations result in syntactic interferences, and given the fact that Indo-Iranians have been in a substratal contact situation with speakers of an almost certainly non-IE culture, will the parametric values of Proto-Indo-Iranian and Proto-Iranian differ from those of their western Indo-European relatives? In order to answer this question, we will now establish the parametric values first of Proto-Iranian and then of Proto-Indo-Iranian.

---

<sup>14</sup>See Főrizs (2016) for a case study of how concrete physical BMAC elements may be reflected in Vedic hymns.

## 4.2 Step 1: The syntaxeme inventory

### 4.2.1 Proto-(Indo-)Iranian

Before the reconstructions can be carried out, take notice of the fact that the 9 premodern Iranian languages represent at least 5 different chronological layers. As described in Chapter 3, Old Avestan was spoken in the (mid) 2nd millennium BCE, Young Avestan sometime between 1000 and 500 BCE, Old Persian between roughly 500–330 BCE and the Middle Iranian languages (notwithstanding the fact that a few inscriptions date to the 2nd and 1st centuries BCE) were living languages in the 1st millennium CE, with Chorasmian being a special case due to its even later attestation. See Figure 4.1 for a visualization of the different chronological layers.

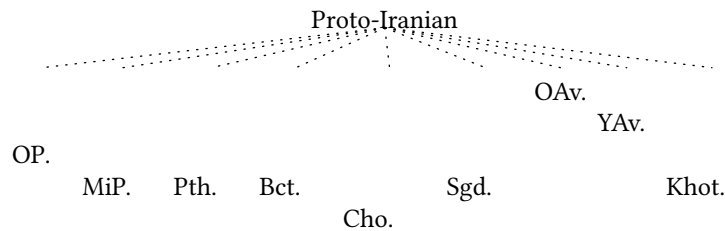


Figure 4.1: Premodern Iranian languages: relative chronological layers.

The chronological gap between the last common ancestor of all Iranian languages, Proto-Iranian, and Old Avestan, Young Avestan and Old Persian is not as great as the time span between PIr. and the Middle Iranian languages. The default assumption would be that the three Old Iranian languages – and among them OAv. more so than YAv., and YAv. more so than OPers. – are more conservative than the younger Middle Iranian languages, because the Old Iranian languages simply did not have as much time to develop and diverge from the inherited patterns or archetype than their younger cousins. However, this does not imply that OAv. should be analyzed as a living fossil or unchanged continuation of Proto-Iranian. Changes can have occurred also here, and it will be necessary

to weigh the evidence for each parameter anew to make plausible conjectures.

#### 4.2.1.1 Unambiguous reconstructions

Several parametric values appear to be straightforwardly reconstructible for the Proto-Iranian level, as all premodern Iranian languages evince the same settings. In such cases, there should be no doubt about the Proto-Iranian values. The parameters identical in all languages are the following: FGP +, FGN +, FSN +, FNN +, DPQ –, DIN –, AST +, FFS +, ADR +, AER 0, ARR +, GAD –, GFS –, PAP 0, PCL +, GSP –, GCN –, NOA +, NGO +, NOE + and GLI –. A graphical visualization of the parametric reconstruction for each of these aforementioned parameters, implying a flat hierarchy due to the same values, might look as follows:



Figure 4.2: Parametric syntactic reconstruction: flat hierarchy (X = some parameter).

We can assume with a very high degree of confidence that Proto-Iranian had grammaticalized person (both in the verbal and pronominal domains) and grammaticalized number (parameters FGP and FGN), and that the number marking was morphologically expressed on the noun (FSN and FNN). All premodern Iranian languages evince this strategy. Even though this is by no means trivial as there are several non-IE languages who diverge from Proto-Iranian already with these parameters and on this level (e.g., Mandarin or Japanese), it is of course no new information to anyone familiar with reconstructions of the ‘traditional’ syntax, i.e., morphosyntax of Proto-(Indo-)Iranian.

Likewise, Proto-Iranian almost certainly had no count/mass distinctions by means of differential case marking (as e.g., modern Finnish has). Iranians have been in contact with speakers of Uralic languages for millennia, but the long and



early contact apparently did not result in syntactic convergence (at least with regard to this particular parameter and the behavior it describes). Regardless of the exact location of the *Pre-Proto-Uralic Urheimat*,<sup>15</sup> and whether or not one believes in the autochthonism of Uralic to north-eastern Europe, it is commonly assumed that in the centuries around 2000 BCE (at the latest) a belt of Uralic languages was spoken both to the west and to the east of the northern Urals, i.e. in the boreal zone immediately to the north of the forest-steppe. Numerous IIr. loanwords in Uralic languages are known, many of them only into smaller-scale daughter branches (e.g., borrowed only into Finnic, Mordvin, Mansi etc.).<sup>16</sup> This implies an intensive contact between Uralic varieties and different IIr. contact languages,<sup>17</sup> and that both speech communities/dialect continua must have been

<sup>15</sup>The opinions strongly differ, cf. Häkkinen (2012) for an overview. Häkkinen (2012) has argued, primarily based on early contacts between Uralic and Yukaghir, for a *Pre-Proto-Uralic homeland* in the Altai-Sayan area before 3000 BCE (similarly Janhunen 2009, who assumes a *Pre-Proto-Uralic Urheimat* somewhere in southern-central Siberia, between Lake Baikal and the Sayan mountains). The current distribution of Uralic languages does not contradict this scenario. A rather rapid spread of Uralic languages over large territories of north-western Asia and north-eastern Europe around 2000 BCE can also be explained by the Seima-Turbino phenomenon, whose origins can be traced back to the Altai (Marchenko et al. 2017) and which was roughly contemporaneous with Sintashta and late Abashevo (see Anthony 2007, pp. 443–448).

<sup>16</sup>According to Holopainen (2019, p. 343), this makes “it likely that Uralic had split into several branches by the time of these contacts. Also the fact that many of the Proto-Indo-Iranian loanwords either show a restricted distribution (such as West Uralic *\*wačara*, *\*worači*) or irregular correspondences (*\*asVra*, *\*śasra*, *\*śīta*) can point to the conclusion that Proto-Uralic was fragmenting by the time when contacts with Proto-Indo-Iranian took place.”

<sup>17</sup>For prehistoric contacts between IIr. and early Uralic varieties see the different articles in Carpelan et al. (2001) and the recent dissertation by Sampsa Holopainen (2019). Three to four different layers of IIr. loanwords in Uralic languages are discernible on the basis of characteristic sound changes which took place between PIE and Proto-Iranian: 1. a *Pre-Proto-Indo-Iranian layer*, as exemplified by Proto-West-Uralic *\*kekrä* ← Pre-IIr. *\*kekro-* (cf. PIE *\*k<sup>w</sup>ek<sup>w</sup>lo-* > Skt. *cakrá-* ‘wheel, cycle’), see Holopainen (2019, pp. 118–119); 2. a *Proto-Indo-Iranian layer*, e.g., Proto-Uralic *\*šišta* ‘beeswax’ ← PIA./PIIr. *\*čišta-* (Skt. *madhu-śišta-* ‘beeswax’, lit. ‘what is left over’), Holopainen (2019, pp. 249–250); 3. perhaps an *Indo-Aryan layer* (although this may just be the same as the above-mentioned *PIIr. layer*, given that a pre-Indo-Aryan dialect of PIIr. was the source language for the borrowings into Uralic), e.g., PU. *\*anti* or *\*onta* ‘grass’ ← PIIr./PIA. *\*(H)ánd<sup>h</sup>as-* > Skt. *ándhas-* ‘herb’, cf. Holopainen (2019, pp. 55–58) for this word; and 4. finally loanwords from different *Iranian* languages (Proto-Iranian, Scythian, Sarmatian/Alanic), such as Proto-Ugric/Proto-Uralic *\*sīr(a)ńa* ‘gold, metal’ and Proto-Uralic *\*serńä* ‘gold’ ← PIr. *\*dzaranīa-* > YAv. *zaraniia-* ‘gold’, Bactr. *ζapo*, Oss *zærin/zærijnæ* etc. (cf. in contrast Skt. *hīraṇya-* ‘gold’),

close enough to interact. Interestingly, Uralic languages appear to have been nearly exclusively the receiving part in this interaction, and the apparent lack of Uralic loanwords in IIr. languages renders any structural influence of Uralic on PIIr. and PIr. highly improbable, unlike the situation with regard to the BMAC language. This is confirmed by parameter DPQ which has a positive value in Finnish and other Uralic languages but a negative in Proto-Iranian, to judge by the premodern Ir. varieties.

Similarly, theoretically possible contact with ancient Semitic languages either did not result in any interferences or (perhaps more likely) there was no contact at all due to different locations at which Proto-Iranian and early Semitic varieties were spoken (still separated by the Kopet Dagħ mountains around 1800–1600 BCE, the time span of PIr.). This lack of interference manifests itself in two parameters of which the settings differ between Semitic (+) and PIr. (–), viz. DIN and GCN. All 9 premodern Iranian languages have a negative value with regard to these parameters, signaling that PIr. nouns (a) did not have special inflectional forms of nouns depending on the occurrence of certain morphemes in D (as is the case with the Arabic nunation) and (b) that there was likewise no distinctive morphological marking on the noun when occurring with a genitive argument (Semitic construct state).

Parameters AST and FFS relate to structured adjectives (i.e., NP/DP internal ones, as contrasted with freely occurring ones) and feature spread (agreement) between head noun and dependent adjectives. Both are set positively in all 9 languages, implying that PIr. had adjectives occurring in NP/DP internal position and showing inflection for number, gender and case. Likewise, all languages had adjectives functioning as free reduced relative clauses. Again, this is no new information.

Relative clauses, assumed to be ‘base-generated’ in a pre-NP position (which is also the typical pattern for modern Indo-Aryan and Dravidian languages), apparently were systematically crossed over by the whole NP/DP, implying a sur-

---

see Holopainen (2019, pp. 232–234).

face manifestation to the right of all other nominal arguments and modifiers (NP–RelCl). This does not imply that relative clauses always had to occur to the right of the NP, but merely that PIr. had the (frequently used) possibility to do so. Correlative relative clauses, e.g., could also precede the noun, but since all 9 premodern languages have the consistent possibility to place relative clauses to the right of NPs, parameter ADR must be reconstructed with a positive value for Proto-Iranian. Languages in which relative clauses are not systematically crossed over by the NP (Hindi, Marathi etc., but also Basque) can still exhibit postnominal relatives as an effect of extraposition (which is the case in Hindi and Marathi: finite relatives are postposed, whereas participial ones are not). Parameter AER asks whether in ADR negative languages at least a subset of relative clauses can be postposed to the rest of the NP (Hindi and Marathi yes, Basque no; see the appendix to Longobardi et al. 2013, p. 15). Since PIr. had systematic postnominal relative clauses (N–RelCl), parameter AER becomes irrelevant.

Parameter GAD, defining whether or not the language had an adpositional genitive (like, e.g., modern Pashto), unambiguously must be reconstructed with a negative value. PIr. had no adpositional genitive, but instead an inflectional one. Speaking of genitives, parameter GFS, asking whether or not there is phi-feature agreement from a non-free genitive to the noun (head marking), must be reconstructed with a negative value for Proto-Iranian. Nouns did not incorporate a dedicated morpheme signaling gender or number of the genitival argument. A free genitive likewise did not agree with its head noun (GSP –). Genitives in premodern Iranian languages usually were dependent NPs which had their own inherent phi-features (gender or number), so that this value is set negatively. There was no spread of internal phi-features like gender or number from the noun to dependent genitives, unlike dependent adjectives. An illustrating example is YAv. (Y. 57.3) *ahe yasna yazatanqm* ‘by his (GenS) sacrifice of/to the gods (GenO), showing that the head noun *yasna* neither influences the first nor the second genitive in terms of gender or number, as both entail their own information.

Since Proto-Iranian certainly had no article (see the next section), let alone a suffixed one (as is found, for instance, in modern Norwegian, Swedish or Bulgarian), parameter PAP (post-affix possessives) is (or was) irrelevant. Possessives, however, certainly could occur as bound morphemes cliticizing on the noun or another stressed item, prominently on the first stressed word in a clause (Wackernagel position). Hence parameter PCL (clitic possessives) should be reconstructed with a positive value. Again this is neither new nor trivial, as many modern non-IE languages like Finnish, Hungarian, Arabic or Basque (but also modern western IE languages) have no clitic possessives and thus a different parametric value. We can conclude that ca. 3800 years ago in the steppe and (semi-)deserts of southern Central Asia, Proto-Iranians/Aryans expressed possession *inter alia* by means of clitic possessives.

Common to all 9 premodern Iranian languages was also the possibility to raise nouns at least over adjectives. The situation is different with modern Iranian languages: Pashto does not have this possibility, as here the noun always is NP-final and the language in general is rigidly head-final. There certainly was no generalized linker in Proto-Iranian as in modern Persian, though the languages differ as to whether they allowed only for verbless relative clauses or were on a grammaticalization pathway leading to linkers.

With that being said, the unambiguous evidence has been presented and reconstructions have been given for the respective parameters. The remaining parametric values force us to weigh the arguments anew in each case for the settings of the proto-language, as the individual languages differ.

#### 4.2.1.2 Ambiguous evidence

We begin (or rather continue the discussion) with the N-raising parameters. Ranking higher than adjectives, the next nominal arguments in the NP-internal hierarchy are numerals, with cardinal numbers ranking higher than ordinals. The evidence is inconclusive, and often the data are not very significant, given the fact that e.g., OAv. has only one single instance of a Noun Phrase with a

cardinal number. Similarly, Old Persian has only a few attestations of cardinals, and not a single one of a structured (i.e., attributive) ordinal. Consider Figures 4.3 and 4.4 for the individual settings of parameter NOO (noun raising over ordinal numerals) and NOC (noun raising over cardinal numerals):

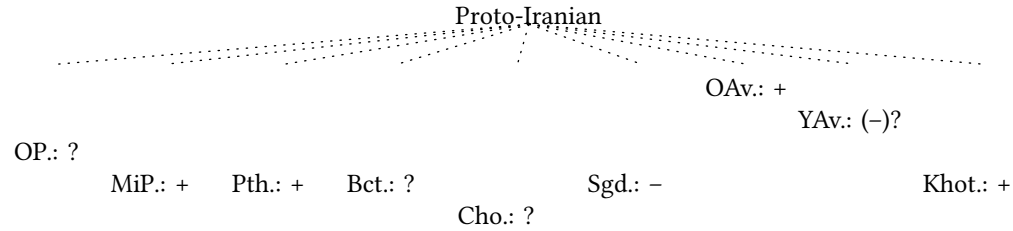


Figure 4.3: Parameter NOO: Noun raising over ordinals.

With regard to parameter NOO, Old Persian, Bactrian and Chorasmian yielded no useful data; with regard to NOC, Old Avestan must be left out of the discussion.

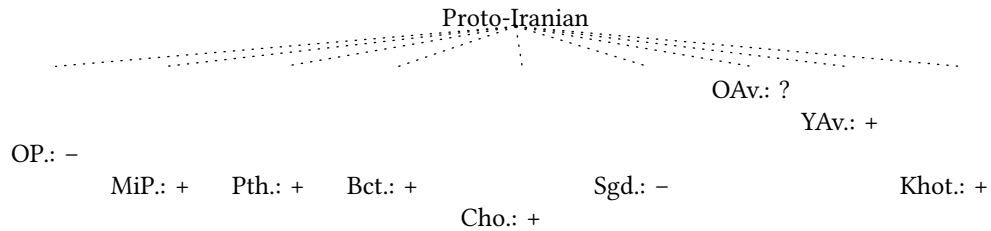


Figure 4.4: Parameter NOC: Noun raising over cardinals.

An educated guess would be that Proto-Iranian had positive values for each of the two parameters. There are two reasons for this assumption: 1.) If there is indeed a ‘base-generated’ order of nominal constituents – irrespective of how exactly it may be hardwired in our brains and mental processes –, human beings will tend to fall back to this order, if their particular language does not already evince it. If it is a natural order, and if all humans share some basic cognitive strategies, then deviations from it, no matter how they originally came into be-

ing,<sup>18</sup> will provoke an unconscious drift back towards a more natural order. Over time, this will result in a tendency to come back to the unmarked pattern, and in the case of N-raising parameters, this means NPs being rigidly head-final without the ability to be raised over arguments. This situation is met in Pashto. Apparent deviations from this diachronic trend (many modern western Iranian languages, with Farsi as archetype) can then be explained (and in fact have been explained) as being the result of language contact in a linguistic macro-area with languages that mutually reinforce their ‘unnatural’ patterns and prevent the fallback to (in this case nominal) head finality.<sup>19</sup>

2.) Considering the evidence, it is conspicuous that the Middle Iranian *Randsprachen* Middle Persian, Parthian and Khotanese show positive values for both parameters. One possible explanation could be language contact with languages that had the ability to raise N over numerals (implying that the Iranian languages did not have this possibility beforehand), another an internal development (at least with regard to the Middle Western Iranian languages, with the linker construction entailing a right-branching behavior), but the third possibility is a trivial *Randsprachenarchaismus*. In that case, the centre of the then much larger Iranian speaking area – the regions along the Oxus/Amu Darya – had innovated

---

<sup>18</sup>This relates to the question of language change in general and opens up a completely new discussion. Suffice to say here that in my view traditional accounts of language change (analogy, economy principles etc.) cannot fully explain the fact that typologically highly unusual patterns (click consonants, ejectives; basic OVS order in transitive clauses etc.) can emerge at all in the first place. I always took it for granted (similar to Sarah Thomason’s ideas on language change) that language change predominantly arises out of the strong human desire to create a group identity. Humans everywhere on this planet create groups or collectives and distinguish themselves from humans belonging to other communities – in fact that is the very reason why decoration, adornment, tattoos, body modifications, clothes and (hair) fashion exist at all. ‘We’ want to differ from ‘them’. Language signals group identity not on a visible, but on an audible level. This becomes all the more tangible (or audible), if a particular group speaks in such a highly distinct way that it differs from all other groups in the greater vicinity. It is obvious that this is the only explanation that can account for the fact that apparent deviations from ‘base-generated’ or natural orders exist at all, despite strong tendencies to fall back to the latter.

<sup>19</sup>See Stilo (2005, 2012). But to be fair, one could likewise argue (and Stilo has argued) that the head finality in Pashto and other modern eastern Iranian languages is also due to language contact with languages from the Indian area.

parametric values, whereas the periphery had preserved the original and inherited situation. The Old Iranian languages are either inconclusive as to their settings (YAv. NOO –, NOC +) or have no reliable data (OAv. in terms of NOC with just one example and OPers. with a handful).

Considering the fact that N-raising over nouns is, at least in the theoretical model or rather concept employed here, marked and considering further that it is most prominently visible in the Middle Iranian periphery, it seems likely that this is an inheritance from Proto-Iranian. And since the argument of theoretical markedness applies to N-raising in general, it seems plausible that Proto-Iranian probably also had the possibility to raise N over demonstratives. Consider Figure 4.5:

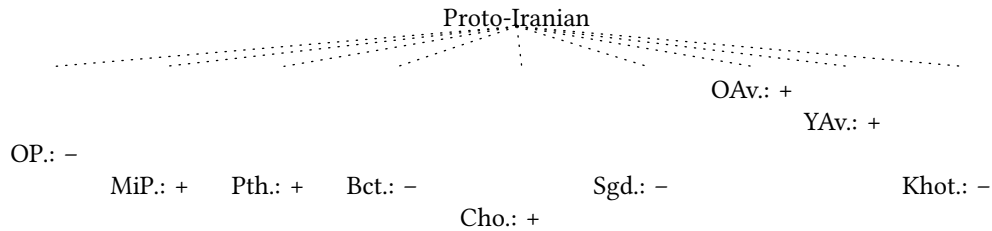


Figure 4.5: Parameter NDE: Noun raising over demonstratives.

In this case, the fact that both Old and Young Avestan could raise nouns over demonstratives, as is true also of three of the six Middle Iranian languages (again Middle Persian and Parthian, but this time also Chorasmian; interestingly not Khotanese), combines historical weight with the majority argument. Five languages make use of this possibility, and among them the two oldest ones. Apparently, Proto-Iranian could raise nouns over *any* nominal argument, showing a theoretically marked pattern. This could explain, why, for instance, Khotanese did not have this possibility anymore. Khotanese was in the midst of a long-term drift back to a head-final order either completely without N-raising, or with a reduced version (only before numerals). The same is true for Bactrian and Sogdian. Old Persian shows an unexpected behavior given the Middle Persian situation;

it may be the case that the corpus simply was too small and too stereotypical to show the real parametric settings of Old Persian.

The next parameter to be discussed and reconstructed is FGG, relating to grammaticalized gender. With the exception of Middle Persian and Parthian, and bearing in mind the Bactrian situation – Bactrian still had vestiges of gender distinctions in the early Kushan era, but quickly lost them afterwards –, both the majority argument and the fact that all three Old Iranian languages have gender contrasts clearly speaks in favor of a reconstruction with a positive value for FGG in Proto-Iranian. This is likewise not unexpected. The same is true for CPS (plural spread to cardinals). All languages have positive values, although Sogdian and Chorasmian *sensu stricto* do not have feature spread of plurality to elements after numbers, but rather feature spread of ‘numerality’. In any case numerals cause non-singular marking on nouns occurring together with them. Nouns then are inflected for plural or dual (or a dual decoupled from its original limitations), and the same situation – morphological marking of NPs with numerals for non-singularity – can be reconstructed for Proto-Iranian.

Much more interesting is parameter DGP (grammaticalized partial definiteness). Six of the nine languages have a negative value, among them the three Old Iranian representatives:

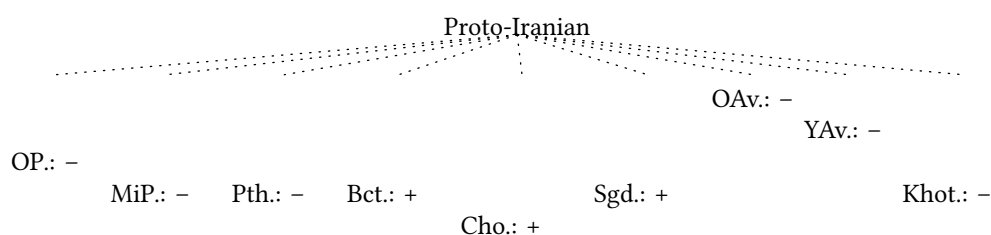


Figure 4.6: Parameter DGP: Grammaticalized partial definiteness.

A negative value for Proto-Iranian can be reconstructed without difficulty. As stated in the individual sections in Chapter 3, a negative value of this parameter renders 14 parameters (DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI,



PDC, PAP, PHS, TDC and DOA) irrelevant or rather neutralizes them. All of these hierarchically lower parameters depend on obligatory definiteness marking in the language (either anaphorical, DGP, or general, DGR). We can assume with a high degree of confidence that there was no obligatory anaphoric marking of nouns mentioned previously in the discourse in Proto-Iranian. Facultative marking by means of demonstratives, 3rd person pronouns or (verbless) relative clauses was almost certainly possible in Proto-Iranian, but it was not mandatory. 6 of the 9 premodern daughter languages have neither specialized articles nor definiteness-affixes on nouns (or nominal arguments).

However, another aspect arouses interest. There is a cluster of languages showing systematic definiteness marking by means of articles, viz. the three central Asian languages Sogdian, Chorasmian and Bactrian. All three languages have one thing in common: they were spoken in areas in which culturally and genetically different people had lived before Aryan tribes moved there during the early 2nd millennium BCE. Recall from section 4.1 that the first contact of Aryans (Sintashta-Petrovka) with the towns of Central Asia occurred via Sarazm in the Zeravshan valley. The area along the Zeravshan valley was later known as Sogdiana, and Sogdian was spoken here. Chorasmian was spoken along the lower reaches of the Amu Darya, and here, too, ‘early Iranian farmers’ (a term from the current archaeogenetic literature, see Narasimhan et al. 2019) had cultivated the land before originally semi-nomadic Aryans with eastern European ancestry settled there to become farmers themselves. Bactrian, finally, was spoken in one of the two core regions of the BMAC, Bactria.

The fact that these three languages have etymologically different article systems based on either the old relative pronoun or demonstratives (Bactr. and Chor. < Plr. *\*iǎ-/iā-*, Bactr. and Sogd. *\*imǎ-/imā-*, only Sogd. *\*hau*, *\*aṣa-*) speaks in favor of an individual development instead of an inherited pattern. In principle, there are only two plausible possibilities: 1.) One of these three varieties innovated articles as a result of an internal development, and the other two (in whichever constellation) imitated this; or 2.) the original stimulus was an

external, non-Iranian one. In that case, it may be explained as a contact-induced phenomenon – syntactic interference in other words, typical for a substratal or adstratal situation. There are two possible contact languages:

1. The article systems are the result of an imitation of (classical or Koine) Greek patterns, being the indirect consequence of Alexander's campaign into Central Asia and the establishment of Greek/Macedonian colonies (Ἀλεξάνδρεια Ἐσχάτη in the Ferghana valley; Ἀλεξάνδρεια Ἀραχωσίας, better known as Kandahar in modern-day Afghanistan). This could, perhaps, explain the Bactrian and Sogdian calques, but not the Chorasmian one. What is more, why did Middle Persian and Parthian not develop article systems, if Greek was the model? The Seleucid Empire with its hellenization was the perfect venue for an amalgamation of Greek language and culture with native Iranian elements – but nothing happened, at least in terms of articles.
2. The fact that those, and only those, Middle Iranian languages which were spoken in former territories of 'early Iranian farmers' – Bactria, Zerafshan valley, Chorasm –, combined with the fact that PIIr. adopted vocabulary from a sedentary culture, might indicate that the 'early Iranian farmers' not only were the source for terminology related to agriculture/irrigation in oases, but that they also spoke a language (or rather a set of genetically related languages) which had articles or overt determiners. This, then, could have become relevant only in the long term, influencing the intruding (I)Ir. peoples and languages and causing them to develop articles.

There is one problem with this idea. As stated before, both Old and Young Avestan were probably spoken in southern Central Asia as well (though not necessarily, and also not likely in the same region). Likewise, Vedic (or rather Indo-Aryan in general) inherited the vocabulary from the common IIr. substrate situation, but neither the two Avestan varieties, nor Vedic, nor the three other Middle Iranian languages show any evidence of syntactic interference with the non-IE

Oxus–Zerafshan languages (at least in terms of articles). Perhaps the contact was not long enough to yield results in terms of syntactic borrowing, or the language contact was rather a borrowing than a real substratum scenario (see Table 4.1).

In the latter case, if the commonly accepted term ‘Indo-Iranian substrate’ is just a case of infelicitous wording, the contact between Aryans and ‘early Iranian farmers’ in Central Asia between 2000–1600 BCE may instead have been a borrowing scenario, in which case we might be satisfied with the fact that there was no syntactic convergence. If early Indo-Iranians were just passing through, or making their living as semi-nomads outside of the urban centres without a real substratum situation, the lack of an emergence of articles would not be unexpected. The situation with regard to those Aryan colonists who settled down in the formerly non-IE speaking areas of the Zerafshan valley and along the Oxus/Amu Darya<sup>20</sup> would then have been different in nature. By becoming sedentary themselves, they would have had much more contact with the native populations than their semi-nomadic cousins (e.g., the Sakas). In the long term, perhaps over the course of a full millennium or more (assuming that the native non-IE languages survived that long), this real substrate situation would then have led to the emergence of article systems in the precursors of later Sogdian, Chorasmian and Bactrian. Indirectly, and if the emergence of articles was not a completely internally motivated development, this would or could tell us something about the syntactic profile of the native non-IE Central Asian language(s).

In any case Proto-Iranian had a negative value of DGP and hence no articles or overt determiners. Parameter CGB, defining whether or not singular count nouns are inherently unbounded or not, shows an interesting pattern. The three Old Iranian languages have negative values for this parameter, as do Khotanese and Sogdian. Middle Persian, Parthian, Bactrian and Chorasmian, on the other

---

<sup>20</sup> Add to this the ancient satrapy/region of Margiana or rather the Murghab river and the oasis of Merv. Unfortunately, we know next to nothing about the native language(s) spoken there, in particular nothing about Old or Middle Iranian varieties. In the event that substantial textual remains (inscriptions) are, perhaps, one day uncovered as a result of archaeological excavations, it is likely that they will show an Iranian language with articles, notwithstanding Middle Persian or Parthian overlays/superstrates.

hand, cluster together as having developed unbounded readings of singular count nouns and accompanying use of an element based on the numeral ‘one’ to signal boundedness (see Fig. 4.7). In other words: The Middle Iranian languages show a syntactic isogloss separating East and West, but ignoring phonology-based divisions (Bactrian and Chorasmian are traditionally classified as eastern Iranian languages).

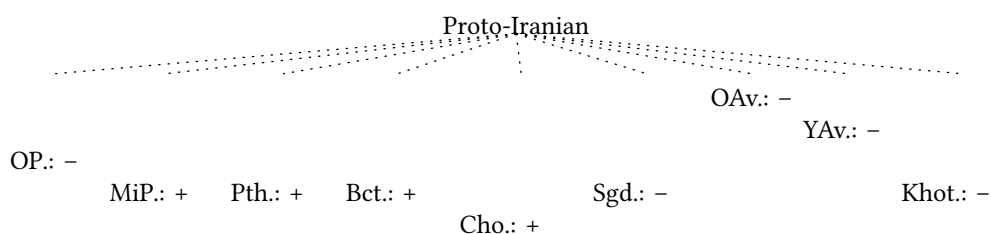


Figure 4.7: Parameter CGB: Grammaticalized boundedness.

The Proto-Iranian situation probably was reminiscent of the former group of languages; singular count nouns were not inherently unbounded and did not require the numeral ‘one’ (or a derivative of it) to overtly signal boundedness, as this was inherently expressed through the singular inflectional endings. Instead, an unbounded reading of singular count nouns could probably be marked via the collective ending PIr. *\*-H*, morphologically homophonous with the neuter plural and being a relic inherited from PIE (< *\*-h<sub>2</sub>*), used together with singular verb agreement. Since parameter CGB is thus reconstructed with a negative value, the dependent parameter CCN must have been neutralized in Proto-Iranian.

The next parameters are those pertaining to genitives. Consider Figure 4.8.

Proto-Iranian had a uniform genitive, which was inflectional, not bound to any structural positions, could occur inside the NP or separated from it, and which could be freely iterated inside but also outside of the structured NP core. Being an extended version of the free genitive found in many modern languages, both GFR and GUN must be reconstructed with positive values for Proto-Iranian. Note that Old Persian, Khotanese and Chorasmian have inflectional genitives

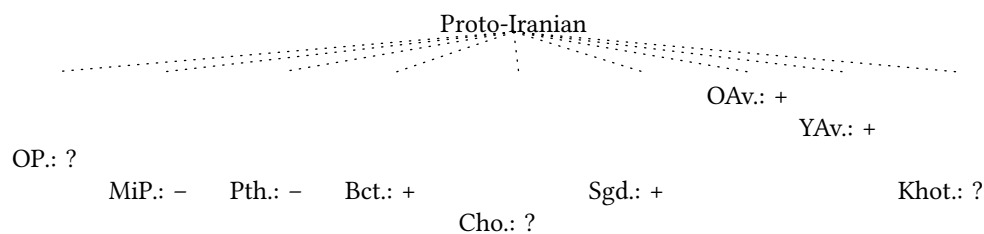


Figure 4.8: Parameter GUN: Uniform genitive.

not bound to any structural positions, but that I was unable to find recursive genitives to the right of a raised head noun (N-FreeGen-FreeGen; probably an artifact of the limited corpora and the nature of the texts). Strictly speaking, they must be left out of the discussion (marked by question marks in Figure 4.8). The possibility to raise the whole NP over free genitives is also reconstructible, with the latter then occurring to the right of the NP (see Figure 4.9):

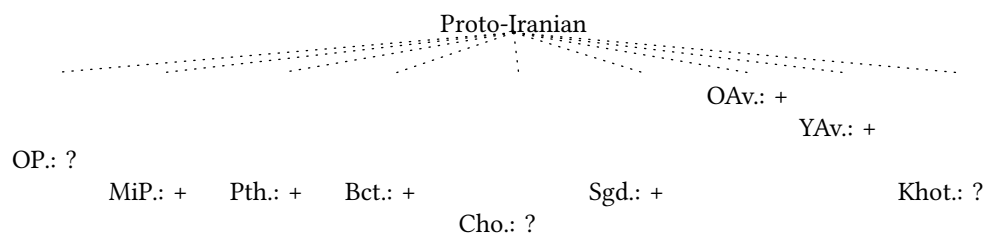


Figure 4.9: Parameter GPR: DP/NP over free genitive.

Since Proto-Iranian had a uniform genitive, parameter GFO, asking for a specific postadjectival genitive (*genitivus obiectivus*) becomes irrelevant, as its function is already covered by the uniform genitive.

The first parameter not straightforwardly reconstructible is APO. APO defines whether or not the language had adjectival possessives (such as Vedic *má-maka-* ‘my’, *asmāka-* ‘our’, *yusmāka-* ‘your’ or Old Avestan *ma-* ‘my’, *θβa-* ‘your’ etc. Consider Figure 4.10.

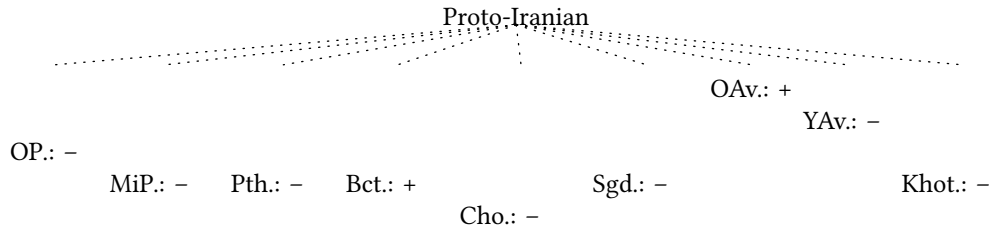


Figure 4.10: Parameter APO: Adjectival possessives.

Bearing in mind the majority rule, a negative value should be reconstructed for APO in Proto-Iranian, implying that the language did not have the possibility to express possession by means of adjectival possessives. However, most Indo-Europeanists would probably refrain from proposing this (given the fact that Latin, Greek, Gothic or Armenian all have adjectival possessives), and in this case the inner-Iranian lack of adjectival possessives (apart from Old Avestan and Bactrian) may be compensated for by the Indo-Aryan evidence. Both Vedic and Old Avestan have adjectival possessives (see Schmidt 1978, pp. 85–86), and this concordance could be brought forward in favor of a common inheritance from Proto-Indo-Iranian times. I would therefore opt for a positive value of APO in Proto-Iranian, but note that this reconstruction is not based on Iranian data, but on the Ilr. basis of comparison (thus being an anticipation of the next Section).

There were probably no adjectival genitives in Proto-Iranian, as becomes clear from Figure 4.11. Only Bactrian has adjectival genitives, making it plausible that this is an innovation (note that AGE will become irrelevant, if the language does not allow for adjectival possessives).

Five parameters remain to be discussed (recall that many others were neutralized due to a negative setting of higher ranking parameters like DGP). Parameters TPL and TSL ask for ‘strong locality’. The observation behind these parameters relates to the fact that deictic demonstratives sometimes behave differently from discourse-oriented (anaphoric/cataphoric or topical) ones. Parameter TPL defines whether or not at least some demonstratives (at least deictic ones) are at-

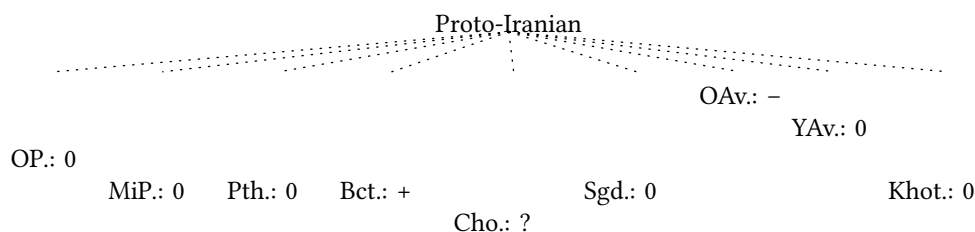


Figure 4.11: Parameter AGE: Adjectival genitives.

tracted toward the D-area (whereas others are not and surface lower in the NP), and TSL further specifies this by stating whether or not the system is uniform and *all* demonstratives are moved to the D-area. The Iranian situation is varied, consider Figure 4.12.

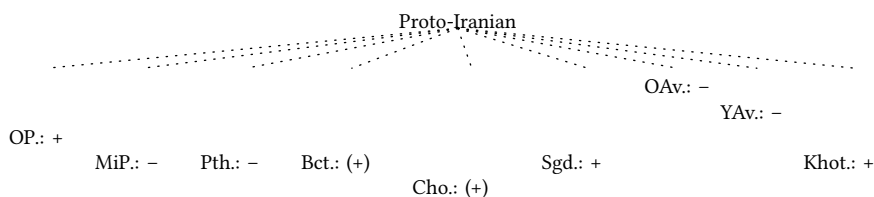


Figure 4.12: Parameter TPL: Strong partial locality.

Note that a negative value of TPL implies neutralization of TSL, as is the case with Old and Young Avestan, Middle Persian and Parthian. The bracketed values of Bactrian and Chorasmian are due to an implicational relationship with parameter CGR. The latter must not be positive; otherwise TPL becomes irrelevant or rather predictable (and CGR is positive in Bactrian and Chorasmian).

A priori, it is possible that TPL had a positive value in Proto-Iranian. But given the fact that nouns probably could be raised over demonstratives in Proto-Iranian, and since in OAv. and YAv. demonstratives do not systematically surface in the D-area, a reconstruction of this parameter with a negative value in PIr., with a subsequent neutralization of TSL, seems more likely.

The last parameters are those asking for a linker. Consider Figures 4.13 and 4.14.

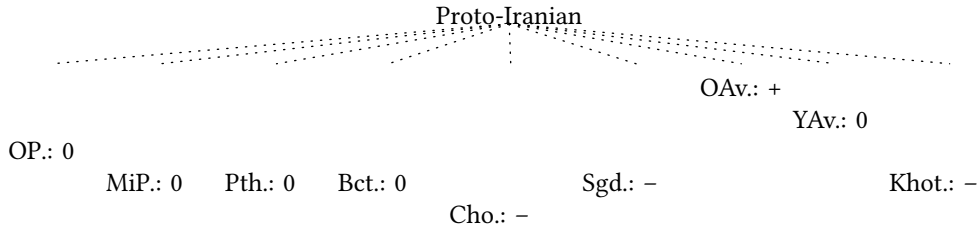


Figure 4.13: Parameter VRC: Verbless relative clauses.

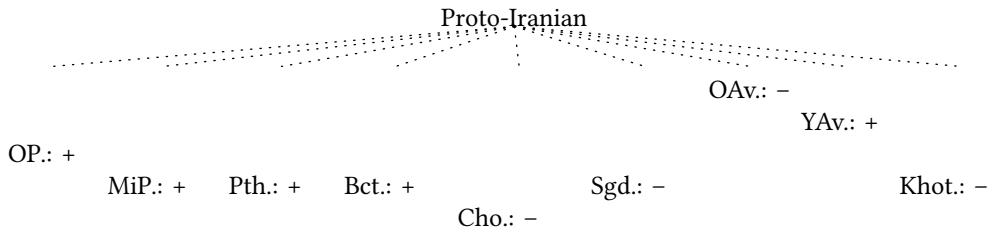


Figure 4.14: Parameter FLI: Facultative linker.

As discussed in Chapter 3, OAv. had no linker, but verbless relative clauses, whereas YAv., OPers., MiPers., Parthian and Bactrian had facultative linkers. The majority argument would suggest to reconstruct a positive value for parameter FLI for Proto-Iranian, signaling that nouns could be linked to dependent arguments by means of a facultative linker. However, the grammaticalization process leading to the modern Persian or Kurdish linker constructions is well studied (see Seiler 1960, Haider and Zwanziger 1984, Yakubovich 2020 or Viti 2013, 94ff.), so that we can safely assume that Proto-Iranian did neither have a generalized nor a facultative linker, but instead could specify nouns by means of verbless relative clauses with or without *tractio relatiui* (VRC +, FLI –, GLI –).



#### 4.2.1.3 The (Indo-)Iranian parametric features

Previously, it has been stated that the OAv. situation need not be identical with the PIr. one. However, apart from one parameter (NOC, and recall that OAv. simply yielded not enough data), the values are in fact the same. It is an open question whether this is just valid for NP parameters (and here only for the small set of this study) or whether it reflects a general conservativity of Old Avestan. In the verbal domain, there could be differences between the parametric settings which can be reconstructed on the basis of the 8 other premodern Iranian languages and the values of Old Avestan. At this point in time, we do not know; the reconstruction of verbal parameters is a desideratum. With regard to the NP parameters covered by my study, OAv. is indeed a language not differing from its (perhaps) 200–300 years older predecessor Proto-Iranian. The only parameters posing some uncertainty are APO, defining whether or not Proto-Iranian had adjectival possessives and FLI, stating whether or not facultative linkers were allowed. But here the comparison with Vedic sheds light on the probable situation in the common parent language, as both Ved. and OAv. have adjectival possessives and only verbless relative clauses. Speaking of Vedic and the comparison with Indo-Aryan, the question arises whether the reconstruction of Proto-Iranian can be taken one step further and be projected back to Proto-Indo-Iranian by means of a comparison with Vedic.

As indicated before, a rough terminus ante quem for the period of time in which PIr. (or Proto-Aryan) may have been spoken is provided by (1) the existence of an Ilr. superstrate in the Near East between ca. 1450–1350 BCE,<sup>21</sup> and by (2) the oldest preserved IA. and Ir. texts, i.e., the Ṛgvedic hymns and the Old Avestan literature. Both are usually dated to the middle or to the latter half of the 2nd millennium BCE, so that a conservative inference would be that PIr. must have disintegrated around 1600–1500 BCE at the latest, and most likely earlier.

---

<sup>21</sup>Which shows typical Indo-Aryan features (Vedic deities, lexical features etc.); best known is the Indo-Iranian/Indo-Aryan superstrate of the otherwise Hurrian-speaking state of Mitanni (cf. Mayrhofer 1966, 1982 or Witzel 2001, pp. 53–55), but the Kassites may also have had an Indo-Aryan element among their ruling class (and the greater Hyksos, one wonders, as well?).

Personally, I would reckon with ca. 2100–1800 BCE for Proto-Indo-Iranian (Sintashta) and ca. 1800–1600 BCE for Proto-Iranian, but there is certainly room for discussion about these dates or time frames.

Are there any differences between Proto-Iranian and Proto-Indo-Iranian in terms of syntax? Phonologically, it is clear that all Iranian languages share several innovations not found in Indo-Aryan (but partly in the Nuristani languages) that must go back to a common proto-language – e.g., loss of the aspiration (or more likely, breathy voice) of the old mediae aspiratae or the development of stops to fricatives in certain environments (Mayrhofer 1989). But is this also valid for syntactic features?

The parametric settings of the last common ancestor of all Iranian languages, Proto-Iranian, have now been reconstructed. Vedic values must be compared with the Proto-Iranian ones in order to ascertain which values the last common ancestor of all Iranian, Indo-Aryan and Nuristani languages had.<sup>22</sup> See Table 4.2 for a tabular overview of the parametric values of all 9 premodern Iranian languages, those reconstructed for Proto-Iranian, and the Vedic ones.

---

<sup>22</sup>I will equate the Vedic situation with the Proto-Indo-Aryan one. I previously noted that one should not carelessly equate the Old Avestan situation with the Proto-Iranian one, but since it turned out that Proto-Iranian indeed had the same parametric NP values as Old Avestan, it may be acceptable to take Vedic as the only reference point from the Indo-Aryan branch. I am aware of the methodological weak point here, as one would normally expect the reconstruction of Proto-Indo-Iranian to entail a comparison of several daughter languages from all three branches (or at least of the oldest Indo-Aryan and Iranian ones, as their members are attested in premodern times, in contrast with Nuristani).

Title/Description	Precondition(s)	Abbrev.	OAv.	YAv.	OP.	Kho.	Sgd.	Cho.	Bet.	Pth.	MIP.	Pir.	Ved.
1 Grammm. Person		FGP	+	+	+	+	+	+	+	+	+	+	+
2 Grammm. Number	FGP must be +	FGN	+	+	+	+	+	+	+	+	+	+	+
3 Grammm. Gender	FGN +	FGG	+	+	+	+	+	+	(+)	-	-	+	+
4 Feature spread to N	FGN +	FSN	+	+	+	+	+	+	+	+	+	+	+
5 number on N (bounded nouns)	FSN +	FNN	+	+	+	+	+	+	+	+	+	+	+
6 Grammm. partial definiteness	DGP +	DGP	-	-	-	-	-	-	-	-	-	-	-
7 Grammm. definiteness	DGP +	DGR	0	0	0	+	+	+	-	0	0	0	0
8 Strong Person	FGP +, DGR +, NDE -	NSD	0	0	0	0	+	(+)	+	0	0	0	0
9 Free null partitive Q	FSN +	DPQ	-	-	-	-	-	-	-	-	-	-	-
10 Grammm. distal art.	FSN - or FNN - or DGR +	DDA	0	0	0	0	+	-	(+)	0	0	0	0
11 Def.-checking N	DGR +	DCN	0	0	0	0	-	-	(-)	0	0	0	0
12 Def. spread to N	DCN +, NSD -	DSN	0	0	0	0	0	0	-	0	0	0	0
13 Def. on relatives	DGR +	DGR	0	0	0	0	-	-	(-)	0	0	0	0
14 D-controlled infl. on N	FSN +	DIN	-	-	-	-	-	-	-	-	-	-	-
15 Plural spread from Cardinals	FSN +	CPS	+	+	+	+	+	(+)	+	+	+	+	+
16 Grammm. boundedness		CGB	-	-	-	-	-	-	+	+	+	+	-
17 Strong article	DGR +, FNN +, CGB -	CGR	0	0	0	0	-	-	(+)	0	0	0	0
18 Bounded-checking N	CGB +	CCN	0	0	0	0	0	0	-	+	+	+	+
19 null-N-licensing article	FSN - or FNN - or DCN -, NOD + or NSD +	DNN	0	0	0	0	?	(+)	+	0	0	0	0
20 Structured Adjective Phrases		AST	+	+	+	+	+	+	+	+	+	+	+
21 Feature spread to structured APs	FSN +, AST +	FPS	+	+	+	+	+	+	+	+	+	+	+
22 Feature spread to predicative APs	FGN +	FSP	+	+	+	+	+	+	-	+	+	+	+
23 D-controlled infl. on adjectives	NSD -, FFS +	ADI	0	0	0	0	-	-	-	0	0	0	0
24 DP over relatives		ADR	+	+	+	+	+	+	+	+	+	+	+
25 RelCl. extrap.	ADR -	AER	0	0	0	0	0	0	0	0	0	0	0
26 Free reduced RelCl	ADT +	ARR	+	+	+	+	+	+	+	+	+	+	+
27 Adpositional Genitive		GAD	-	-	-	-	-	-	-	-	-	-	-
28 Free Gen.		GUN	+	?	+	+	+	?	+	+	+	+	+
29 Uniform Gen.	GFR +	GUN	+	+	(+)	+	+	+	+	+	+	+	+
30 DP over free Gen.	GFR +, ADR +	GFR	+	?	?	?	?	?	+	+	+	+	+
31 GenO	GUN must not be +	GFO	0	0	(0)	0	0	(0)	0	+	+	+	+
32 Gen-feature spread to N		GFS	-	-	-	-	-	-	-	-	-	-	-
33 D-checking possessives	DGR +, NSD + or CGR not +	PDC	0	0	0	0	-	-	(+)	0	0	0	0
34 Adjectival poss.		APQ	+	-	-	-	-	-	-	-	-	(+)	+
35 Post-affix poss.	DCN +	PAP	0	0	0	0	0	0	0	0	0	0	0
36 Clitic poss.		PCL	+	+	+	+	+	+	+	+	+	+	+
37 N-feature spread to pron. poss.	FFS + or AST -, PAP + or PCL +	PHS	0	0	0	0	+	+	+	0	0	0	0
38 N-feature spread to free Gen.	FFS +, GFR +, PHS must not be -	GSP	-	-	-	-	-	(-)	-	-	-	-	-
39 Adjectival Gen.	APQ +	AGE	-	0	0	0	0	0	+	0	0	-	-
40 Poss.-checking N	GFS -	GCN	-	-	-	-	-	-	-	-	-	-	-
41 Strong partial locality	FSN - or FNN +, CGR must not be +	TPL	-	-	+	+	+	(+)	(+)	-	-	-	-
42 Strong locality	TPL must not be -	TSL	0	0	0	+	+	(-)	(+)	0	0	0	0
43 D-checking demonstratives	FSN - or DGR +, TPL must not be -	TDC	0	0	0	0	-	-	-	0	0	0	0
44 N over Demonstratives	FGP +	NDE	+	+	-	-	-	+	-	+	+	+	+
45 N over Cardinals		NOC	?	?	?	?	-	?	?	+	+	+	+
46 N over Ordinals	NOC -	NOO	+	(-)?	+	+	-	?	?	+	+	?	?
47 N over adjectives	NOO - or NGS -, Npp -	NOA	+	+	+	+	+	+	+	+	+	+	+
48 N over GenO	GFO must not be -, NOA - or AST -	NGO	+	+	+	+	+	+	+	+	+	+	+
49 N over external arguments	NGO - or (GFO -, NOA - or AST -)	NOE	+	+	+	+	+	+	+	+	+	+	+
50 Definiteness on APs	DGP +, postnominal APs	DOA	0	0	0	0	-	-	-	0	0	0	0
51 Verbless relative clauses	FLI -	VRC	+	0	0	0	-	-	-	0	0	0	+
52 Facultative linker	GLI -	FLI	-	+	+	-	-	-	+	+	+	+	-
53 Generalized linker		GLI	-	-	-	-	-	-	-	-	-	-	-

Table 4.2: Parametric Grid: NP parameters in Vedic, Old and Middle Iranian languages and Proto-Iranian.

As can be seen, there is no difference (apart from isolated uncertainties) between Vedic, Old Avestan and Proto-Iranian. Apparently, the configurational syntax of NP parameters of Proto-Iranian was identical to that of its predecessor Proto-Indo-Iranian. The data do not allow for any other conclusion. After the Vedic tribes had crossed the Hindu Kush, several phonological innovations spread among the remaining, mutually intelligible Indo-Iranian dialects that gave them a distinctly Iranian appearance, thus changing their phonological systems from Proto-Indo-Iranian to Proto-Iranian. Yet with regard to nominal configurational syntax, these Proto-Iranian dialects did not innovate, as otherwise different parametric values should be reconstructible on the base of the daughter languages. Whether this is also valid for verbal parameters or more fine-grained NP parameters than covered in this study is an open question at this point. For the time being, we can conclude that the syntactic equivalent to a phoneme inventory of Proto-Iranian appears to be the same as that of Proto-Indo-Iranian, at least in terms of nominal parameters.

In essence, the nominal syntax of Proto-(Indo-)Iranian was characterized by (case, number, gender) agreement between nouns and dependent adjectives,<sup>23</sup> whereas there was no phi-feature spread of the noun to genitives. Instead, the latter could have own dependent adjectives characterizing them which were then inflected for gender or number according to the inherent properties of their genitival head nouns. Free adjectives, i.e., those occurring as participial adjuncts outside the NP core, could function as reduced relative clauses<sup>24</sup>. Nominal ar-

<sup>23</sup>Alfieri (2011) (to name but one study in this vein) is mistaken by assuming that Vedic and Avestan did not have real adjectives or a noun/adjective distinction. According to Alfieri, the oldest Ir. languages preserve an old, allegedly IE situation with only two (instead of three, as in, e.g., Latin) lexical classes, viz. verbal roots and primary nouns, without or only a few genuine adjectives. It is true that adjectives have the same inflectional endings as nouns in Ved. and Av. so that there is no inflectional difference between nouns and adjectives, but on the other hand, gender marking (by means of phi-feature agreement) only affects adjectives (regardless of lexicalized gender differences as in Ved. *nár-* ‘man, hero, warrior’ : *nârī-* ‘woman’ or Av. *nar-* ‘man’ : *nāirī-* ‘woman, wife’). This is a clear signal that adjectives were a separate class of words. Furthermore, one wonders how to reconcile the fact that the Caland system is securely reconstructible for PIE with an alleged lack of a distinctive class of adjectives in PIE or its daughters.

<sup>24</sup>For the phenomenon as such, see Baker and Vinokurova (2009), Cinque (2010), Kayne (1994),

guments could optionally be linked to a head noun by means of verbless (and right-branching)<sup>25</sup> relative clauses with complete (case, number, gender) agreement between head noun and relative pronoun. Contrary to widespread assumptions,<sup>26</sup> the possibility to raise nouns over *any* argument, including numerals and demonstratives, is securely reconstructible for Proto-(Indo-)Iranian. Since it results in a marked right-branching surface word order (in the theoretical model employed in this study), deviating from unconscious or ‘base-generated’ NP patterns,<sup>27</sup> it is understandable why several daughter languages lost this ability, if not backed up by a linguistic environment reinforcing this marked pattern and preventing the loss (as was the case in the Near East with Akkadian, Elamite or Urartian as contact languages with nominal right-branching).<sup>28</sup>

Sleeman (2011) or several articles in Le Feuvre et al. (2017).

<sup>25</sup>See Seiler (1960, pp. 194–201), who states that there are only few exceptions to the right-branching of verbless relative clauses with case attraction in Avestan (Yt. 13.1, V. 13.2). For verbless relative clauses in other IE languages, see Probert (2015, pp. 407–414).

<sup>26</sup>For instance, Yakubovich (2020, p. 95) has “the impression that the dependent-head word order was the most common one in the Iranian-speaking area and is to be reconstructed as prototypical for Proto-Iranian nominal syntax”.

<sup>27</sup>Concerning this order, see Longobardi (2001b), Cinque (2005) or Dryer (2018). For the sake of completeness: It would go beyond the scope of this study, but I simplified the theoretical discussion to some extent. Modern generative analyses assume that there is raising of the whole NP (or VP) besides or apart from simple N (or V) raising (see e.g., Alexiadou et al. 2007). This is called ‘snowballing’ or pied-piping and means that each consecutive step in the derivation/raising moves a larger constituent to the left instead of moving only N (or V). *In nuce*, N takes ever larger phrasal parts with it when raised over higher elements; this can then result in surface orders like N-Adj-Num-Dem, i.e., seemingly mirrored NP structures contradicting the concept of a left-branching, head final ‘base-generated’ order. See Cinque (2005, pp. 321–325) for an account of such superficially aberrant or deviating orders. I abandoned the respective parameter (NPP, nominal pied-piping) of Longobardi et al. (2013), because I did not want this study to be overly complex. *Qui nimium probat, nihil probat*.

<sup>28</sup>Nonetheless, note that also here a diachronic tendency to fall back to a left-branching typology shines through. A telling example is the old (and probably borrowed) title ‘king of kings’. Old Persian still has N-Gen, *xšāyaθiya xšāyaθiyānām* (e.g., DBa 1–2), whereas Middle Persian and Modern Persian have a preposed genitive (MiPers. *šāhān šāh*, Farsi *šāhan-šāh*). Left-branching apparent archaisms in otherwise predominantly (or nearly exclusively) right-branching languages need not be archaisms. Quite the opposite: the generative NP model predicts or rather suggests that there is an unconscious impetus to fall back to a left-branching strategy in right-branching languages, whereas left-branching languages do not have such a tendency. In a strict sense, NPs with N at the right edge tell us nothing about the prehistory of the particular lan-

That the parametric value signaling this strategy (NDE +) has been preserved in several daughter languages is an archaism and probably tells us that the parent language much more often and much more regularly used this strategy. If it was a rare strategy already in P(I)Ir. times, we might expect that the unconscious bias to fall back to a much more restricted N-raising typology would have resulted in daughter languages without at least the highest two or three N-raising possibilities (N over demonstratives and N over cardinal/ordinal numerals) by the time of their attestation/preservation. Both Old Avestan and Vedic had the possibility to raise nouns over any argument, implying that N-raising over hierarchically higher ranking elements was not rare in Proto-(Indo-)Iranian. We do not know what spoken Proto-Indo-Iranian sounded like, but I would predict that the surface order of NPs was almost evenly distributed between right- and left-branching NPs (perhaps even with a slight inclination to the right-branching strategy). Also the fact that verbless relative clauses that were to develop into the *ezāfe* construction consistently are right-branching in Old Avestan, Young Avestan, Old Persian and Vedic, with N surfacing to the left of the elements introduced by the relative pronoun, tells us something about the language – Proto-(Indo-)Iranian – from which this strategy was inherited.<sup>29</sup> And by the way – syntax is not about frequency, but about syntactic options irrespective of how often they are chosen by speakers.

---

guage, whereas NPs with a raised N in otherwise rigidly left-branching languages tell us a lot about previous stages of this language (e.g., Latin, verbal basic (= most frequent) transitive order SOV, but NPs like *tribunus plebis*, *magister equitum*, *orbis terrarum*, *praefectus urbis* etc.).

<sup>29</sup>Alternatively, one could also consider the possibility that the right-branching (proto-)linker construction was a calque of a pattern found in some non-IE contact language. Yakubovich (2020) regards Elamite as a possible model for the Persian linker – which does not convince me, on both morphological and syntactical grounds –, whereas Windfuhr (2009a, p. 28) reckons with Elamite and Urartian as possible models. The fact that Old Avestan, Young Avestan, Old Persian and Vedic have verbless relative clauses clearly speaks in favor of a strategy inherited at least from Proto-Indo-Iranian (Proto-Indo-European is another issue, considering the *attractio relativi* of the classical languages Greek and Latin), so that a loan or syntactic inference with non-IE languages is not necessary to account for the grammaticalization pathway observable in Western Iranian – an internal development of the *ezāfe* is the explanation with the least additional assumptions, in accordance with Occam's Razor.

### 4.2.2 Proto-Indo-Iranian = Proto-Indo-European?

To continue this line of thought, it would make good sense to assume that already Proto-Indo-Iranian was in the midst of a long-term drift away from this marked strategy, and that a predecessor of this language must also have allowed for a nominal right-branching typology. The ancestral language of Proto-Indo-Iranian, however, is (late) Proto-Indo-European, and based exclusively on the Indo-Iranian situation one might surmise that (late) PIE was a right branching language in terms of nominal configurational syntax as well.<sup>30</sup> This could best account for the fact that Vedic and Old Avestan still had positive values for parameters NDE and NOC.

Cross-linguistically, there is a statistical tendency<sup>31</sup> that combinations of a basic transitive order SOV (which is usually reconstructed for PIE) and a right-branching nominal typology (N-Gen) are very rare – Modern Persian evinces it, but WALS lists only 26 instances of the combination SOV/N-Gen worldwide. Add to this the fact that both Vedic and Old Avestan could also raise verbs over all verbal arguments (see examples 19a and 59c), one might be tempted to propose a basic (theoretically marked, hence often given up) VP order Verb–Subject–Object or rather VSO for PIE,<sup>32</sup> with subjects usually occurring in form of suffixal endings on the verb (as rightly noted by Hale 2018). A combination of consistently right-branching VSO and N-Gen is found today in 77 languages, among which the Insular Celtic languages Irish and Gaelic, but also Berber languages or several Austronesian varieties.

<sup>30</sup>Recall the ‘imperishable fame’, Homeric κλέος ἄφθιτον (Il. 9.413) and Vedic *śrávas... akṣitam* (RV 1.9.7bc) with N-Adj. Whether this reflects a synchronically marked order in Vedic or Greek is debatable (parametrically not, but pragmatically?), but I would go as far as saying that this inherited NP is a perfect archaism reflecting the default or unmarked N-Adj order of (late) PIE.

<sup>31</sup>See WALS online (Dryer and Haspelmath 2013), combination of features 81A: Order of Subject, Object and Verb and 86A: Order of Genitive and Noun.

<sup>32</sup>Thus reviving the idea of Hirt (1937, pp. 252–258) or Miller (1975). Again, consider Vedic *áhann áhim* (RV 1.32.1 etc.) ‘he slew the serpent’, Gk. κτείνει ... ὄφιν (Pindar *Pythian* 4.249) ‘he killed the serpent’, Av. *yō janaṭ ažiṃ dahākəm* (Y. 9.8) ‘who slew Aži Dahāka’. This would imply that these cognate VPs probably do *not* reflect a marked word order, as Watkins (1995, p. 302) wants us to believe.

More data, particularly more on the verbal behavior of old IE languages, are necessary to back up this idea, but if we reverse the long-term diachronic drift back towards restricted N-raising and left-branching strategies, as observable in Iranian, it may be a plausible conjecture to reconstruct at least late PIE as a predominantly right-branching language. It is likewise not implausible to assume that after the disintegration of PIE more or less everywhere (with the exception of the Near East → Western Iranian and British Isles → Insular Celtic) the N-raising (and probably V-raising) was quickly lost within, perhaps, 1000, 1500 or 2000 years. The standard left-branching reconstruction of PIE has a clear advantage in form of its typological unmarkedness (398 instances for SOV/Gen-N in WALS), but it fails to account for the ability of languages like Old Avestan or Vedic to raise nouns and verbs over all arguments. If we take NPs with the order N-Dem or N-Numeral as archaisms and bear in mind the concept of a long-term drift back towards more natural patterns (in the generative model), a right-branching (late) PIE may be a plausible concept – with NPs and, probably, VPs or whole sentences as well. And again, note that it is completely irrelevant how frequently this option was chosen. All that matters is the fact that high raising was probably possible.

### 4.2.3 Proto-Indo-European: adding Greek and Latin

But perhaps this was a premature deduction, and the Proto-Indo-Iranian parametric profile does not reflect the Proto-Indo-European situation due to the allegedly substratal contact with central Asian non-IE languages. The fact that Proto-Indo-Iranian did not have articles or overt determiners (as indicated by Vedic and Old Avestan), whereas the three Middle Iranian central Asian languages Bactrian, Sogdian and Chorasmian did, could also be explained by means of a later, internally motivated Iranian innovation and isogloss. In that case it might be that Proto-Indo-Iranian had undergone other changes which we could mistake for archaisms. A comparison with other Indo-European branches and other ancient IE languages is the only means to decide whether or not the values



obtained for Proto-Indo-Iranian were inherited from and reflect the situation of Proto-Indo-European or whether they differ in one or more points from values of genetic relatives outside the Indo-Iranian branch.

To cut a long story short: they do not differ, apart from one single value. Of course a reconstruction of Proto-Indo-European values will be all the more reliable, if (as far as possible) all ancient Indo-European languages are included in a comparison of nominal macro- and mesoparametric values. For a start, however, only Greek and Latin may suffice. In Chapter 3, the parametric values of the oldest forms of these two branches, i.e., Archaic (Mycenaean and Epic) Greek and Old Latin, have been determined. See the Appendix for a parametric Feature Matrix (5.1) listing their values beside the Indo-Iranian ones covered by this study.

The only difference to Proto-Indo-Iranian, Vedic and Old Avestan pertains to adjectival genitives which existed in Latin and Greek, but not in the oldest Indo-Iranian representatives. Vedic, Old Avestan, Old Latin and Archaic Greek had basically the same parametric profile (Young Avestan slightly differs), at least with regard to the configurational syntax of NPs. Classical Greek will resemble Chorsmian if analyzed (DGP and DGR +), but Archaic Greek closely aligns with Latin, Vedic and Old Avestan. Concerning the possibility of adjectival genitives, it seems best to me to postpone a reconstructive decision in favor of the Indo-Iranian or Latin/Greek situation until more branches and languages have been parametrically analyzed. But apart from that, the parametric reconstruction of Proto-Indo-Iranian appears to be a good indicator as to which parametric settings Proto-Indo-European may have had. It remains to be shown in future studies whether this is also true for more fine-grained parametric analyses of both verbal and nominal nature.

Greek and Latin also tip the balance in favor of a rebuttal of a genuine substratal contact situation with the central Asian donor language(s) of the so-called ‘Indo-Iranian substrate’. The contact with the BMAC (or other Central Asian cultures/speech communities) did not affect the (nominal) configurational syn-

tax of Proto-Indo-European, and should better be regarded as a mere borrowing scenario. A substratal contact should have resulted in syntactic interferences. Either the syntactic profile of the donor language(s) was nearly identical to that of Proto-Indo-European, or there was no substratal language contact situation.

One point should be addressed in this Section as well. It is usually claimed that PIE must have been a language with no, little or only discourse-configurationality due to its rich inflectional morphology. The direct effect of this would have been frequent splitting of constituents, known as *tmesis* in the verbal domain and *hyperbaton* in the nominal one. Keydana (2018, p. 2205) devoted a half page to the issue which is worth quoting in full:

“Since determiners are not obligatory and no other empirical evidence for DPs has yet been given, we assume a simple NP structure for PIE. Hints at the internal structure of Vedic NPs can be found in Keydana (2013), who in an investigation into event nominals in the language of the Rigveda observed that no more than one argument of the event nominal can be realized in the NP (...).

Adjectives agree with nouns in the NP, the only exception being nouns in the dual, which are combined with adjectives in the plural (...), obviously due to a later development. The serialization of modifier and head noun is open to variation. Old juxtapositions like Vedic *dámpati-* (besides *pátir dán*), Avestan *dāng paiti-*, Greek δεσπότης < PIE \**déms póti-* may be taken as a hint that the modifier preceded the noun in PIE (...).

Hyperbata are the result of dislocations out of NPs. Material may be dislocated to the left into the DF-slot or to the right. While the target slots of these dislocations are easily named, the process as such is not yet understood: Neither do we know what exactly triggers right dislocation, nor are we in a position to identify factors for and possible constraints on extracting material out of NPs (...).

Maybe the process of dislocation of elements out of NPs (as is indeed a very common surface phenomenon in Vedic, Avestan, Ancient Greek or Latin) is the result of a mixed system. If we assume that Proto-Indo-European once had a

high N-raising capacity, we may surmise that this, perhaps, even went as far as allowing for N-raising out of the structured NP core. N may then have had the possibility to surface before or rather to the left of phrase-initial elements such as adpositions.<sup>33</sup> The situation then would have resembled that of other nominal elements with the possibility to surface outside the NP core, such as relative clauses, non-structured adjectives (participles functioning as adjectival relative clauses) or free genitives. What we perceive as left- or right-dislocation may actually have been the ultimate N-raising in origin.

Since nominal adjuncts like relative clauses, non-structured adjectives/participles or free genitives were not bound to any structural position inside the Noun Phrase, a fully raised noun might have had the same freedom to appear anywhere. N would have been completely decoupled from the actual phrasal core and only connected with it by means of phi-feature agreement visibly expressed via case suffixes. In fact, such hypothetical N-raising out of the structured NP core is a plausible way to explain the fact that Proto-Indo-European clearly must be reconstructed, based on the situation in its daughters, as a double-marking language instead of a head- or dependent-marking language in the sense of Nichols (1986). Double-marking would have been necessary in cases of N-raising out of the core; double-marking served to signal the affiliation of the headless NP to the straying noun, and this situation is well preserved in languages like Vedic, Old and Young Avestan or Homeric Greek (on which see Devine and Stephens 2000). In cases with N-raising leading to a surface appearance of N anywhere inside the structured NP core, double-marking was redundant and actually not motivated.

In the course of time, the theoretically predicted impulse to fall back to head-final patterns could have led to a mixed system in several daughter languages of Proto-Indo-European. Mixed system means that the basic process of extraction of an NP-internal element was retained, but since N appeared more and more

---

<sup>33</sup>Provided that Proto-Indo-European had adpositions, or at least certain case forms of nouns which were on a grammaticalization pathway leading to eventually uninflected adpositions (LOC.SG *h<sub>2</sub>ent-i* ‘on the front’ vel sim.).

often in NP-final position, the original freedom and raising did not involve the head noun anymore. The pattern remained in an abstract sense, the language still allowed for N-raising out of the structured core, but the constituents or elements involved changed: instead of nouns, nominal arguments like adjectives or numerals could be raised. This process could already have started in Proto-Indo-European times, and it was certainly not completed by the time of the creation of the Vedic hymns or the Gathas.

I would call it a mixed system, because the dislocation of phrase-internal elements was not restricted to nouns anymore, but already possible for other elements. Furthermore, this unexpected behavior of argument-raising instead of N-raising may reflect the cross-linguistic pattern of (theoretically) more natural orders with N being inside or rather at the right edge of a structured phrasal core instead of being separated from it by several intervening words. Discontinuous or distracted NPs, hyperbata, might thus be analyzed as the surface result of a mixture of an old parametrically governed pattern and a large-scale diachronic trend towards head-finality (although the inherited, ‘pure’ pattern of only N-extraction still shines through in the daughter languages).<sup>34</sup>

In the end, this would have led to phrases like Lat. *magnā cum laude* ‘with great praise’, in which only *laude* represents the structured NP core, *cum* is a phrase-external adposition and *magnā* an extraposed or dislocated element. According to the original pattern or prototype of this hyperbaton configuration, a phrase *\*laude cum magnā* might be reconstructed for earlier stages of Latin, based on the parametric approach chosen here and its intrinsic implications.

This may explain the phenomenon of dislocation without a recourse to information structure or pragmatic explanations. Even though the latter can be relevant factors, they need not be. They are additional assumptions and factors not actually necessary in a full-scale parametric approach. If we want to understand and explain syntactic phenomena, an approach relying only on syntax will be preferable to one taking into account extra-syntactic factors as well (*non sunt*

<sup>34</sup>For instance, in Old Avestan (Y. 28.1) *maniāuš ... spəntahiā* ‘of the bounteous will’.

*multiplicanda entia sine necessitate*).

To conclude this section, Proto-Indo-European probably was a language with high raising of verbal or nominal heads, illustrated by the following sentence from Plautus, *Merc.* 588: *Sumne ego homo miser, qui nusquam bene queo quiescere?* ‘Am I not a wretched mortal who can rest quietly nowhere?’. The surface word order here is highly archaic and reflects very old parametric settings, both verbally and nominally (as is well known, yes/no-questions in many modern IE languages still evince a verb-initial order).

The idea of Roberts (2007, pp. 364–367) or rather Roberts (2021, pp. 505–514) of how to carry out a parametric syntactic reconstruction has now been demonstrated. The syntactic equivalent to a phoneme inventory has been reconstructed for both Proto-Indo-Iranian and Proto-Indo-European. Can we be satisfied with the results? Is that sufficient? Or do the real reconstructions only start from here?

### 4.3 Step 2: Is there a syntactic equivalent to sound laws?

Now that the Proto-(Indo-)Iranian syntaxeme inventory for Noun Phrases has been reconstructed, regular developments of individual parameters to Iranian daughter languages can be discerned. Similar to the fact that we can reconstruct two series of palatals for Proto-Indo-Iranian (primary and secondary ones) and that we know the regular developments of these sounds to the individual daughter languages (e.g., Proto-Indo-European *\*k̑* > Proto-Indo-Iranian *\*ć* > Avestan *s*), the question arises whether there are cases in which unexpected or rather divergent outcomes of basic linguistic units like phonemes exist. For instance, the development of basic or plain PIE *\*k* in Khotanese resembles that of Avestan, i.e., it developed to a sibilant *s*, cf. Khotanese and Avestan *sata-* ‘hundred’. However, when occurring before the glide *\*u*, i.e., in the combination *\*k̑u*, it re-

sulted in Khotanese /f/ <śś> but Avestan (and in most other Iranian languages as well) /sp/, cf. Khot. *biśśa-*, Av. *vīspa-* ‘all’ (< Proto-Indo-Iranian \**uićua-* > Vedic *vīśva-*). Another famous example is Brugmann’s Law: PIE \**o* regularly became \**a* in Proto-Indo-Iranian. However, if occurring in open (and probably accented) syllables, i.e., if it was followed by just one consonant and another vowel, it developed into \**ā*. The handbook example is Proto-Indo-European \**doru* ‘wood’ > Vedic *dāru*. Obviously, a specific context prevented the normal development.

Is there anything similar observable with parameters? Can we discern instances of parametric settings that should not normally exist in the individual daughter languages, forcing us to refine the individual developments? In other words: Parametric setting or value X would be the expected one in language Y, but in a particular subset of cases it does not seem to apply – the value is set differently here. In such cases we could feel encouraged to review the context, bearing in mind that sounds (the basic phonological units) frequently develop differently than expected in certain environments and that the same may be true for parameters. NPs that are as parallel as possible in at least two different languages may be a reasonable starting point to discern potential differences. Such parallel NPs may possibly be syntactic cognates.

### 4.3.1 Syntactic Cognates

In a general sense, cognates are linguistic structures (words, morphemes, phrases, clauses) with a certain, fixed string of phonemes or a particular configuration of parametric settings (and a particular meaning) that are inherited from a common proto-language and preserved in at least two daughter languages. Their constituents, i.e., the basic units that are arranged in a particular order or constitution, may have undergone individual developments, generally or in certain contexts, which becomes evident in Vedic *soma-* and Avestan *haoma-* which both continue Proto-Indo-Iranian \**sauma-* ‘stimulating (ephedra) juice/drink’.

As stated before, this study is not the first to combine concepts from modern generative linguistics with a reconstructive agenda. Walkden (2014), the other

important study with a generative approach to syntactic reconstruction, asserted that syntactic reconstruction always entails a lexical reconstruction, taking as its premise or rather starting point a very strict interpretation of the Borer-Chomsky conjecture. To recap, this (plausible) hypothesis states that “all parameters of variation are attributable to differences in the formal features of particular items (e.g. the functional heads) in the Lexicon”. Walkden inferred from this that the reconstruction of syntactical aspects will only be possible if cognate lexical material in genetically related daughter languages is compared. However, I do not think that this is necessarily correct, and I will now explain why.

Based on the previous discussion we know the normal parametric settings for Vedic and the two Avestan languages. We would expect that nouns could be raised over their arguments in all three Old Indo-Iranian languages, irrespective of whether the NP was independent or in a genitival relation to another NP – compare Young Avestan (Yt. 13.91) *staota ašahe yaṭ mazištaheca* ‘The prayer of Aša, the greatest’ with Vedic (RV 10.66.4) *ṛtām mahád* ‘the great Truth/*Ṛta*’. The Proto-Indo-Iranian predecessor was *\*Hṛta- maj<sup>(h)</sup>Ha-* with a raised noun. Several other NPs like this are listed in Table 4.3, based on the “Konkordanz C” as assembled by R. Schmitt and B. Schlerath (Schlerath 1968, pp. 148–164).

Nr.	Iranian	Indo-Aryan	Comment
1	Y. 13.1 <i>ratūm ā yānqnm</i> 'time of the women'	RV 5.46.8 <i>ṛtúḥ jānīnām</i> 'time of the women'	N-Gen
2	Y. 16.1 <i>dātārəm vohunqm</i> (cf. Y. 65.12, Vr. 11.12, Vd. 19.17, Vd. 22.1) 'giver of goods'	RV 8.51.5 <i>dātā vāsūnām</i> (cf. also RV 2.22.3, 6.23.3, 7.20.2, 10.55.6) 'giver of goods'	N-Gen
3	Y. 50.10 <i>raocā xʷəng</i> 'the rays of the sun' but Yt. 6.1 <i>huaarə raoxšne</i> 'the Sun's light'	RV 3.22.3 <i>rocané (...)</i> <i>sūryasya</i> 'in the sun's realm of light' RV 1.14.9 <i>sūryasya rocanāt</i> 'from the sun's realm of light'	N-Gen (Gen-N)
4	Y. 44.17 <i>vāxš aešō</i> 'effectual voice' but note P. 26 <i>aešō vāxš</i> Y. 29.9 <i>anaešəm (...)</i> <i>vācəm</i>	RV <i>vācam (...)</i> <i>iṣīrām</i> 'vigorous speech'	N-Adj OAv.: predicative use
5	Yt. 10.136 ( <i>aēuuā</i> ) <i>caxra zaranaēna</i> 'on (one) golden wheel'	RV 6.56.3 <i>cakrām hiraṇyāyam</i> 'the golden wheel'	N-Adj
6	Y. 57.27 <i>auruuaṇtō aurša raoxšna</i> 'white, shining racers'	RV 4.15.6 <i>ārvantam (...)</i> <i>sānasīm aruśām</i> 'horse that wins a prize, (like) the red (child of the sky) ...'	N-Adj(-Adj)
7	Yt. 10.31 <i>raθβīia vaca</i> 'with (temporally) proper words'	RV 1.190.2 <i>ṛtvīyāḥ (...)</i> <i>vācaḥ</i> 'speeches according to the season'	N need not be raised Cf. also A 1.11

Table 4.3: Lexically identical syntactic cognates of the oldest Iranian and Indo-Aryan representatives, based on Schlerath (1968, pp. 148–164).

These are perfect syntactic cognates, both in terms of parametric settings and lexical identity. Note that differing case endings of the head noun (e.g., accusative in Young Avestan, nominative in Vedic) signal the argument or adjunct function of verbal complements, embedding NPs into their higher clausal setting. They are irrelevant for analyses of NP-internal syntactic aspects. Only the configurations of the NPs themselves are of interest, as they are defined by a certain subset of parametric values.

Examples like these probably come closest to the usual concept(s) of syntactic cognates, irrespective of the particular theoretical framework (Construction Grammar etc.). And at first sight, this is no invalid assumption. At second sight, however, I would like to point out that a phonological analogon to such seemingly perfect cognates may be seen in sequences of two phonemes that are phonologically very close (e.g., Greek *du* : Luvian *ttu* : Hittite *dā* in Greek ὀδύνη 'pain, harm, sorrow', Aeolic ἐδύν-, Luvian *attuwal-* and Hittite *idālu-* 'evil'. But what about Armenian *rk* as in *erkn* 'birth-pangs'? A trained Indo-Europeanist will know that Armenian *rk* is a regular descendant of Proto-Indo-European \**du*, and that a correspondence between Greek *du* : Armenian *rk* may be fully reg-



ular, even though this must be far from obvious to the uninitiated. Real cognates, no matter whether based on phonemes or syntactic units, must allow for synchronic arbitrariness in the individual languages and can, but need not be obviously similar at first sight. In other words, syntactic cognates may be fully parallel NPs, but they can also have a superficially divergent appearance. Only a detailed parametric analysis can differentiate between true syntactic cognates and chance resemblances, and can reveal underlying perfect cognates which, however, superficially seem to differ fundamentally. To give an example, and contrary to intuition, but defined by N-raising parameters, nouns need not be raised; they can also surface after adjectival or genitival arguments, as shown by Table 4.4.

Again, the NPs listed in this Table are lexically identical syntactic cognates. Let us first consider NPs with a genitival argument, as found in examples 1, 2 and 3 of Table 4.3. These NPs reflect positive values of parameters FGN, FSN, FNN (morphological number marking on N), a negative value of DGP (no anaphoric definiteness marking) and CGB (boundedness), positive values of parameters GFR, GUN and GPR (uniform inflectional genitive, and the whole DP/NP can be raised over the ‘free’ genitive), negative GAD (no adpositional genitive), negatively set AGE (no adjectival genitive) and positive NGO (N over genitives). Other parametric values (structured genitivus obiectivus, parameters defining person or gender marking, adjectival behavior, relative clause surfacing etc.) are irrelevant for the genitival NPs of examples 1, 2 and 3. Their formal ‘skeleton’ can be stated as:

The settings of Table 4.5 constitute a blueprint, allowing us to configure NPs with one genitival argument. This configuration implies that the noun in Proto-Indo-Iranian could (but did not have to) be raised over the genitive, so that the genitival element could surface both before and after its head noun (visible in example 3 of Table 4.3).<sup>35</sup> The genitival relation was marked by means of a ‘uniform’ inflectional ending (i.e., it did not change according to the position of

---

<sup>35</sup>This may be compared with allophonic variation.

Nr.	Iranian	Indo-Aryan (or Iranian cognate)	Comment
1	Y. 1.16 <i>šōiθranqmca gaiiaoitiŋqmca</i> Y. 2.16 <i>šōiθrāscā gaōiiaoitišca</i> Yt. 8.42 <i>asō.šōiθrāscā gaōiiaoitišca</i> 'pasture land' (lit. 'home pasture')	RV 6.47.20 <i>agavyūtī kṣétram</i> 'infertile land/country void of pastures'	-ca in the Iranian examples; Ved. alpha privative
2	Y. 10.21 <i>haoməm zāirīm bəṛəzaŋtəm</i> 'the high, golden/yellow Haoma' Vd. 19.19 <i>haomasca zāiriš bəṛəzō</i>	RV 10.170.1 <i>bṛhāt (pibatu) somyám mādhu</i> 'the high Soma mead' RV 9.75.1 <i>bṛhán [sómah]</i> but RV 9.103.4 <i>sómah ... háriḥ</i> 'Soma, the yellow-colored/hued' RV 10.96.6 <i>sómāḥ hárayaḥ</i>	'Sauma' is a N in Ir., but Adj. in RV 10.170.1
3	Y. 19.2 <i>vīspa vohu</i> 'all good (things)'	RV 1.31.9 <i>vāsu vísvam</i> 'all good (things)' but RV 1.58.7 <i>vísvešām ... vásūnām</i> RV 1.113.7 <i>vísvasya ... vásvaḥ</i> RV 1.84.20 <i>vísṽ ... vásūni</i> RV 8.46.16 <i>vísvešām ... vásūnām</i> (Y 34.7 <i>vanhəuš vaēdānā mananḥō</i> Y. 34.9 <i>vanhəuš əvistī mananḥō</i> )	Vedic archaic order only in RV 1.31.9
4	Y. 28.5 <i>manascā vohū vaēdāmnō</i> '(I will see) the Vohu Manah as a knowing one' cf. Y. 28.11 <i>ašəm ... manascā vohū</i> Y. 51.18 <i>mananḥō ʔ vanuḥiš vido</i> Y. 49.10 <i>manō vohū urunascā</i> <i>ašāunqm</i> 'the Good Thought and the souls of the Righteous'		
5	Y. 31.6 <i>haiθīm mąθrəm</i> 'right word'	RV 1.67.5 <i>mántrebhiḥ satyāiḥ</i> 'with true words'	
6	Y. 31.13 <i>cašmāŋg θβisrā</i> 'with a sparkling eye'	RV 5.8.6 <i>tvešám cákṣuḥ</i> 'sparkling eye'	
7	Y. 32.3 <i>daēuuā vīspānḥō</i> 'all gods'	RV 1.3.8 <i>vísve devāsaḥ</i> 'all gods'	
8	Y. 46.19 ( <i>mīzdām ...</i> ) <i>parāhūm</i> '(reward ...) [of] the Other Life; [auf das künftige Leben bezüglich]'	RV 1.140.8 <i>āsum páram</i> 'higher spirit'	
9	Y. 51.20 <i>hazaošānḥō vīspānḥō</i> 'all like-minded (gods)'	RV 1.131.1 <i>vísve sajošasaḥ devāsaḥ</i> 'all like-minded gods'	cf. RV 1.186.2
10	Yt. 8.18 <i>aspāhe (...)</i> <i>aurušahe</i> 'of a white horse'	RV 7.57.6 <i>arušāsaḥ ásvāḥ</i> 'red horses'; also 7.97.6	Different order in RV 5.59.5
11	Yt. 13.52 <i>asušca aspō</i> but Yt. 17.12 <i>aspānḥō ... āsauuō</i> 'swift/fast horse(s)'	RV 7.71.5 <i>ášúm ásvam</i> RV 5.55.1 <i>ášvaiḥ (suyámebhiḥ) áśúbhiḥ</i> 'id.'	Hom. ὠκέες ἵπποι also RV 10.78.5, 8.13.11

Table 4.4: Lexically identical cognate NPs of the oldest Iranian and Indo-Aryan representatives with irrelevant surface differences in terms of word order, based on Schlerath (1968, pp. 148–164).

the genitive within or outside the structured NP core) and the noun itself was morphologically marked for number. This pattern is valid for Young Avestan, Old Avestan and Vedic, and also for Proto-Indo-Iranian, but not, for instance, for modern Pashto. Pashto has another configuration (see Table 4.6), differing from the one of its older cousins and reconstructed ancestor Proto-Indo-Iranian in several values (inter alia parameter GFO must be added).

Translated into plain language, Pashto NPs with basic genitival arguments

	FGN	FSN	FNN	DGP	CGB	GFR	GUN	GPR	GAD	AGE	NGO
OAv.	+	+	+	–	–	+	+	+	–	–	+
YAv.	+	+	+	–	–	+	+	+	–	–	+
Ved.	+	+	+	–	–	+	+	+	–	–	+
PIIr.	+	+	+	–	–	+	+	+	–	–	+

Table 4.5: Basic genitival NPs (Old Avestan, Young Avestan, Vedic and Proto-Indo-Iranian), parametric configuration.

	FGN	FSN	FNN	DGP	CGB	GFR	GUN	GPR	GFO	GAD	AGE	NGO
PIIr.	+	+	+	–	–	+	+	+	0	–	–	+
Psht.	+	+	+	–	–	+	–	–	–	+	0	–

Table 4.6: Basic genitival NPs (Proto-Indo-Iranian and Pashto), parametric archetype/configuration.

are much more restricted than their Proto-Indo-Iranian predecessors. Nouns are morphologically marked for number as well, but genitives can only occur as free, recursive, adpositional genitives to the left of the whole NP. This surface structure is the direct result of these configurational settings.<sup>36</sup> Pashto thus markedly differs from its ancient relatives and its own predecessor. From a purely configurational point of view and considering the fact that lexical identity is not necessary, Young Avestan (Y. 13.1) *ratūm ā yānqnm* ‘time of the women’, Vedic (RV 5.46.8) *ṛtúḥ jānīnām* ‘id.’ and Pashto *de dārs (de dawre) wákht* ‘time (period) of study’ (lit. ‘of study [of period] time’, see Tegey and Robson 1996, p. 185) would be syntactic cognates. But one could also take any other example of an NP with one basic genitival argument (e.g., Pashto *de asád (de) plār* ‘Asad’s father’, see Tegey and Robson 1996, p. 185, or RV 1.124.5 *gāvām jānitṛī* ‘the mother of the cows’ or Y. 29.2 *taša gēuš* ‘the fashioner of the cow’).

Since nominal macro- and mesoparameters with scope over a very great part of the lexicon are compared – all genitives, all nouns –, lexical identity is super-

<sup>36</sup>Note that parameter AGE (adjectival genitives) is neutralized in Pashto due to an implicational relationship with another parameter not relevant for genitival NPs (APO). On the other hand, parameter GFO must be set with a negative value, because Pashto has no uniform genitive. The opposite is true for its ancient relatives: due to the presence of a uniform genitive, the old Indo-Iranian languages had a neutralized GFO.

fluous. Relevant is only the fact that the languages to be compared have NPs with one noun and one dependent genitive. A default assumption would be that the parameters stored in the functional heads of various lexical items scattered throughout the whole lexicon are identical for all NPs of this type. We can expect that this is not true for parameters with a smaller scope (micro- and nanoparameters). Investigations also taking into account such parameters pertaining only to a few lexical items necessarily must compare lexically identical phrases or syntactic structures (as was the case with Walkden 2014). But in macro- or mesoparametric comparisons, lexically cognate phrases do not lend greater credence to the argument.

Since all parameters are the same for a certain lexical class, a macro- or mesoparametric correspondence set can compare abstract types of phrases, as long as the same classes of constituents or words are found in these types of phrases. This can be compared with the fact that, for instance, phoneme sequences like PIE *\*k<sub>u</sub>* retain their nature as consonant cluster or sequence of two different consonants, but there are different phonemes (configurations of basic units) in different daughters, such as the development in ‘central’ or ‘non-peripheral’ Iranian to *\*sp* > Pashto *sp* (e.g., PIIr. *\*ćuakas* > *\*spakah* > Pashto *spay* ‘dog’ [NOM.SG], *\*aćuaH* > *\*aspā* > Pashto *aspa* ‘mare’ [NOM.SG]), whereas other daughters have different configurations (as was discussed in the introductory Section of Chapter 3). Macro- and mesoparametric syntactic cognates are cognate types of NPs (or VPs, for that matter) and reconstructions aim at their specific configurations, irrespective of varying lexical material. Phonemes are the same across the lexicon and it is not evident why the same should not be true for macro- or mesoparameters.

The abstract sum of all parametric settings can be visualized in form of a parametric Feature Matrix which functions as the syntactic equivalent to a phonological Feature Matrix. On the other hand, the specific configurations valid for certain structures such as NPs with one simple genitival argument (or NPs with two free, recursive genitives etc.) rather resemble phonotactic restrictions. Just

as phonotactics deals with permissible combinations of phonemes (syllable structure, consonant clusters, vowel sequences etc.), configurational syntax defines the types of syntactic structures that are allowed in a particular language.

### 4.3.2 Syntaxemes: syntactic phonemes

Certain configurational settings should always result in the same surface appearances, no matter whether the lexical material is identical or not in the daughter languages – similar to phonotactic rules guiding the appearance of roots and morphemes. Relevant is the presence of distinct types of larger syntactic structures, similar to aspects of edge phonotactics which define licit bi- or tripartite word- or syllable-initial or final clusters like (from a very abstract point of view, and with P = any stop/plosive, R = resonant, C = consonant, F = fricative, O = obstruent) #PR-, #PP-, #PPC-, #PsC- or -RO#, -OF#, -RCs# etc.

The syntactic equivalent to such subtypes of different consonant clusters in specific contexts are, for example, basic NPs with only one structured adjective, or NPs with one genitival argument, or NPs with a genitive and demonstrative, NPs with demonstrative, structured adjective and dependent genitive etc. In as much as notations like #PR- are macrocategories subsuming various subtypes (in the case of Proto-Indo-European, *\*pn-*, *\*dn-*, *\*ǵn*, *\*ǵ<sup>h</sup>n-*, *\*b<sup>h</sup>r-*, *\*dl-* etc.), the syntactic macrocategory is a specific subtype of NPs, and the parametric configuration relates to the question of which basic units (~ phonemes) are allowed in which combinations. Since this study focuses on nominal macro- and mesoparameters with scope over a great or even the greatest part of the relevant subsection of the lexicon (all nouns, all adjectives, all nouns in a genitival relation etc.), it is not evident why lexical identity should be necessary.

The stem of a word (at least in Proto-Indo-European terms) consists of a root and an optional derivational suffix (e.g., PIE *\*-o-*, *\*-ro-*, *\*-u-*, *\*-men-*, *\*-ter-* etc.) attached to this root. Is it possible to transfer this concept to syntax? Taking the analogy to phonotactics one step further, let us assume that a syntactic ‘root’ might be a VP with one (intransitive) or two (transitive) or even more central

arguments (NPs), on which several ‘affixes’ (other phrases or dependent clauses) could optionally be attached to change the semantics of the whole complex, leading to fine-grained nuances (cf. Proto-Indo-European *\*dh<sub>3</sub>-tér* ‘one whose role is to give’ and *\*déh<sub>3</sub>-tōr* ‘one who is a giver by virtue of having actually given something’; semantic reconstructions as per Fortson 2010, p. 124). Similar to the fact that PIE had root presents (e.g., *\*h<sub>1</sub>es-* ‘be’) and root nouns (e.g., *\*ped-* ‘foot’), we may assume that it had root clauses (a simple VP with arguments), but since also derived types of words can be reconstructed, syntactic analogues to derived words should have been present as well. In my view, it is possible to conceive of syntactic structures as resembling certain types of

- bare roots (e.g., PIE *\*CeC* as in *\*pet-* ‘fly’, *\*ped-* ‘foot’ or *\*d<sup>h</sup>eg<sup>uh</sup>-* ‘burn’; *\*CReC* as in *\*d<sup>h</sup>uer-* ‘door’, *\*sneh<sub>2</sub>-* ‘sew’, *\*suep-* ‘sleep’ or *\*sreu-* ‘flow’; *\*CReRC* as in *\*kreuh<sub>2</sub>-* ‘gore’, *\*sueh<sub>2</sub>d-* ‘sweet’ or *\*mleuh<sub>2</sub>-* ‘speak’);
- suffixed verbs (e.g., R/root + affix/suffix, such as PIE present formations like *\*R(é)-je/o-*, e.g., visible in *\*spék-je/o* ‘to look, view’ or *\*R(z)-ské/ó-* as in *\*g<sup>u</sup>m-ské/ó-* ‘to come’; causative *\*R(o)-éje/o-* like in *\*mon-éje/o-* ‘to remind’ (‘to let think’); s-aorist *\*R(ē)/R(é)-s-* as in *\*d<sup>h</sup>éig<sup>h</sup>-s-/d<sup>h</sup>éig<sup>h</sup>-s-* ‘to shape, mold’), or
- suffixed nouns (e.g., athematic nouns built with *\*-uer/uen-*, *\*-men-*, *-sen-*, *-ten-* or *\*-ter-* and *\*-tor-*, or thematic nouns in plain *\*-o-*, or *\*-mo-*, *-tlo-*, *-tro-* etc.).

Words in Proto-Indo-European had the structure R (+ S) + E, i.e., root (+ optional suffix) + ending, notwithstanding special cases like nasal infix presents, reduplication etc. Whereas nominal inflectional endings signal the relation of NPs to their verbal head (and verbal ones, among others, person), the syntactic equivalent to endings is unclear – provided that such a parallelism is indeed possible. If so, might other phrases serve a similar function? Is discourse-related or pragmatic information the syntactic equivalent to inflectional endings?

Even though such theoretical questions cannot be satisfactorily answered in the present study, I would nonetheless like to put them up for discussion. Further progress in the identification, analysis and comparison of syntactic cognates crucially depends on a new conceptualization of syntax – one that is reminiscent of phoneme-based analyses of and approaches to form-meaning pairings stored in the lexicon of particular languages. Words have fixed and variable parts in many languages of the world, and it might be a promising idea to search for a syntactic analogue, dependent upon an elaboration of all basic syntactic units or rather syntactic equivalents to phonemes. I will leave this line of thought open at this point in time and only focus on the syntactic equivalent to word stems in the following discussion.<sup>37</sup>

Under the assumption that VPs are syntactic equivalents to roots, and taking as a starting point the idea that there is a structural parallelism to phonotactics, one possibility may be to conceive of VPs as right edge (or ‘word’ or ‘root’ final) combinations of ‘syntactic phonemes’ or rather clusters of basic units (in this case verbal parameters); NPs would then form the left edge (‘syllable onset’, or ‘word-initial’ clusters vel sim.), so that an intransitive syntactic root could be defined as NP–VP or VP–NP.

However, a more elegant parallelism to roots and words will regard VPs or rather their verbal heads as syntactic equivalents to vowels which form the prosodic head of syllables (ignoring syllabic resonants for the moment). NPs may then form both syllable-/root-initial and -final clusters, and nominal constituents

<sup>37</sup>Another interesting question relates to the possibility of a syntactic equivalent to ablaut/apophony. ‘Syntactic reduplication’ is another interesting case in point, as is the question of syntactic equivalents to infixes in the morphological domain. Can phenomena like the Chorasmian ‘principle of anticipation’ (Henning 1955, p. 48) be regarded as cases of ‘syntactic reduplication’? In Chorasmian, a pronominal suffix is attached to any first syntactic unit of a clause that anticipates an NP surfacing later (the pronominal clitic is obligatory, the NP not), cf. QR 12 *hāβr-ina-hi-di, yā duyd-a-mi* <h’βr-n.<sup>i</sup>y.-d<sup>i</sup>, y’ δy°d.-a’.-m<sup>i</sup>> ‘I have given **her** to you, **my daughter**’ (Durkin-Meisterernst 2009, p. 360). Another illustrating example can be seen in QR 157: *ka-mi-hi hiβr-ī-kām-ī ī zirnī nē(n)*... <k.my.h h’βr.-yk’-m.<sup>i</sup>’y zrny ny> ‘If you will give me this gold (ī zirnī nē(n) lit. ‘the gold this’)...’ (Durkin-Meisterernst 2009, p. 367). Another question that comes up: Is parataxis the syntactic parallel to compounding?

(nouns, adjectives, genitives, demonstratives, numerals) may be equated with different types of consonants. A transitive syntactic ‘root’ ( $\sim$  *\*ped-*) can then be stated as NP<sub>A</sub>–VP–NP<sub>P</sub> (with NP<sub>A</sub> = agent and NP<sub>P</sub> = patient).<sup>38</sup> This concept implies that the syntactic ‘words’ we should focus or reconstructive efforts on are specific arrangements of phrases with specific parametric configurations each.

Real syntactic correspondence sets can then only be assembled, if analyses of the most important parameters (macro- and mesoparameters) of the two basic types of ‘syntactic sounds’ – NPs *and* VPs – have been carried out. Before we can try to recover such a syntactic ‘root’ or ‘word stem’ (ignoring the ‘ending’),<sup>39</sup> we first need to carry out parametric analyses, determining which types of phrases with which parametric configurations are allowed in a set of genetically related languages (paralleling edge phonotactics). The present study is a first step, focusing on NPs – so to say the consonantal syllable onsets and codae of syntax. But without a subsequent study taking VPs into consideration, we can only state which types of syntactic equivalents to consonant clusters were allowed in Proto-Indo-Iranian and Proto-Indo-European. This is not enough. Full syntactic ‘word’ stems require additional, verbal ( $\sim$  vocalic) information.

But there is a caveat: if certain types of NPs are equivalents to certain types of consonant clusters (or arrangements of phonemes), a major difference will pertain to the number of basic units forming these clusters. Proto-Indo-European consonant clusters consisted of two or three phonemes (e.g., *\*h<sub>2</sub>stér-* ‘star’, *\*krót-u-* ‘insight, intelligence’ or *\*b<sup>h</sup>léǵ<sup>h</sup>-mon-* ‘the one of the sacred formulation, priest’), and not of a whole array of basic units as in Table 4.5 or 4.6. I do not know of an consonant cluster in an Indo-European language consisting of 11 or 12 phonemes.

Up to this point, I have equated parameters with phonemes. However,

<sup>38</sup>Note that the medial placement of the VP does not indicate that VPs are base-generated in second position, but that this representation was chosen simply to parallel a normal PIE root structure with vowels in medial position (*\*CVC*).

<sup>39</sup>Recall the fact that we usually also reconstruct stems and not finite word forms.



if we tentatively equate parametric values not with phonemes, but with distinctive features combining to phonemes, the parallel to consonants and consonant clusters will be a perfect one. Phonological models explain variation between different phonemes as being due to different compositions of smaller distinctive features. Some features pertain to the place and manner of articulation, others to voice etc. A subset of distinctive features combines, in specific configurations, to certain phonemes. Phonemes can be described as bundles of different phonological features (some with greater scope, others with smaller), and this can be visualized in feature bundles like the following (reduced) one for German /d/.

$$/d/ = \begin{bmatrix} C \\ - \textit{cont} \\ + \textit{voice} \\ + \textit{coronal} \\ + \textit{ant} \\ - \textit{labial} \\ - \textit{dorsal} \\ - \textit{lateral} \\ - \textit{strident} \end{bmatrix}$$

The Appendix gives an example of a phonological Feature Matrix. A quick glance reveals that the parallelism to the parametric Feature Matrix we arrived at is striking. If parameters are a syntactic parallel to distinctive phonological features, they can be expected to form bundles. Certain syntactic constituents such as (in the nominal domain) nouns, adjectives, genitival arguments, demonstratives, cardinals etc. may be equated with different types or classes of sounds (vowels, glides, nasals and liquids, affricates, non-coronal and coronal obstruents etc.). Individual phonemes are defined by specific configurations of phonological features, and some features may pertain only to one specific class or only one phoneme of a class (e.g., a manner feature like *nasal* is only relevant for nasals, a place feature like *round* is relevant for labial non-coronal obstruents and a few vowels, but only for /m/, /pf/ among nasals and affricates, and utterly irrelevant for other sounds). What if the same is true for syntactic parameters?

Several parameters together define genitival NPs, as described above, but if we reduce the number of 12 separate features to two bundles – one feature bundle for a (basic) genitival element ( $\sim$  one single phoneme), and another one for the head noun ( $\sim$  another phoneme) – we will have a syntactic equivalent to a sequence of two consonants. If this concept is correct, syntaxemes will not be single parameters, but bundles of parameters, pertaining to different phrasal components. See Fig. 4.15 for a visualization of this concept.

$$\left[ \begin{array}{c} Gen \\ + FG_N \\ + FG_G \\ + FS_N \\ + FN_N \\ - DG_P \\ - CG_B \\ + GF_R \\ + GU_N \\ 0 GF_O \\ - GA_D \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FG_N \\ + FG_G \\ + FS_N \\ + FN_N \\ - DG_P \\ - CG_B \\ + GP_R \\ + NG_O \end{array} \right]$$

Figure 4.15: Proto-Indo-Iranian, basic genitival NPs, syntaxemic representation (e.g., Y. 50.10 *raocā x<sup>v</sup>āṇg* ‘the rays of the sun’ or RV 3.22.3 *rocané (...) sūryasya* ‘in the sun’s realm of light’).

The values are those of reconstructed Proto-Indo-Iranian, which, as we have seen, are synonymous with those of Old Avestan and Vedic. Prima facie, more parameters define the genitival argument than the higher-ranking nominal head, but this is just an artifact of the representation. In principle, also phonemes should be stated as feature bundles taking into account all settings (i.e., a whole column of the phonological feature matrix), but in practice only relevant features are given (for instance, I omitted, in the case of /d/ above,  $-son$ ,  $-syll$ ,  $0back$ ,  $0high$  etc.) The same pertains to syntactic parameters; both columns or feature bundles constitute a selection of relevant features for this type of syntactic elements, yielding a specific configuration. In theory, a whole column of the parametric feature matrix could be given in each case, but in practice this can be

reduced to relevant features only. These minimalized feature bundles might be regarded as the real syntaxemes.

The Pashto equivalent to the Old Avestan (and Vedic and Proto-Indo-Iranian) syntaxeme cluster is visualized in Figure 4.16.

$$\left[ \begin{array}{c} Gen \\ + FG_N \\ + FG_G \\ + FS_N \\ + FN_N \\ - DG_P \\ - CG_B \\ + GF_R \\ - GU_N \\ - GF_O \\ + GA_D \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FG_N \\ + FG_G \\ + FS_N \\ + FN_N \\ - DG_P \\ - CG_B \\ - GP_R \\ - NG_O \end{array} \right]$$

Figure 4.16: Pashto, basic genitival NPs, syntaxemic representation (e.g., *de asád (de) plâr* ‘Asad’s father’, following Tegey and Robson 1996, p. 185).

What exactly does that concept involve? Is the difference between Old Avestan and Pashto genitival NPs similar to one between sequences of, say, Old/Proto-Iranian *\*fš-* and Pashto *wž-* (cf. *\*fšančij-* > Pashto *wžənj*, *žmanj* ‘comb’; see Skjærvø 1989b, p. 404) in the two languages? Or rather to the difference between *\*-rz-* and *-žd-* (*\*brzah* > Pashto *(w)užd* ‘long’, or *\*rarzajā-* > *režd-edāl* ‘tremble’)? The differences in terms of the parametric configurations between Pashto and Old Avestan are not insignificant. To what extent do they correspond to and differ from phonemes? Are we observing differences comparable to phonological changes in terms of voice or place/manner of articulation, or whole class changes as in the latter example (*-rz-* : *-žd-*)? To what extent do syntactic feature bundles resemble different phonemes?

Given that the concept is correct, Old Avestan and Vedic had the same syntactic ‘phoneme’ with regard to genitival arguments of NPs/nouns, and Old Latin and Archaic Greek as well. With regard to nouns in genitival NPs, all four ancient Indo-European languages agree, too; the parametric feature bundle for ba-

sic genitival NPs, i.e., those consisting of a noun and a single dependent noun, given in Figure 4.15, is valid for all of them. Compare this with the fact that all four mentioned Indo-European languages had basic clusters of the type *-nt-* (notwithstanding secondary palatalisations, as e.g. in the Greek feminine participles etc.).

Relevant for historical-comparative purposes is the heritability and reconstructibility of such different types of syntaxeme configurations. If the same types of NPs are found in genetically related languages, irrespective of lexical identity of the phrasal (or rather syntactic) components, we can surmise that the same type of NPs was present in the parent language and that we are dealing with cognates in an abstract sense – cognate syntactic equivalents of consonant clusters, of syllable onsets or codae, so to speak. The particular subtype of phrases or syntactic bodies can then be reconstructed for the parent language, and the relevant question relates to the exact configuration of this subtype of NPs (or other syntactic structures), depending on the situation in the daughter languages.

Let us now define the configuration of simple NPs with only one structured adjective. Consider examples 4, 5, 6 and 7 of Table 4.3 as well as Table 4.4; note again that even though these tables list lexically identical phrases, lexical identity is not necessary – it only serves an aesthetic purpose. Irrespective of surface differences in noun placement (defined by N-raising), there is feature spread of the noun to structured adjectives in terms of gender and number (as well as case) and neither boundedness nor anaphoric or other definiteness marking are relevant for the configuration of NPs (see Figure 4.17).

Note that the parametric configuration for nouns in adjectival NPs is not identical to that of genitival NPs (NOA : NGO and GPR). The adjectival configuration is valid for Old Avestan, Vedic and Proto-Indo-Iranian, but Archaic Greek and Old Latin differ in having a positive value for parameter AGE (adjectival genitives). This is one single difference. Can we compare this with a difference between *\*t* and *\*d*, or between *\*p* and *\*f*, or between *\*k̑* and *\*k*? Is this similar

$$\left[ \begin{array}{c} Adj \\ + FGN \\ + FGG \\ + AST \\ + FFS \\ + APO \\ - AGE \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FGN \\ + FGG \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NOA \end{array} \right]$$

Figure 4.17: Proto-Indo-Iranian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (e.g., RV 6.56.3 *cakráṃ hiraṇyáyam* ‘the golden wheel’).

to a voicing contrast, or to a change in the place or manner of articulation? Did Proto-Indo-Iranian have the innovated form, or rather its western cousins? We do not yet know. Only subsequent studies determining first the parametric values of other ancient (and perhaps modern) Indo-European languages and then comparing their configurations will lead to further clarification. Right now, all that can be said is that the ancient Indo-Iranian languages and two of their western Indo-European sister languages did not substantially differ in terms of the syntactic configuration of two very basic types of NPs.

The two parametric archetypes discussed so far define the configuration of NPs with plain (i.e., only nominal) genitival arguments and NPs with structured adjectives. Both are frequent types of NPs in ancient Indo-European and Indo-Iranian languages, and it was reasonable to introduce my conceptualization of syntactic cognates or rather syntaxemic clusters by means of these two types. To proceed further, we will now establish which configurations pertain to other basic nominal syntaxeme sequences of Proto-Indo-Iranian, i.e., to those with demonstratives, numerals and verbless relative clauses. Figure 4.18 visualizes the syntaxemic representation of a basic NP with a demonstrative pronoun in Old Indo-Iranian languages and Proto-Indo-Iranian.

$$\begin{bmatrix} Dem \\ + FGP \\ + FGN \\ + FGG \\ - DGP \\ - TPL \end{bmatrix} \text{ — } \begin{bmatrix} N \\ + FGN \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NDE \end{bmatrix}$$

Figure 4.18: Proto-Indo-Iranian, basic NP with demonstrative, syntaxemic representation (e.g., Y. 36.6: *imā raocā* ‘these lights’ or Y. 32.13; 45.3,4 *aṇhōuš ahiiā* ‘of this world’).

Note that only in demonstratives the highest parameter FGP (person marking) becomes relevant in the nominal domain. On the other hand, NPs with cardinal numbers had a very simple configuration in Proto-Indo-Iranian (see Figure 4.19).

$$\begin{bmatrix} Card \\ + FGN \\ + FGG \\ + CPS \end{bmatrix} \text{ — } \begin{bmatrix} N \\ + FGN \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NOC \end{bmatrix}$$

Figure 4.19: Proto-Indo-Iranian, basic NP with cardinal numeral, syntaxemic representation (e.g., Khotanese, Z. 22.162 *ratanyau haudyau* ‘(full of) seven jewels’).

It is a remarkable fact that some numerals were inflected for gender in Proto-Indo-Iranian (e.g., Vedic *tisrās*, Young Avestan *tišrqm* ‘three.F’, Vedic *cātasras*, Young Avestan *čataṇrō* ‘four.F’), whereas others were not (as is well-known, this is one of several [Italo-]Celtic–Indo-Iranian isoglosses).<sup>40</sup> Maybe parameter FGG

<sup>40</sup>Vendryes (1918) pointed out that several archaic lexical elements are shared by Indo-Iranian in the East and Italic as well as Celtic in the West of the Indo-European Kulturkreis. This is usually explained as representing *Randsprachenarchaismen* of a special legal or religious sphere (e.g., Latin *rēx* ‘king’ = Old Irish *rí* = Vedic *eka-rāt* ‘ruling alone’, Old Irish *rígain* = Vedic *rājñī* ‘queen’, Latin *iūs* ‘law’ = Vedic *yóh* or the famous correspondence between Latin *crēdō* ‘I believe’, Old Irish *creitim* and Vedic *śraddhā* ‘trust’). The feminine forms of ‘3’ and ‘4’ must be mentioned in this context as well (Old Irish *cetheoir*, Welsh *pedair* ‘four.F’).

should be left out of the feature bundle for cardinals and better be regarded as a micro- or nanoparameter (or a new nanoparameter should be posited) pertaining only to a few numerals.

The last syntaxemic representation of a basic NP type refers to NPs with verbless relative clauses (see Figure 4.20).

$$\begin{bmatrix} N \\ + FGN \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NOE \end{bmatrix} \text{ --- } \begin{bmatrix} Rel \\ + FGN \\ + FGG \\ + VRC \\ + ADR \end{bmatrix} \text{ --- } \begin{bmatrix} N \\ + FGN \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NOE \end{bmatrix}$$

Figure 4.20: Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 31.6 (West 2011, p. 147) *mąṛrām yim hauruuatātō* ‘a spell [which is one of] integrity ...’).

The difference with regard to nouns always concerns a parameter defining the raising (here NOE; another parameter, ADR, pertains to the relative pronoun); otherwise the nominal core remains stable. If, e.g., a demonstrative is also present in the particular noun phrase to be syntaxemically analyzed, the nominal parameter to be set and included will be NDE. Consider Figure 4.21 which also illustrates the fact that elements can be combined to describe larger structures – in this case, an NP involving a verbless relative clause *and* a demonstrative. In contrast, Figure 4.22 presents a different order of the same phrase, in accordance with the base-generated order. This is only a notational variation, the NP as such is analyzed the same way in both cases. The variability in order is the only real difference to sequences of phonemes – whereas morphemes or lexemes actually consist of fixed phoneme sequences or rather a concatenation of basic phonological units, syntactic structures rather resemble an orrery with mobile elements. It is irrelevant how the parametric bundles are horizontally arranged, as this is, anyway, only a convenient way to visualize these abstract syntactic building blocks and their mutual interplay.

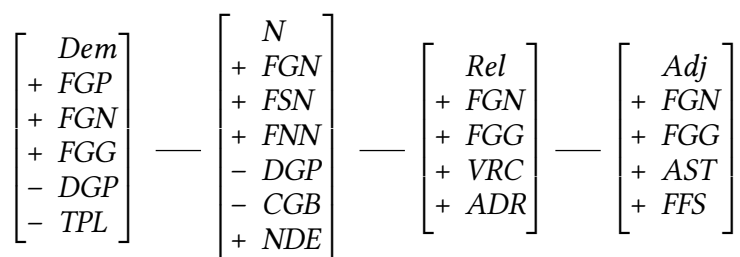


Figure 4.21: Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 35.4 *tāiš šiiāoθānāiš yāiš vahištāiš* ‘with those deeds that (are) the best’).

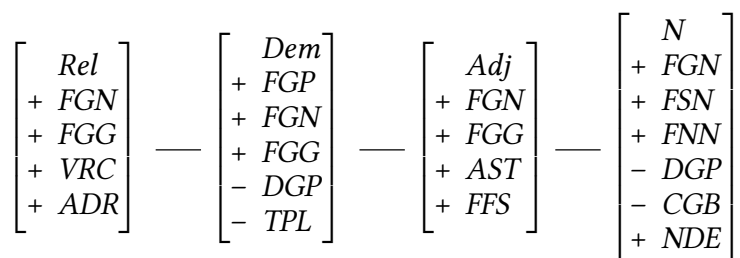


Figure 4.22: Proto-Indo-Iranian, basic NP with a verbless relative clause, syntaxemic representation (e.g., Y. 35.4 *tāiš šiiāoθānāiš yāiš vahištāiš* ‘with those deeds that (are) the best’); base-generated order.

Figures 4.21 and 4.22 show only one NP which consists of four feature bundles or rather syntaxemes. Probably the limit is around 6 *different* elements, notwithstanding repetitions of one class or type of elements (e.g. recursive genitives). Consider e.g., the complex Young Avestan NP (Yt. 13.91) *staota ašahe yaṭ mazištaheca vahištaheca sraēštaheca* ‘The prayer of Aša, the greatest and best and most beautiful’ in Figure 4.23. I would render this with only four different elements, because in my eyes repeated elements like the ones of this phrase are reminiscent of phonological lengthening or geminated consonants (vel sim.). Qualitatively, the elements differ only in four aspects, the quantitative aspect is best ignored in this analysis.

A complete syntactic ‘syllable’, let alone a ‘word’, would also entail, as hypothesized above, a verbal core, head or nucleus (whatever the terminology),



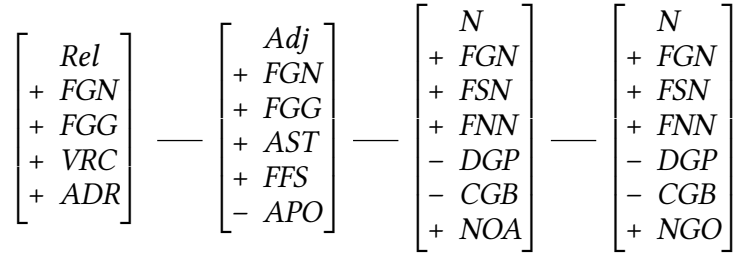


Figure 4.23: Young Avestan, complex NP with a verbless relative clause being in a genitival relation to a head noun, syntaxemic representation (e.g., Young Avestan [Yt. 13.91] *staota ašahe yaṭ mazištaheca vahištaheca sraēštaheca* ‘The prayer of Aša, the greatest and best and most beautiful’).

but an elaboration of this aspect must be left for future studies. For a start, NP elements can be defined according to their specific configurations of relevant parameters. The Proto-Indo-Iranian configurations have been given above, and on that basis the enormous surface variation with regard to NPs one encounters when reading the Vedic hymns or Old Avestan gathas appears to be nothing more than the result of a few very basic configurations of bundles of features and the N-raising they allow for.

Other, more complex NPs – for instance, a noun phrase consisting of a head noun and a genitival argument further defined by an adjective (e.g., Y. 45.4 *aṇhōuš ahiiā vahištəm* ‘the best one in this world’ with an NP<sub>Gen</sub><NP order (or rather [N–Dem]–Adj), see West 2011, p. 121) imply more varied parametric settings than those sufficient for basic types of NPs. This could be discussed at length, but the point should be clear by now. Specific subsets of parameters can be selected from all parametric values of Table 4.2 to account for different types of NPs, from moderately complex ones to those involving several arguments and adjuncts.

The most interesting observation is that it is possible to create an extremely large set of different combinations with a very limited set of syntaxemes or rather parametric feature bundles:

N -V;

N -V-[N];

Dem-N -V-[N];

Dem-N -V-[Dem-N];

Dem-N -V-[Dem-Adj-N];

Dem-Card-N -V-[Dem-Adj-N];

Dem-Card-N -V-[Dem-Adj];

Dem-Card-N -V-[Gen-Adj-N]-[Gen-N] etc.

This parallels the behavior of phonemes. Moderate phoneme inventories of 20–30 sounds create immense lexical bodies with tens of thousands of different items. A few atomic ingredients and you can create extremely variegated strings of these units, with just one substitution of a unit through another leading to a completely new word (e.g., PIE *\*pet-* ‘fly’ : *\*ped-* ‘foot’).

Syntactic structures such as clauses must be seen and analyzed in this way. Nominal elements like adjectives, nouns, genitives or numerals form syntactic equivalents of syllable onsets and codae, and verbal elements syntactic ‘vowels’; combined, they can be arranged in ultimately finite, but nonetheless manifold ways. So far, only the syntactic equivalent to consonants has been investigated, and the really interesting agenda still lies ahead of us. As soon as the full set of syntaxemes is known, i.e., also verbal ones, a search for patterns can be carried out. These patterns may consist of divergent clusters of ‘consonants’ in specific contexts – divergent in the sense that their constituents show unexpected configurations. Context-sensitive developments concerning phonemes are conventionally known as sound laws, and context-sensitive developments of syntaxemes may constitute syntactic laws. But that is a task for another study. What is now left for us to do is to demonstrate that parameters, or bundles of them, really are syntaxemes providing us with minimal pairs.

### 4.3.3 Syntactic Minimal Pairs

Syntactic minimal pairs are the equivalent to phonological minimal pairs (English *hat* : *had*; *red* : *dead* etc.). If parametric feature bundles are indeed syntactic equivalents to phonemes, being the smallest meaningful syntactic units, the difference in just one parametric value must manifest itself in a different behavior of NPs. This is exactly what can be observed in RV 10.96.6 *dāsā vṛtrāṇi* (*āryā*) ‘the Dasic (and Aryan) enemies’ (or ‘the hostile Dasas’) as compared with Archaic Greek (Il. 10.326) νῆ Ἀγαμεμνομένην ‘Agamennon’s ship’ (or ‘the Agamemnonian ship’) or Il. 2.658 (etc.) βίη Ἡρακλείη ‘for Herakles’ force’ or ‘for the Herakleian force’. The Greek NPs allow for a genitival reading, whereas the Vedic one does not. Consider Figures 4.24 and 4.25.

$$\left[ \begin{array}{c} Adj \\ + FGN \\ + FGG \\ + AST \\ + FFS \\ + APO \\ - AGE \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FGN \\ + FGG \\ + FSN \\ + FNN \\ - DGP \\ - CGB \\ + NOA \end{array} \right]$$

Figure 4.24: Vedic, adjectival NPs with genitival reading, syntaxemic representation (e.g., RV 10.96.6 *dāsā vṛtrāṇi* (*āryā*) ‘the Dasic (and Aryan) enemies’).

On the other hand, NPs with different surface orders of the noun and its arguments (N-Gen : Gen-N, Dem-N-Adj : Dem-Adj-N) are no minimal pairs as long as the compared languages allow for the same amount of N-raising. A minimal pair in such situations could only be brought forward, if one of the languages to be compared had completely parallel parametric values in all relevant cases (e.g., in NPs with structured adjectives, identical settings of adjectival parameters such as FFS or AST), yet not allowing for N-raising over, say, adjectives (whereas the other languages would permit N-raising here).

Superficially similar NPs must always be analyzed in terms of their whole

$$\left[ \begin{array}{c} Adj \\ + FG_N \\ + FG_G \\ + AS_T \\ + FF_S \\ + AP_O \\ + AG_E \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FG_N \\ + FG_G \\ + FS_N \\ + FNN \\ - DGP \\ - CGB \\ + NOA \end{array} \right]$$

Figure 4.25: Archaic Greek, basic adjectival NPs (attributive, not predicative), syntaxemic representation (e.g., Il. 2.658 (etc.) βίη Ἡρακλείη ‘for Herakles’ force’ or ‘for the Herakleian force’).

relevant parametric configuration. Even though NPs may look strikingly similar at first glance, their outward appearance can be the result of two or more underlying configurational differences. For instance, Old Persian (DZc 6) *bumiyā vazrakāyā* ‘great earth’ and Middle Persian *zamīg ēw wuzurg (ud istabr)* ‘a great (and solid) earth’ (Durkin-Meisterernst 2014, p. 361), despite looking similar, do not constitute a minimal pair, because two parameters are different: CGB, defining grammaticalized boundedness of the noun and FGG, asking for gender marking. The parametric setup of the Old Persian syntaxeme configuration would be as in Figure 4.26, whereas the Middle Persian equivalent is depicted in Figure 4.27.

$$\left[ \begin{array}{c} Adj \\ + FG_N \\ + FG_G \\ + AS_T \\ + FF_S \\ - AP_O \\ 0 AG_E \end{array} \right] \quad \text{---} \quad \left[ \begin{array}{c} N \\ + FG_N \\ + FG_G \\ + FS_N \\ + FNN \\ - DGP \\ - CGB \\ + NOA \end{array} \right]$$

Figure 4.26: Old Persian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (DZc 6 *bumiyā vazrakāyā* ‘great earth’).

But strictly speaking, nominal syntaxemes on their own are not sufficient for

$$\begin{bmatrix} Adj \\ + FGN \\ - FGG \\ + AST \\ + FFS \\ - APO \\ 0 AGE \end{bmatrix} \quad \text{---} \quad \begin{bmatrix} N \\ + FGN \\ - FGG \\ + FSN \\ + FNN \\ - DGP \\ + CGB \\ + NOA \end{bmatrix}$$

Figure 4.27: Middle Persian, basic adjectival NPs (attributive, not predicative), syntaxemic representation (*zamīg ēw wuzurg (ud istabr)* ‘a great [and solid] earth’, following Durkin-Meisterernst 2014, p. 361).

establishing real minimal pairs. Usually one compares whole words. What we have done so far was a comparison of the syntactic equivalents to two consonants. A comparison of Vedic *-tt-* and Avestan *-st-* does not constitute a proper minimal pair; it can be part of it, and in fact it can entail the exact difference one is interested in, but in reality several cognate words containing this cluster or sequence are necessary (e.g., Vedic *vittá-* and Avestan *vista-* ‘known’ beside other examples). This implies that whole clauses with cognate syntaxemes are required in the syntactic domain. The current situation resembles one in which we knew, for instance, that Vedic allowed for sequences like *-tt-*, whereas Avestan instead had another configuration *-st-*; but the exact contexts in which these sequences of different phonemes occurred would require the study of larger structures otherwise completely identical in order to search for possible context-sensitive special correspondences between the two compared languages. As long as one can only analyze, describe and reconstruct consonants and their rules of ordering (phonotactics) – for instance, *\*b<sup>h</sup>r* was a legitimate syllable onset in the proto-language, basic *\*g<sup>h</sup>* as well (or a coda), and a cluster *\*nt*, too –, the way to *\*b<sup>h</sup>rg<sup>h</sup>ent-* ‘mighty, high’ (Vedic *brhánt-*, Avestan *bərəzant-*) is far from obvious. Vowels are an integral part, and without a proper parametric analysis yielding syntactic equivalents to vocalic phonemes real cognates cannot be given.

The process of linguistic reconstruction involves several subsequent steps. At first, a phoneme inventory for a group of genetically related languages to be compared must be set up, and in a second step the particular values of the relevant basic units of the proto-language can be reconstructed. Building upon an idea of Ian Roberts, this study aimed at a demonstration of the fact that the same can be achieved for syntactic aspects by means of parameters. In a subsequent step not yet attainable, specific developments can come into focus. Given the structural parallelism between parameters and phonemes or rather phonological features, I propose that syntactic equivalents to sound laws are very likely to exist (probably being structurally different from Wackernagel's and Behaghel's so called laws and virtually identical to sound laws).

However, the important caveat relates to the fact that sound laws can only be detected when full words are compared. A comparison of syllable onsets or codae would lead nowhere. The discussion should have made clear that it is too early to posit context-dependent special developments of syntactic equivalents to phonemes. We do not have, as yet, a full parametric analysis of any clause, since the nucleus, the VP, has not been covered by this study. If only word-initial consonant clusters were known, it would be very difficult or even impossible to discern sound laws, as sound laws depend on a comparison of whole words. Very likely the same is true for syntactic equivalents to words. NPs are an integral part of them, but only if the whole picture is taken into account, if all parametric values are compared, certain context-sensitive special developments may be detected. I have shown the pathway and the method for doing this, and no doubt a future study aiming at VPs and combining VP and NP parameters will achieve this, but as yet it is too early to posit any syntactic laws.

If we had full knowledge of clausal macro- and mesoparameters and if we had carried out an analysis of the respective settings in Iranian languages, we would probably have detected one or more context-sensitive developments in each language. They must exist. If so, we could then had made use of this knowledge and reconstructed rather concrete formal syntactic skeletons with

very specific parametric configurations, and thus whole clauses instead of just parts of them as we can do now. Our knowledge of syntactic laws would then also help us in detecting previously overlooked syntactic cognates, as we could predict their syntaxemic structure by means of a mechanical transposition of the values and configurations of one language into those of another (Vedic *brhánt-* = Avestan \_\_\_\_?). In principle, the method of syntactic reconstruction is not very complicated, it just depends on a full account of nominal *and* verbal macro- and mesoparameters.

# Chapter 5

## Conclusion

“[T]he first law of comparative grammar is that  
you’ve got to know what to compare”

---

Calvert Watkins (1976, p. 312)

This dissertation focused on the hypothesis that syntax can be reconstructed using the traditional Comparative Method and structural analogues to distinctive phonological features, so-called syntactic parameters. The central thesis can be defined as: *Syntactic structures must be broken down into their smallest meaningful units or building blocks; syntax must be phonemicized, as only then the Comparative Method can be applied.*

In other words, any linguistic reconstruction depends on a previous deconstruction of the object of research. It does not make any difference whether one is interested in the reconstruction of words or syntactic structures – both must be compartmentalized. Words consist of synchronically arbitrarily concatenated phonemes, and these phonemes in turn are bundles of distinctive phonological features. These phonological bundles – phonemes – are the central point for an application of the Comparative Method. Previous approaches to syntactic reconstruction either reflect an insufficient understanding of the working procedure of linguistic reconstruction or of modern syntactic theories, and the present con-



tribution tried to overcome both of these hurdles. My study aimed at a demonstration of the feasibility of a solid, rigorous syntactic reconstruction meeting the standards of Indo-European linguistics.

The so far unresolved question of a syntactic analogue to phonemes was the greatest obstacle for an application of the Comparative Method to syntactic research objects. I proposed to regard syntactic parameters, or rather thematically grouped bundles of such parameters, as ‘syntactic phonemes’. Syntactic parameters differ from phonemes and phonological features in one important respect at least: they are not shaped by the physical properties of the human vocal tract. Instead, they are formal, logically motivated choices which derive their existence from some ‘base-generated’ proto-utterance and general categories like animacy, number, gender, person etc. which – as seems likely given the results of the psycholinguistic and typological research of the last decades – appear to be cognitive concepts common to and shared by all healthy humans. If all humans share these basic cognitive categories, a system of logical ramifications of ever finer syntactic choices starting from these basic concepts may be regarded as a syntactic equivalent to the sounds the human vocal tract can produce.

Parameters are by definition discrete in nature, and because they form a system of abstract logical choices deriving from some basic cognitive mechanisms and categories common to all healthy members of our species, they are also finite in number and cross-linguistically comparable due to their universal applicability or prevalence. The parallelism to phonological features – and bundles of phonological features forming phonemes – can be seen in the overall finite number of syntactic parameters and the fact that both parameters and phonological features are characterized by their discrete nature.

This discrete nature of parameters shines through in minimal pairs of syntactic structures differing only in one parametric value. In phonology, it is common practice to work out or rather define the different phonemes of a particular language by means of minimal pairs that differ only in one phonological aspect (voiced : voiceless, breathy voice : creaky voice, etc.). If a difference in

just one parametric value or setting allows for the differentiation of syntactic structures, this may justifiably lead us to the conclusion that ‘syntactic phonemes’ are nothing else than bundles of thematically grouped parameters. Phonemes can be differentiated according to their particular composition of distinctive phonological features, and ‘syntactic phonemes’ or rather ‘syntaxemes’ may parallel this with different parametric configurations resulting in different ‘syntaxemes’. From a synchronic point of view, individual languages use only an arbitrary subset of all theoretically possible phonemes, but once the diachronic dimension is taken into account, the seemingly chaotic composition turns out to be the result of regular prehistoric processes. I contend that the same pertains to parametric bundles.

If this basic parallelism between phonemes and bundles of parameters is accepted, it may follow that several ‘syntaxemes’ form larger syntactic structures, structurally comparable to the fact that several concatenated phonemes form syllables or words. It is not evident why syntactic parameters and the specific composition of parameters should be different in that respect. In other words, bundles of parameters may be the abstract building blocks of syntax, like phonemes in the case of roots, morphemes or whole words.

If that is true, the process of syntactic reconstruction can be expected to resemble phonological reconstruction: At first, one compares genetically related languages in terms of their parametric configurations, determining the individual parametric settings ( $\sim$  phonemes). In a second step, one reconstructs the values of their last common ancestor, similar to a reconstruction of a phoneme inventory. Afterwards one can trace particular changes to individual daughters and, lastly, search for special developments in certain contexts. If such are indeed found, together with an identification of the conditions that cause aberrant developments of parametric values, they may force us to modify the initial generalizations of diachronic developments from the proto-language to its daughter(s) and to refine the reconstructions in a heuristic process.

There is no difference in terms of structural properties between syntactic,

parameter-based reconstruction and phonological reconstruction. Both aim at an elaboration of the specific constitution of basic units – a phoneme inventory in terms of the basic phonological units (phonemes) and a parametric feature matrix in terms of the basic syntactic units (parameters). Likewise, both parameter-based syntactic reconstruction and phoneme-based phonological reconstruction try to uncover regular developments, general and context-sensitive, of the respective basic units of both domains from a proto-language to its daughters. The current dogma in IE linguistics is that

“[t]he goal of IE studies is not the reconstruction of utterances, but that of linguistic competence. The reconstructed roots, words, or affixes are entries in the mental lexicon of an ideal PIE speaker, the phonological or morphological rules for manipulating them part of his grammar. Likewise, PIE syntax is not concerned with actual strings, but with the structure of complex syntactic objects and constraints on the wellformedness of such objects.” (Keydana 2018, p. 2196)

This assertion is partly correct and partly incorrect. The “linguistic competence” is the set of abstract rules relevant for a particular language. If we want to focus our attention on this, we will need rules that can be formulated in such a way that they can be compared across language families, ideally in form of lists or matrices. Parameters are a reasonable approach to formalize syntactic rules, facilitating an objective comparison of the syntactic behavior of genetically related and unrelated languages. Parameters can thus be used for both reconstructive and typological-comparative purposes.

It is, on the other hand, not evident why it should be impossible to use parameters also for reconstructions of very concrete syntactic structures. If it is possible to reconstruct the syntactic equivalent of a phoneme inventory, and if the regular developments to a set of genetically related languages are known, one can analyze a larger syntactic structure in a particular language with regard to its parametric constitution and configuration and then make an educated guess – or even a prediction – on its configurational type in the other languages. The

search for cognates will be an easy task once the full set of parameters, i.e., nominal and verbal macro- and mesoparametric ones, will be known.

In theory, one could differentiate between two types of syntactic cognates. The first one would consist of collocations, phrases or clauses built upon and with lexically identical material. This interpretation would resemble that of previous approaches to syntactic reconstruction, e.g., those focusing on collocations (the traditional IE approach), but also those rooted in (Radical) Construction Grammar (see most recently Barðdal and Eythórsson 2020; Barðdal et al. 2020; Gildea et al. 2020 etc.) and the generative one of Walkden (2014).

The second variant of syntactic cognates would consist of abstract syntactic structures, such as distinct types of NPs (noun + adjective, noun + demonstrative + adjective; noun + demonstrative + cardinal + adjective + genitive etc.). The lexical material would then be irrelevant, only the abstract structure (which parameters are involved?) would be compared in terms of the parametric configuration (which values?) in languages X, Y, Z.

As long as one restricts the parametric comparison and reconstruction to macro- and mesoparameters (as was done in this dissertation), this differentiation does not make much sense, as macro- or mesoparameters determine the syntactic behavior of large parts of the language's lexicon (all probes, all phase heads, all heads of a given natural class such as nouns). If all compared parameters have such a great scope, lexical identity of concrete compared phrases will not be mandatory, as all phrases of the same abstract type will have the same parametric values. The only additional benefit of lexically and structurally identical phrases is an aesthetic one. Even with microparameters lexical identity should not in all cases be necessary; only nanoparameters require the comparison of identical lexical material. My proposed solution to the 'correspondence problem' in syntax is therefore two-fold: abstract types of larger structures consisting of the same constituents suffice as long as one compares and reconstructs rather general properties; lexically identical or corresponding structures (phrases) will

be necessary, if one is interested in nanoparametric values.

Walkden (2014) took as his starting point a very strict reading of the Borer-Chomsky Conjecture, leading him to the conclusion that “syntactic reconstruction is lexical reconstruction” (l.c., p. 113). I do not think that this is necessarily true; it depends on which syntactic features one wants to reconstruct – those with scope over a small part of the lexicon or those with scope over a great or even the greatest part of it. If we should always restrict concrete reconstructions to cases of lexical identity, the point of arbitrariness – rarely addressed by proponents of syntactic reconstruction (but see Daniels 2021 for an exception, particularly p. 409–411) – would be grossly ignored. Syntactic comparanda must be units that can be arranged in a synchronically arbitrary way, similar to phonemes.

As to the causes or mechanisms of syntactic change, this study took an agnostic point of view. Considering the fact that the reconstructions of phonemes and concrete lexicalized strings of phonemes (words, morphemes) are independent of the possible intrinsic (or extrinsic) factors responsible for changes from a proto-language to individual daughters, it is not relevant *why* and *when* syntactic structures change over time. Whether this happens during language acquisition (I-language) by means of reanalysis or later in life by means of interferences between different I-languages (different registers, language contact etc.) is irrelevant. Language change does happen, and so does syntactic change. Speculations about causes will be unnecessary, if one just aims at the reconstruction of former parametric settings. It is only necessary to give an answer to the questions of *whether or not* and *how* the syntactic behavior of a proto-language differed from that of its daughters, and this study relied on the time-tested Comparative Method and a set of neutral questions, parameters, to approach this problem.

Noun Phrases can be classified according to many other, much more fine-grained parametric differences than covered in this dissertation. My study should be regarded as a first attempt to demonstrate the potential of parametric ana-

lyses for reconstructive purposes, and I hope that I have been able to show this. Differing from other approaches to syntactic reconstruction, parametric studies are falsifiable and reproducible and thus resemble classical reconstructions of sequences of phonemes. A refined and exhaustive analysis of the parametric settings of all ancient Indo-European languages will enable us to understand how the syntactic rules of PIE and its daughter languages mutually interacted to yield seemingly free surface orders, besides or even irrespective of information structure. It is my firm conviction that a determination of PIE as a configurational, non-configurational or discourse-configurational language depends on fine-grained parametric analyses. Going far beyond a mere description of the syntactic surface patterns of individual daughters, the ability to analyze the syntactic behavior of languages in parametric terms will be a major advancement in our attempts to reconstruct the Proto-Indo-European language – we may soon have the ability to predict the Proto-Indo-European shape of syntactic structures based on an analysis of the configuration of cognate structures found in one or more daughter languages.

To conclude this book, a plausible answer to the question of what to compare in the sense of Watkins' 'first law of comparative grammar' are logical choices defining the syntactic behavior of languages. Parameters are the comparanda of syntax, and specific strings or arrays of individual parametric feature bundles – syntaxemes – its cognates.



# Appendix

## 5.1 Parametric Feature Matrix

Feature matrix 5.1 allows for a comparison of premodern Iranian NP properties with those of selected other Indo-European languages (Vedic, Greek, Latin, but also Modern Persian and Pashto). The matrix lists syntactic parameters in the form of parallel columns, similar to lists or ‘grids’ of distinctive phonological features. This database can be used for comparative purposes on an abstract, quasi phonological level (as if one were to compare the phonological inventories of languages, only with a syntactic module: NPs).

The parameters, their abbreviations, and the logical relationships are based on Longobardi et al. (2013), though with minor changes (I both omitted and added a few parameters). All data were collected by myself – Farsi and Pashto were already covered by Longobardi et al. (2013) and Ceolin et al. (2020), but I checked the settings (I gave several examples of parametric values in Chapter 2). Abbreviations: OAv. = Old Avestan; YAv. = Young Avestan; OP. = Old Persian; Kho. = Khotanese; Sgd. = Sogdian; Cho. = Chorasmian; Bct. = Bactrian; Pth. = Parthian; MiP. = Middle Persian; PIr. = Proto-(Indo-)Iranian; Ved. = Vedic; Lat. = (Old) Latin; Gk. = (Archaic) Greek; Far. = Farsi; Psh. = Pashto.

Parametric values: + signals positive setting; – negative setting; 0 neutralization or irrelevance due to some other parameter; ? uncertainty or lack of unambiguous data; brackets () indicate predictability or a parametric setting not 100% matching the definition.



	Title/Description	Precondition(s)	Abbrev.	OA	Yav.	OP.	Kho.	Sgd.	Cho.	Bet.	Pth.	MIr.	Plr.	Ved.	Gk.	Lat.	Far.	Psh.
1	Gramm. Person		FGP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2	Gramm. Number	FGP must be +	FGN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3	Gramm. Gender	FGN +	FGG	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
4	Feature spread to N	FGN +	FSN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
5	number on N (bounded nouns)	FSN +	FNN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	Gramm. partial definiteness	DGP +	DGP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Gramm. definiteness	DGP +	DGR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Strong Person	FGP +, DGR +, NDE -	NSD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	Free null partitive Q	FNN +	DPQ	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	Gramm. distal art.	FSN - or FNN - or DGR +	DDA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Def-checking N	DGR +	DCN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	Def. spread to N	DCN +, NSD -	DSN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	Def. on relatives	DGR +	DOR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	D-controlled infl. on N	FSN +	DIN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15	Plural spread from Cardinals	FSN +	CPS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16	Gramm. boundedness		CGB	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17	Strong article	DGR +, FNN +, CGB -	CGR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	Bounded-checking N	CGB +	CCN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	null-N-licensing article	FSN- or FNN - or DCN - or NSD +	DNN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	Structured Adjective Phrases		AST	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
21	Feature spread to structured APs	FSN +, AST +	FIS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
22	Feature spread to predicative APs	FGN +	FSP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
23	D-controlled infl. on adjectives	NSD -, FIS +	ADI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	DP over relatives		ADR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
25	RelCl. extrap.	ADR -	AER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	Free reduced RelCl	ADT +	ARR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
27	Adpositional Genitive		GAD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Free Gen.		GFR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
29	Uniform Gen.		GUN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
30	DP over free Gen.	GFR +, ADR +	GPR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
31	GenO		GFO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	Gen-feature spread to N		GFS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
33	D-checking possessives		PDC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	Adjectival poss.	DGR +, NSD + or CGR not +	APQ	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
35	Post-affix poss.		PAP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	Clitic poss.		PCI	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
37	N-feature spread to pron. poss.	FIS + or AST -, PAP + or PCI +	PHS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	N-feature spread to free Gen.	FIS +, GFR +, PHS must not be -	GSP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
39	Adjectival Gen.	APQ +	AGE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
40	Poss.-checking N		GCN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
41	Strong partial locality	FSN - or FNN +, CGR must not be +	TPL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
42	Strong locality	TPL must not be -	TSI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	D-checking demonstratives	FSN - or DGR +, TPL must not be -	TDC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	N over Demonstratives	FGP +	NDE	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
45	N over Cardinals		NOC	?	-	-	-	-	-	-	-	-	-	-	-	-	-	-
46	N over Ordinals	NOC -	NOO	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
47	N over adjectives	NOO - or NGS -, NPP -	NOA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
48	N over GenO	GFO must not be -, NOA - or AST -	NGO	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
49	N over external arguments	NGO - or (GFO -, NOA - or AST -)	NOE	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
50	Definiteness on APs	DGP +, postnominal APs	DOA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	Verbless relative clauses	FLI -	VRC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	Facultative linker		FLI	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
53	Generalized linker		GLI	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5.1: Parametric Feature Matrix: nominal macro- and mesoparameters of Old and Middle Iranian languages compared with those of other early attested IE languages.

## 5.2 Phonological Feature Matrix

As stated in Chapter 4, parameters might turn out to be the syntactic equivalent of distinctive phonological features. Consider the phonological features chart on the next page.<sup>1</sup> The parallelism to a macro- and mesoparametric ‘grid’ is obvious. Different configurations of phonological features constitute certain phonemes, and the same may pertain to parameters on the syntactic level.

---

<sup>1</sup>Downloaded from <https://www.artoflanguageinvention.com/papers/features.pdf>, last accessed on 8-14-2023.



## 5.3 Farsi

Farsi and Pashto were covered in Longobardi et al. (2013) and Ceolin et al. (2020), respectively. I familiarized myself with the NP parameters and their properties by means of a check of the parametric settings ascribed by Longobardi and colleagues to Farsi and Pashto. I tried to understand how the parameters worked and I wanted to find examples of concrete NPs showing individual settings. Neither Longobardi et al. (2013) nor Ceolin et al. (2020) present any examples demonstrating their parametric settings, so that the following pages may fulfill this function, together with the remarks in Chapter 2.

Modern Persian is a member and representative of the southwestern Iranian group of languages. Due to its millennia-long status as *lingua franca* of the Iranian world, Persian is today spoken in three variants, namely Farsi (Iran), Dari (Afghanistan), and Tajik or Tajiki (Tajikistan). I will concentrate here only on Farsi, because in general the three varieties are very close to each other and Farsi is undoubtedly best known and best described of all three, but note that particularly Tajik(i) has a few typical eastern features (triple deixis, plural marking on nouns after cardinals) as an obvious contact phenomenon with eastern Iranian languages.

In many cases, I presented examples of individual parametric settings in Chapter 2 of this book. I will not repeat glossed examples here and instead focus on additional information not presented in Chapter 2. The discussion is not as exhaustive as the sections on premodern Iranian languages and the three Indo-European relatives of Chapter 3, but it may nevertheless serve its function to back up the values given in the parametric feature matrix.

1. Gramm. Person (FGP): set to yes. Farsi distinguishes person in the verbal domain and has a two-way contrast in anaphors (*in* ‘this’ and *ān* ‘that’)<sup>2</sup> and a three-way contrast in terms of personal pronouns, cf. the respective entries/chapters in Paul (2019), Windfuhr and Perry (2009) or Lazard

---

<sup>2</sup>Tajiki has a three-way deictic contrast between *in* ‘this’, *on* ‘that’ and *vay* ‘that’.

(1992).

2. Gramm. Number (FGN): yes. Farsi distinguishes between singular and plural and marks plural by means of suffixes on the noun (generalized plural marker *-hā* and a specific animate plural marker *-ān*). Cf. e.g., Windfuhr and Perry (2009, p. 483): *deh-hā xeyli rāh nistand* ‘The villages are not far.’
3. Gramm. Gender (FGG): no. Farsi has no gender distinctions.
4. Feature spread to N (FSN): yes. Cf. parameter FGN; the feature number is expressed on the noun.
5. Number on N (bounded nouns) (FNN): yes. Cf. FNN or FSN.
6. Gramm. partial definiteness (DGP): no. This is an extremely influential parameter, a negative value neutralizes 14 other parameters. As discussed in Chapter 2, Farsi has no *obligatory* definiteness marking of nouns or nominal arguments, only a facultative one. Hence this parameter gets a negative value, and logically and hierarchically dependent parameters – DGR, NSD, DDA, DCN, DSN, DOR, CGR, DNN, ADI, PDC, PAP, PHS, TDC and DOA – are neutralized.
7. Free null partitive Q (DPQ): no. There are no count/mass distinctions by means of differential case marking as in Finnish.
8. D-controlled infl. on N (DIN): no. Despite being in contact with Semitic languages for two and a half thousand years, there is nothing similar to the Arabic/Semitic nunation in Persian, i.e., inflectional forms of the noun do **not** depend on the occurrence of certain morphemes in the determiner slot beyond phi-feature concord.
9. Plural spread from Cardinals (CPS): no. Cf. Windfuhr and Perry (2009, p. 537): *in dah ruz-rā* ‘these ten days [direct object]’, without explicit plural morpheme on the noun. Note that this is one of the differences between

Farsi and Tajiki, as Tajiki has the option to mark nouns for plural if the NP contains a cardinal numeral.

10. Grammaticalized boundedness (CGB): set to yes. A bare count singular noun has an unbounded reading in Farsi, and an overt determiner is needed to get a bounded indefinite denotation. I gave an illustrating example in Chapter 2, viz. *gorg koštand* ‘they killed wolves, were wolf-killing’ vs. *(yek) gorg-i koštand* ‘they killed some [or: one] wolf.’
11. Bounded-checking N (CCN): yes. The noun incorporates a marker for bounded indefinite readings as an enclitic affix (Farsi *-i*).
12. Structured Adjective Phrases (AST): yes. Adjectives occur within the NP/DP core area, i.e., they are ‘structured’, cf. *so’āl-hā-ye xub-i* ‘some/certain good questions’ (Windfuhr and Perry 2009, p. 471), with noun, plural marker, *eżāfe* marker, adjective and the clitic marker signaling a bounded indefinite reading of the whole NP.
13. Feature spread to structured APs (FFS): no. Structured adjectives are not inflected in Farsi, in stark contrast to other IE languages.
14. Feature spread to predicative APs (FSP): no. Predicative adjectives are likewise not inflected, cf. e.g., Windfuhr and Perry (2009, p. 487): *man gorosne (am)* ‘I am hungry.’
15. DP over relatives (ADR): yes. Relative clauses consistently surface to the right of all other NP arguments and modifiers, irrespective of relative clause-internal markedness or variation as expressed by the placement of the relative clause before or after the matrix clause predicate. Cf. Windfuhr and Perry (2009, pp. 504–505):

## (66) Relative clauses in Farsi

## a. Unmarked

*ān doxtar-i [ke Ali-rā dust dār-ad] raft*  
 that girl-IDF REL Ali-DO friend have.PRES-3SG go.PST.3SG  
 ‘(that) the girl whom Ali loves left.’

## b. Contrastive

*ān doxtar-i raft [ke Ali-rā dust dār-ad]*  
 that girl-IDF go.PST.3SG REL Ali-DO friend have.PRES-3SG  
 ‘(that) the girl whom Ali loves left [not the other one].’<sup>3</sup>

This neutralizes parameter AER.

16. Free reduced RelCl (ARR): no. Farsi has no adjectives as free reduced relative clauses.
17. Adpositional genitive (GAD): no. Genitives in Farsi always entail an *eḏāfe* construction.
18. Free Gen. (GFR): yes, recursive genitives are possible, and they are formed with an *eḏāfe* chain, cf. e.g., Windfuhr and Perry (2009, p. 536) *mardom-e mataq-e Rebāt-e Morād-e Xomeyn* people-EZ region-EZ Rebāt-EZ Morād-EZ Xomeyn ‘the people of the region of Rebāt-e Morād in the Xomeyn district’, or (l.c.) *tamām-e ahl-e mahall* whole-EZ people-EZ place ‘all people of the village’.
19. Uniform Gen. (GUN): no.
20. DP over free Gen. (GPR): yes. The *eḏāfe* construction by which free recursive genitives are expressed has the consistent right-branching order of N-EZ-Gen-Ez-Gen-EZ-Gen etc. (e.g., see parameter GFR above), hence genitives occur to the right of NPs/DPs.

<sup>3</sup>The translations of these two examples are those given in Windfuhr and Perry (2009, 504f.). It does not affect the argument of consistent relative clause placement to the right of all NP arguments, but in my view, the translations are wrong. Even though the – optional – relative clause subject marker *u* is missing, Ali is marked as direct object, hence I read the whole complex as ‘The girl who loves Ali left’?

21. GenO (GFO): yes. Cf. e.g., *kār-e pesar* ‘the work of the boy’.
22. Gen-feature spread to N (GFS): no. There is no marking on N.
23. Adjectival poss. (APO): no. There are no adjectival possessives in Farsi. This neutralizes parameter AGE.
24. Clitic poss. (PCL): yes. Cf. *Hasan ketāb-aš* (lit. ‘Hasan book-his’) ‘Hasan’s book’.
25. N-feature spread to free Gen. (GSP): irrelevant due to negative setting of FFS.
26. Poss.-checking N (GCN): no.
27. Strong partial locality (TPL): yes. Deictic demonstratives are attracted to the D-area, i.e., they surface in determiner position. Cf. e.g., *ān mard* ‘that man’.
28. Strong locality (TSL): yes. All demonstratives moved to the D-area.
29. N over Demonstratives (NDE): no. See the discussion in Chapter 2 on N-raising in Farsi.
30. N over Cardinals (NOC): no.
31. N over Ordinals (NOO): yes. Cf. *sāl-e sevom* ‘third year/grade’.
32. N over adjectives (NOA): irrelevant due to setting of NOO.
33. N over GenO (NGO): irrelevant due to setting of NOO.
34. N over external arguments (NOE): irrelevant due to setting of NOO.
35. Definiteness on APs (DOA): irrelevant due to setting of DGP.
36. Verbless relative clauses (VRC): irrelevant due to positive value of GLI.



37. Facultative linker (FLI): irrelevant due to positive value of GLI.
38. Generalized linker (GLI): yes. Cf. the examples above, e.g., parameter GFR. The generalized nature of the linker becomes clear from examples like Windfuhr and Perry (2009, p. 471) *pesar-e az madrase gorixte* ‘boy (which) fled from school (i.e., truant boy)’.

## 5.4 Pashto

Pashto, spoken in Afghanistan, is probably the best known modern eastern Iranian language. The best description of syntactic phenomena is Robson and Tegey (2009), but Tegey and Robson (1996) and Skjærvø (1989b) occasionally yield additional pieces of information.

1. Gramm. Person (FGP): set to yes. Pashto has a triple-deixis system (near *dagh-*, mid *agh-*, far *ugh-*) as well as person in the verbal domain.
2. Gramm. Number (FGN): set to yes. Pashto distinguishes sg. and pl. and additionally has a so-called numerative used with numerals higher than ‘one’ in the direct case (etymologically the old dual). See e.g., Robson and Tegey (2009, p. 726) or Skjærvø (1989a, p. 371).
3. Gramm. Gender (FGG): set to yes. Pashto systematically distinguishes between masculine and feminine gender and marks this via inflectional endings (e.g., *sp-ay* ‘dog’ vs. *sp-əy* ‘bitch’).
4. Feature spread to N (FSN): set to yes. Number marking occurs on the noun (or the feature is ‘spread’ in generative terminology). Cf. Robson and Tegey (2009, pp. 754–755): *de asad agha loy kitâbuna* ‘those big books of Asad’.

5. Number on N (bounded nouns) (FNN): set to yes. See FSN.
6. Gramm. partial definiteness (DGP): set to no, as discussed in Chapter 2. This neutralizes 14 other parameters.
7. Free null partitive Q (DPQ): no, Pashto has no count/mass distinctions by means of differential case marking.
8. D-controlled infl. on N (DIN): no. Nothing similar to the nunation of Arabic.
9. Plural spread from Cardinals (CPS): yes, cf. Robson and Tegey (2009, p. 769): *tsalor wâr-a pəx-e ye* ‘all four of his legs’.
10. Grammaticalized boundedness (CGB): In contrast to Farsi, the reading of a singular-inflected noun as unbounded is not possible in Pashto. I gave an example in Chapter 2. This neutralizes parameter CCN.
11. Structured Adjective Phrases (AST): yes, adjectives occupy structured positions in a DP/NP, cf. Robson and Tegey (2009, pp. 769–770): *yaw ghaṭ wux* ‘a (lit. one) huge camel’.
12. Feature spread to structured APs (FFS): yes, see AST. Another example can be seen in Robson and Tegey (2009, p. 758):

(67) Pashto, feature spread to structured adjectives  
*xkwəl-e njəl-əy aw xədz-a*  
 pretty-PL.F girl-SG.F and woman-SG.F  
 “pretty girl and woman”

13. Feature spread to predicative APs (FSP): yes, cf. Robson and Tegey (2009, p. 759):

(68) Pashto, feature spread to predicative adjectives  
*nə âshə x-a da, nə gwəla*  
 not Asha good-SG.F be.PRES.IPFV.3SG not Gwela

“neither Asha nor Gwela is good”.

14. DP over relatives (ADR): yes, relative clauses follow their head noun, see Tegey and Robson (1996, 206ff.). However, in the more recent treatise Robson and Tegey (2009, p. 759) it is stated that relative clauses “usually” follow their head nouns, implying that they can also precede, but no examples are presented. This neutralizes parameter AER.
15. Free reduced RelCl (ARR): no, Pashto does not have this option.
16. Adpositional genitive (GAD): yes, cf. Robson and Tegey (2009, p. 755): *de asad de plâr la tsaloro dero xâysta luño sara* ‘with Asad’s father’s four very pretty daughters’.
17. Free Gen. (GFR): yes, see the example in GAD, directly above.
18. Uniform Gen. (GUN): no.
19. DP over free Gen. (GPR): no, free genitives always precede the DP/NP core.
20. GenO (GFO): no, Pashto has no structured (i.e, NP-internal) inflectional object genitive, only a free and adpositional one.
21. Gen-feature spread to N (GFS): no.
22. Adjectival poss. (APO): no. This neutralizes parameter AGE.
23. Clitic poss. (PCL): yes, Pashto has clitic possessives, cf. Robson and Tegey (2009, p. 733): *kitâb me, kitâb ye* ‘my book, his book’.
24. N-feature spread to free Gen. (GSP): no, this does not seem to be the case. Each new free genitive has its own features, not the phi-features of its head noun.
25. Poss.-checking N (GCN): no.

26. Strong partial locality (TPL): yes, deictic demonstratives are attracted to the D-area, see e.g., Tegey and Robson (1996, p. 170):

(69) Pashto, demonstrative in D-area  
*agha spin moṭar de asád nəway moṭár day*  
 that white car of Asad new car is  
 “That white car is Asad’s new car”.

27. Strong locality (TSL): yes, uniform system. All demonstratives surface in the D-area (Tegey and Robson 1996, p. 171).
28. D-checking demonstratives (TDC): neutralized/irrelevant due to negative setting of DGP.
29. N over Demonstratives (NDE): no. N-raising in Pashto has been discussed in Chapter 2.
30. N over Cardinals (NOC): no.
31. N over Ordinals (NOO): no.
32. N over adjectives (NOA): no.
33. N over GenO (NGO): no.
34. N over external arguments (NOE): no. All nominal arguments precede the noun. Pashto is a rigidly head-final language.
35. Verbless relative clauses (EZ1): no, Pashto relative clauses apparently always have an overt predicate.
36. Facultative linker (FLI): no.
37. Generalized linker (GLI): no.



# Bibliography

- Abaev, Vasilij I (1964). *A Grammatical Sketch of Ossetic [Grammatičeskij očerk osetinskogo jazyka]*. Edited by Herbert H. Paper. Translated by Steven P. Hill. Indiana University Research Center in Anthropology, Folklore and Linguistics, Publ. 35. The Hague: Mouton.
- Abney, Stephen Paul (1987). "The English Noun Phrase in its Sentential Aspect". PhD thesis. Massachusetts Institute of Technology. URL: <http://www.ai.mit.edu/projects/dm/theses/abney87.pdf>.
- Adams, James N. (1976). "A typological approach to Latin word order". In: *Indogermanische Forschungen* 81, pp. 70–99. doi: <https://doi.org/10.1515/9783110243239.70>.
- Adger, David (2003). *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Ahn, Byron (2016). "Syntax-phonology mapping and the Tongan DP". In: *Glossa: A Journal of General Linguistics* 1.1. doi: <http://doi.org/10.5334/gjgl.39>.
- Aikhenvald, Alexandra Y. (2006). "Grammars in contact: A cross-linguistic perspective". In: *Grammars in contact: A cross-linguistic typology*. Ed. by Alexandra Y. Aikhenvald and Robert M. W. Dixon. Oxford: Oxford University Press, pp. 1–66.
- Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou (2007). *Noun Phrase in the Generative Perspective*. Studies in Generative Grammar 71. Berlin/New York: Mouton de Gruyter. doi: <https://doi.org/10.1515/9783110207491>.
- Alfieri, Luca (2011). "A Radical Construction Grammar Approach To Vedic Adjective". In: *Rivista degli studi orientali, Nuova Serie* 84.1, pp. 241–255. URL: <https://www.jstor.org/stable/43927270>.
- Allan, Rutger J. (2012). "Clause Intertwining and Word Order in Ancient Greek". In: *Journal of Greek Linguistics* 12.1, pp. 5–28. doi: <https://doi.org/10.1163/156658412X649733>.
- (2014). "Changing the Topic: Topic Position in Ancient Greek Word Order". In: *Mnemosyne* 67.2, pp. 181–213. doi: <https://doi.org/10.1163/1568525X-12341050>.

- Allen, William Sidney (1953). "Relationship in comparative linguistics". In: *Transactions of the Philological Society* 52.1, pp. 52–108.
- Allentoft, M., M. Sikora, and K. Sjögren et al. (2015). "Population genomics of Bronze Age Eurasia". In: *Nature* 522, pp. 167–172. DOI: <https://doi.org/10.1038/nature14507>.
- Anderson, John (1993). "Parameters of syntactic change: a notional view". In: *Historical Linguistics: Problems and Perspectives*. Ed. by Charles Jones. London: Routledge, pp. 1–42. DOI: <https://doi.org/10.4324/9781315845814>.
- Anthony, David W. (1995). "Horse, wagon & chariot: Indo-European languages and archaeology". In: *Antiquity* 69.264, pp. 554–565. DOI: [10.1017/S0003598X00081941](https://doi.org/10.1017/S0003598X00081941).
- (2007). *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes shaped the Modern World*. Princeton, Oxford: Princeton University Press.
  - (2013). "Two IE phylogenies, three PIE migrations, and four kinds of steppe pastoralism". In: *Journal of Language Relationship* 9, pp. 1–21. URL: [https://www.jolr.ru/files/\(104\)jlr2013-9\(1-21\).pdf](https://www.jolr.ru/files/(104)jlr2013-9(1-21).pdf).
  - (2019). "Archaeology, Genetics, and Language in the Steppes: A Comment on Bomhard". In: *Journal of Indo-European Studies* 47.1 & 2, pp. 175–198.
- Anthony, David W. and Donald A. Ringe (2015). "The Indo-European Homeland from Linguistic and Archaeological Perspectives". In: *Annual Review of Linguistics* 1, pp. 199–219. DOI: <https://doi.org/10.1146/annurev-linguist-030514-124812>.
- Azami, Cheragh Ali and Gernot Windfuhr (1972). *A Dictionary of Sangesari. With a Grammatical Outline*. Tehrān: Serkat-e Sahāmi-ye Ketābhā-ye Jibi.
- Baker, Mark C. (1996). *The Polysynthesis Parameter*. Oxford Studies in Comparative Syntax. Oxford, New York: Oxford University Press.
- (2001). *The Atoms of Language. The Mind's Hidden Rules of Grammar*. New York: Basic Books.
  - (2003). *Lexical categories. Verbs, Nouns, and Adjectives*. Cambridge Studies in Linguistics 102. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511615047>.
  - (2008a). "The macroparameter in a microparametric world". In: *The Limits of Syntactic Variation*. Ed. by Theresa Biberauer. Amsterdam: John Benjamins, pp. 351–373. DOI: <https://doi.org/10.1075/la.132.16bak>.
  - (2008b). *The Syntax of Agreement and Concord*. Cambridge Studies in Linguistics 115. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511619830>.
- Baker, Mark C. and Nadya Vinokurova (2009). "On Agent Nominalizations and Why They Are Not like Event Nominalizations". In: *Language* 85.3, pp. 517–556. URL: <http://www.jstor.org/stable/40492892>.

- Balles, Irene (2008). "Principles of syntactic reconstruction and 'morphology as paleosyntax': The case of some Indo-European secondary verbal formations". In: *Principles of Syntactic Reconstruction*. Ed. by Gisella Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 161–186. doi: <https://doi.org/10.1075/cilt.302>.
- Barðdal, Jóhanna (2013). "Construction-Based Historical-Comparative Reconstruction". In: *The Oxford Handbook of Construction Grammar*. Ed. by Thomas Hoffmann and Graeme Trousdale. Oxford: Oxford University Press, pp. 438–457. doi: <https://doi.org/10.1093/oxfordhb/9780195396683.013.0024>.
- (2014). "Syntax and syntactic reconstruction". In: *The Routledge Handbook of Historical Linguistics*. Ed. by Claire Bowerman and Bethwyn Evans. London: Routledge, pp. 343–373. doi: <https://doi.org/10.4324/9781315794013>.
- Barðdal, Jóhanna, Valgerður Bjarnadóttir, Serena Danesi, Tonya Kim Dewey, Thórhallur Eythórsson, Chiara Fedriani, and Thomas Smitherman (2013). "The Story of 'Woe'". In: *Journal of Indo-European Studies* 41.3/4, pp. 321–377.
- Barðdal, Jóhanna and Thórhallur Eythórsson (2012). "Reconstructing Syntax: Construction Grammar and the Comparative Method". In: *Sign-Based Construction Grammar*. Ed. by Hans C. Boas and Ivan A. Sag. Stanford: CSLI Publications, pp. 257–308.
- (2020). "How to Identify Cognates in Syntax? Taking Watkins' Legacy One Step Further". In: *Reconstructing Syntax*. Ed. by Jóhanna Barðdal, Spike Gildea, and Eugenio R. Lujan. Leiden: Brill, pp. 197–238. url: [https://doi.org/10.1163/9789004392007\\_006](https://doi.org/10.1163/9789004392007_006).
- Barðdal, Jóhanna and Spike Gildea (2015). "Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications". In: *Diachronic Construction Grammar*. Ed. by Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer, and Spike Gildea. Constructional Approaches to Language 18. Amsterdam: John Benjamins, pp. 1–50. doi: <https://doi.org/10.1075/cal.18.01bar>.
- Barðdal, Jóhanna, Spike Gildea, and Eugenio R. Lujan, eds. (2020). *Reconstructing Syntax*. Leiden: Brill. doi: <https://doi.org/10.1163/9789004392007>.
- Barðdal, Jóhanna and Thomas Smitherman (2013). "The Quest for Cognates: A Reconstruction of Oblique Subject Constructions in Proto-Indo-European". In: *Language Dynamics and Change* 3.1, pp. 28–67. doi: <https://doi.org/10.1163/22105832-13030101>.
- Barros Damgaard, Peter de, Rui Martiniano, and Jack Kamm et al. (2018). "The first horse herders and the impact of early Bronze Age steppe expansions into Asia". In: *Science* 360.6396. doi: [10.1126/science.aar7711](https://doi.org/10.1126/science.aar7711).
- Bartoněk, Antonín (2003). *Handbuch des mykenischen Griechisch*. Heidelberg: Universitätsverlag C. Winter.
- Bashir, Elena (2009). "Wakhi". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 825–862.



- Batlloori, Montserrat, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca, eds. (2005). *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press.
- Bauer, Anna H. (2014). *The Morphosyntax of the Noun Phrase in Hieroglyphic Luwian*. Brill's Studies in Indo-European Languages and Linguistics, Vol. 12. Leiden: Brill. doi: <https://doi.org/10.1163/9789004260030>.
- Beekes, Robert Stephen Paul (2011). *Comparative Indo-European Linguistics*. (2nd revised ed. by Michiel de Vaan). Amsterdam: John Benjamins.
- Begby, Endre (2017). "Language from the Ground Up: A Study of Homesign Communication". In: *Erkenntnis* 82, pp. 693–714. doi: <https://doi.org/10.1007/s10670-016-9839-1>.
- Behaghel, Otto (1909). "Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern". In: *Indogermanische Forschungen* 25, pp. 110–142. doi: <https://doi.org/10.1515/9783110242652.110>.
- Benedetti, Giacomo (2023). "The Meaning and Etymology of *ārya*". In: *Bhasha. Journal of South Asian Linguistics, Philology and Grammatical Traditions* 2.1, pp. 123–142. doi: [10.30687/bhasha/2785-5953/2023/01/006](https://doi.org/10.30687/bhasha/2785-5953/2023/01/006).
- Benveniste, Émile (1946). *Vessantara-Jātaka: Texte Sogdien, édité, traduit et commenté par É. Benveniste*. Paris: Geuthner.
- Benzing, Johannes (1968). *Das chwaresmische Sprachmaterial einer Handschrift der "Muqaddimat al-Adab" von Zamaxšarī. I. Text*. Wiesbaden: Franz Steiner Verlag.
- (1983). *Chwaresmischer Wortindex. Mit einer Einleitung von Helmut Humbach. Herausgegeben von Zahra Taraf*. Wiesbaden: Harrassowitz.
- Bernstein, Judy B. (1991). "DPs in French and Walloon: Evidence for Parametric Variation in Nominal Head Movement". In: *Probus* 3.2, pp. 101–126. URL: <https://scihub.se/10.1515/prbs.1991.3.2.101>.
- (2001). "The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain". In: *The Handbook of Contemporary Syntactic Theory*. Ed. by Mark Baltin and Chris Collins. Oxford: Blackwell, pp. 536–561. doi: <https://doi.org/10.1002/9780470756416.ch17>.
- (2008). "Reformulating the Determiner Phrase Analysis". In: *Language and Linguistics Compass* 2.6, pp. 1–25. doi: <https://doi.org/10.1111/j.1749-818X.2008.00091.x>.
- Bertrand, Nicolas (2013). "Information Structure and Greek". In: *Encyclopaedia of Ancient Greek Language and Linguistics*. Ed. by Giorgios K. Giannakis. Leiden: Brill, pp. 238–243. doi: [http://dx.doi.org/10.1163/2214-448X\\_eagll\\_COM\\_00000190](http://dx.doi.org/10.1163/2214-448X_eagll_COM_00000190).
- Beschi, Fulvio (2018). "The Ancient Greek sentence left periphery. A study on Homer". In: *Journal of Greek Linguistics* 18.2, pp. 172–210. doi: <https://doi.org/10.1163/15699846-01802003>.

- Biberauer, Theresa and Ian Roberts (2005). "Changing EPP parameters in the history of English: accounting for variation and change". In: *English Language and Linguistics* 9.1, pp. 5–46. doi: 10.1017/S1360674305001528.
- (2015). "The Clausal Hierarchy, Features and Parameters". In: *Beyond Functional Sequence: The Cartography of Syntactic Structures, Volume 10*. Ed. by Uri Shlonsky. Oxford: Oxford University Press, pp. 295–313. doi: <https://doi.org/10.1093/acprof:oso/9780190210588.003.0016>.
- (2016). "Parameter typology from a diachronic perspective: The case of Conditional Inversion". In: *Theoretical Approaches to Linguistic Variation*. Ed. by Ermenegildo Bidese, Federica Cognola, and Manuela Caterina Moroni. *Linguistik Aktuell/Linguistics Today* 234. Amsterdam, Philadelphia: John Benjamins, pp. 259–292. doi: <https://doi.org/10.1075/la.234.10bib>.
- (2017). "Parameter setting". In: *The Cambridge Handbook of Historical Syntax*. Ed. by Adam Ledgeway and Ian Roberts. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, pp. 134–162. doi: 10.1017/9781107279070.008.
- Biberauer, Theresa and George Walkden (2015). "Introduction: Changing views of syntactic change". In: *Syntax over Time: Lexical, Morphological, and Information-Structural Interactions*. Ed. by Theresa Biberauer and George Walkden. Oxford Studies in Diachronic and Historical Linguistics. Oxford: Oxford University Press, pp. 1–14. doi: <https://doi.org/10.1093/acprof:oso/9780199687923.003.0001>.
- Bickel, Balthasar, Alena Witzlack-Makarevich, Kamal K. Choudhary, Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky (2015). "The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking". In: *PLOS ONE* 10.8. doi: <https://doi.org/10.1371/journal.pone.0132819>.
- Bielmeier, Roland (1989). "Yaghnōbī". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 480–488.
- Bonmann, Svenja (2017). "Überlegungen zur pronominalen Flexion sowie zu den *m*- und *b<sup>h</sup>*-Kasus des Indogermanischen". In: *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 14.2, pp. 119–189.
- Bonmann, Svenja, Jakob Halfmann, Natalie Korobzow, and Bobomullo Bobomulloev (2023). "A Partial Decipherment of the Unknown Kushan Script". In: *Transactions of the Philological Society* 121.2, pp. 293–329. doi: 10.1111/1467-968X.12269.
- Borer, Hagit (1984). *Parametric Syntax. Case Studies in Semitic and Romance Languages*. Berlin Boston: De Gruyter Mouton. doi: <https://doi.org/10.1515/9783110808506>.
- Bornkessel-Schlesewsky, Ina and Matthias Schlesewsky (2009). "The Role of Prominence Information in the Real-Time Comprehension of Transitive Constructions: A

- Cross-Linguistic Approach". In: *Language and Linguistics Compass* 3.1, pp. 19–58. doi: <https://doi.org/10.1111/j.1749-818X.2008.00099.x>.
- Bošković, Željko. (2005). "On the locality of left branch extraction and the structure of NP". In: *Studia Linguistica* 59.1, pp. 1–45. doi: <https://doi.org/10.1111/j.1467-9582.2005.00118.x>.
- (2007). "What will you have, DP or NP?" In: *Proceedings of the 37th Annual Meeting of the North East Linguistic Society*. Ed. by Emily Elfner and Martin Walkow. Amherst, MA: GLSA, University of Massachusetts, pp. 101–114.
- Bosson, Georg (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. *Ars Linguistica* 14. Commentationes analyticae et criticae. Tübingen: Gunter Narr Verlag Tübingen.
- Bowern, Claire (2008). "Syntactic change and syntactic borrowing in generative grammar". In: *Principles of syntactic reconstruction*. Ed. by Gisela Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 187–216. doi: <https://doi.org/10.1075/cilt.302.09bow>.
- Boyce, Mary (1964). "The use of relative particles in Western Middle Iranian". In: *Indo-Iranica: Mélanges présentés à Georg Morgenstierne à l'occasion de son soixante-dixième anniversaire*. Ed. by G. Redard. Wiesbaden: Harrassowitz, pp. 28–64.
- (1975). *A reader in Manichean Middle Persian and Parthian: Texts with Notes*. *Acta Iranica* 9. Leiden: Brill.
- Brandenstein, Wilhelm von and Manfred Mayrhofer (1964). *Handbuch des Altpersischen*. Wiesbaden: Harrassowitz.
- Brown, Dorcas R. and David W. Anthony (2019). "Late Bronze Age midwinter dog sacrifices and warrior initiations at Krasnosamarskoe, Russia". In: *Tracing the Indo-Europeans. New evidence from archaeology and historical linguistics*. Ed. by Birgit Anette Olsen, Thomas Olander, and Kristian Kristiansen. Oxford, Philadelphia: Oxbow Books, pp. 97–121. doi: <https://doi.org/10.2307/j.ctvmx3k2h.11>.
- Bruening, Benjamin (2020). "The head of the nominal is N, not D: N-to-D Movement, Hybrid Agreement, and conventionalized expressions". In: *Glossa* 5.1, 1–19]. doi: <http://doi.org/10.5334/gjgl.1031>.
- Brugmann, Karl (1909). *Lehre von den Wortformen und ihrem Gebrauch. 2. Teil, 1. Lieferung: Zahlwörter. Die drei Nominalgenera. Kasus- und Numerusbildung der Nomina. Pronominalstämme und Kasus- und Numerusbildung der Pronomina*. (= Grundriß der vergleichenden Grammatik der indogermanischen Sprachen, Zweite Bearbeitung, volume II). Strassburg: Trübner.
- (1911). *Lehre von den Wortformen und ihrem Gebrauch. 2. Teil, 2. Lieferung: Bedeutung der Numeri beim Nomen und Pronomen. Bedeutung der Kasus. Das Adjektivum. Die Adverbia nach Form und Gebrauch. Die Präpositionen nach Form und Gebrauch*. (=

- Grundriß der vergleichenden Grammatik der indogermanischen Sprachen, Zweite Bearbeitung, volume II). Strassburg: Trübner.
- (1913). *Lehre von den Wortformen und ihrem Gebrauch. 3. Teil, 1. Lieferung: Vorbemerkungen. Verbale Komposita. Augment. Reduplizierte Verbalbildungen. Die Tempusstämme im Allgemeinen. Präsens und starker Aorist. Die s-Aoriste. Das Perfekt und sein Augmenttempus.* (= Grundriß der vergleichenden Grammatik der indogermanischen Sprachen, Zweite Bearbeitung, volume II). Strassburg: Trübner.
  - (1916). *Lehre von den Wortformen und ihrem Gebrauch. 3. Teil, 2. Lieferung: Zusammengesetzte (periphrastische) Tempusbildungen. Die Modusbildungen. Die Personalendungen. Der Gebrauch der Formen des Verbum finitum. Der Gebrauch der Formen des Verbum infinitum. Partikeln im einfachen Satz.* (= Grundriß der vergleichenden Grammatik der indogermanischen Sprachen, Zweite Bearbeitung, volume II). Strassburg: Trübner.
- Brunner, Christopher J. (1977). *A syntax of western Middle Iranian*. Vol. 3. Persian Studies Series. Delmar, NY: Caravan Books.
- Caland, Willem (1897). “Eine übereinstimmung zwischen Vedischem und Avestischem sprachgebrauch”. In: *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 34.3, pp. 456–457. URL: <https://www.jstor.org/stable/40846223>.
- Campbell, Lyle (2020). *Historical Linguistics: An Introduction*. Fourth edition. Edinburgh: Edinburgh University Press.
- Carpelan, Christian, Asko Parpola, and Petteri Koskikallio, eds. (2001). *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at Tvärminne Research Station of the University of Helsinki 8-10 January, 1999*. Mémoires de la Société Finno-ougrienne 242. Helsinki: Société Finno-ougrienne.
- Casaretto, Antje, Jakob Halfmann, Natalie Korobzow, Daniel Kölligan, and Uta Reinöhl (2023). *The morphologically glossed Rigveda - The Zurich annotation corpus revised and extended. Hosted by VedaWeb - Online Research Platform for Old Indic Texts*. Cologne: University of Cologne. URL: <https://vedaweb.uni-koeln.de>.
- Ceolin, Andrea, Cristina Guardiano, Monica Aleandrina Irimia, and Giuseppe Longobardi (2020). “Formal Syntax and Deep History”. In: *Frontiers in Psychology* 11, pp. 1–6. doi: <https://doi.org/10.3389/fpsyg.2020.488871>.
- Ceolin, Andrea, Cristina Guardiano, Giuseppe Longobardi, Monica Alexandrina Irimia, Luca Bortolussi, and Andrea Sgarro (2021). “At the boundaries of syntactic prehistory”. In: *Philosophical Transactions of the Royal Society B* 376 (1824). doi: <https://doi.org/10.1098/rstb.2020.0197>.
- Chantraine, Pierre (1953). *Grammaire homérique. Tome II, Syntaxe*. Paris: C. Klincksieck.
- Cheung, Johnny (2002). *Studies in the Historical Development of the Ossetic Vocalism*. Beiträge zur Iranistik 21. Wiesbaden: Reichert.

- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1995). *The Minimalist Program*. Current Studies in Linguistics 28. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo (2005). “Deriving Greenberg’s Universal 20 and its Exceptions”. In: *Linguistic Inquiry* 36, pp. 315–332. doi: <http://doi.org/10.1162/0024389054396917>.
- (2010). *The syntax of Adjectives: A Comparative Study*. Cambridge, MA: MIT Press. doi: <https://doi.org/10.7551/mitpress/9780262014168.001.0001>.
- Clackson, James (2007). *Indo-European Linguistics: An Introduction*. Cambridge University Press.
- (2017). “The Comparative Method and Comparative Reconstruction”. In: *The Cambridge Handbook of Historical Syntax*. Ed. by Adam Ledgeway and Ian Roberts. Cambridge: Cambridge University Press, pp. 189–206. doi: <https://doi.org/10.1017/9781107279070.010>.
- Claudi, Ulrike (1994). “Word order change as category change: the Mande case”. In: *Perspectives on grammaticalization*. Ed. by William Pagliuca. Current Issues in Linguistic Theory 109. Amsterdam: John Benjamins, pp. 191–231. doi: <https://doi.org/10.1075/cilt.109.04cla>.
- Colarusso, John (2002). “Remarks on the Anau and Niyä Seals”. In: *Sino-Platonic Papers* 124, pp. 35–47. URL: [http://sino-platonic.org/complete/spp124\\_anau\\_niya\\_seals.pdf](http://sino-platonic.org/complete/spp124_anau_niya_seals.pdf).
- Coppola, Marie and Elissa L. Newport (2005). “Grammatical Subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input”. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.52, pp. 19249–19253. doi: <https://doi.org/10.1073/pnas.0509306102>.
- Corbett, Greville (1991). *Gender*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Corver, Norbert (1992). “On deriving certain left branch extraction asymmetries: A case study in parametric syntax”. In: *North East Linguistics Society* 22 (1), pp. 67–84. URL: <https://scholarworks.umass.edu/nels/vol22/iss1/6>.
- Crisma, Paola (1993). “On adjective placement in Romance and Germanic event nominals”. In: *Rivista di Grammatica Generativa* 18, pp. 61–100.
- Crisma, Paola and Giuseppe Longobardi, eds. (2009). *Historical Syntax and Linguistic Theory*. Oxford: Oxford University Press.
- Croft, William (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>.

- Crowley, Terry and Claire Bower (2010). *An Introduction to Historical Linguistics*. Fourth edition. Oxford: Oxford University Press.
- Curnow, Timothy Jowan (2001). "What language features can be 'borrowed'?" In: *Areal Diffusion and Genetic Inheritance. Problems in Comparative Linguistics*. Ed. by Alexandra Y. Aikhenvald and Robert M. W. Dixon. Explorations in Linguistic Typology. Oxford: Oxford University Press, pp. 412–436.
- Dabir-Moghaddam, Mohammad (2006). "Internal and External Forces in Typology: Evidence from Iranian Languages". In: *Journal of Universal Language* 7.1, pp. 29–47. doi: <https://doi.org/10.22425/jul.2006.7.1.29>.
- Daniels, Don (2021). "The issue of arbitrariness in syntactic reconstruction". In: *Folia Linguistica Historica* 42.2, pp. 393–428. doi: <https://doi.org/10.1515/flin-2021-2021>.
- Day, John V. (2001). *Indo-European Origins: The Anthropological Evidence*. Washington, D.C.: Institute for the Study of Man.
- de Vaan, Michiel (2019). "On the homonymy of 'put' and 'suck' in Proto-Indo-European". In: *Indo-European Linguistics* 7, pp. 176–193. doi: <https://doi.org/10.1163/22125892-00701008>.
- Debrunner, Albert and Jacob Wackernagel (1975). *Altindische Grammatik. Band III. Nominalflexion – Zahlwort – Pronomen. Unveränderter Nachdruck der 1. Auflage von 1929/1930*. Göttingen: Vandenhoeck & Ruprecht.
- Degener, Almuth (1993). "Zur Syntax des Khotanischen". In: *Medioiranica. Proceedings of the International Colloquium organized by the Katholieke Universiteit Leuven from the 21st to the 23rd of May 1990*. Ed. by Wojciech Skalmowski and Alois van Tongerloo. Orientalia Lovaniensia Analecta 48. Leuven: Peeters, pp. 45–56.
- (2002). "The Nuristani languages". In: *Indo-Iranian Languages and Peoples*. Ed. by Nicholas Sims-Williams. Proceedings of the British Academy 116. Oxford: Oxford University Press, pp. 103–118. doi: <https://doi.org/10.5871/bacad/9780197262856.003.0004>.
- Delbrück, Berthold (1871). *Der Gebrauch des Conjunctivs und Optativs im Sanskrit und Griechischen*. Halle an der Saale: Verlag der Buchhandlung des Waisenhauses.
- (1878). *Die altindische Wortfolge aus dem Çatapathabrāhmaṇa*. Halle an der Saale: Verlag der Buchhandlung des Waisenhauses.
- (1879). *Die Grundlagen der griechischen Syntax*. Halle an der Saale: Verlag der Buchhandlung des Waisenhauses.
- (1888). *Altindische Syntax*. Halle an der Saale: Verlag der Buchhandlung des Waisenhauses.
- (1907). *Synkretismus. Ein Beitrag zur germanischen Kasuslehre*. Strassburg: Trübner.
- (1893–1900). *Vergleichende Syntax der indogermanischen Sprachen*. 1. Theil (1893), 2. Theil (1897), 3. Theil (1900). (= Grundriß der vergleichenden Grammatik der indogermanischen Sprachen, volumes III, IV, V). Strassburg: Trübner.

- den Besten, Hans (1983). "On the Interaction of Root Transformations and Lexical Deletive Rules". In: *On the Formal Syntax of the Westgermania*. Ed. by Werner Abraham. Linguistik Aktuell, Amsterdamer Arbeiten zur theoretischen und angewandten Linguistik, Band 3. Amsterdam: John Benjamins, pp. 47–131. doi: <https://doi.org/10.1075/la.3>.
- Devine, Andrew M. and Laurence D. Stephens (2000). *Discontinuous Syntax. Hyperbaton in Greek*. New York/Oxford: Oxford University Press.
- (2006). *Latin Word Order. Structured Meaning and Information*. Oxford: Oxford University Press.
- Dik, Helma (1995). *Word Order in Ancient Greek: A Pragmatic Account of Word Order Variation in Herodotus*. Amsterdam Studies in Classical Philology 5. Amsterdam: J.-C. Gieben. doi: <https://doi.org/10.1163/9789004409002>.
- Dressler, Wolfgang (1971). "Über die Rekonstruktion der indogermanischen Syntax". In: *Zeitschrift für vergleichende Sprachforschung* 85.1, pp. 5–22. URL: <http://www.jstor.org/stable/40848307>.
- Dryer, Matthew S. (1992). "The Greenbergian word order correlations". In: *Language* 68.1, pp. 81–138.
- (2013). "Order of adposition and noun phrase". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: <http://wals.info/chapter/85>.
- (2018). "On the order of demonstrative, numeral, adjective, and noun". In: *Language* 94, pp. 798–833. doi: <http://doi.org/10.1353/lan.2018.0054>.
- (2019). "Grammaticalization accounts of word order correlations". In: *Explanation in Typology: Diachronic Sources, Functional Motivations and the Nature of the Evidence*. Ed. by Karsten Schmidtke-Bode, Natalia Levshina, Susanne Michaelis, and Ilja A. Seržant. Berlin: Language Science Press, pp. 63–96. doi: [10.5281/zenodo.2583788](https://doi.org/10.5281/zenodo.2583788).
- Dryer, Matthew S. and Martin Haspelmath, eds. (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: <https://wals.info/>.
- Dul'zon, Andrej P. (1959a). "Geografičeskie nazvanija Zapadnoj Sibiri kak istočnik eë drevnej istorii". In: *Nekotorye voprosy drevnej istorii Zapadnoj Sibiri*, pp. 24–28.
- (1959b). "Ketskie toponimy Zapadnoj Sibiri". In: *Učënye zapiski Tomskogo gosudarstvennogo pedagogičeskogo instituta*, pp. 91–111.
- Durkin-Meisterernst, Desmond (2009). "Khwarezmian". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 336–376.
- (2014). *Grammatik des Westmitteliranischen: Parthisch und Mittelpersisch*. Veröffentlichungen zur Iranistik 73. Wiesbaden: Harrassowitz.

- Èdel'man, Džoy I. (2008). "Xorezmijskij Jazyk". In: *Sredneiranskie i novoiranskie Jazyki*. Ed. by Valentin A. Efimov. Vol. VII. Osnovy Iranskogo Jazykoznanija. Moskva: Izdatel'stvo Firma Vostočnaya Literatura RAN, pp. 6–60.
- Edelman, Joy I. and Leila R. Dodykhudoeva (2009). "The Pamir Languages". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London and New York: Routledge, pp. 773–786.
- Emmerick, Ronald E. (1968). *Saka Grammatical Studies*. London oriental series 20. London: Oxford Univ. Press.
- (1989). "Khotanese and Tumshuqese". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 204–229.
- (2009). "Khotanese and Tumshuqese". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 377–415.
- Emonds, Joseph (1978). "The Verbal Complex V-V' in French". In: *Linguistic Inquiry* 9, pp. 151–175. URL: <https://www.jstor.org/stable/4178050>.
- Erschler, David (2019a). "A new argument for existence of the DP in languages without articles". In: *Journal of Linguistics* 55.4, pp. 879–887. DOI: 10.1017/S0022226719000288.
- (2019b). "Ossetic". In: *The Languages and Linguistics of Western Asia: An Areal Perspective*. Ed. by Geoffrey Haig and Geoffrey Khan. The World of Linguistics 6. Berlin, Boston: De Gruyter Mouton, pp. 851–881. DOI: <https://doi.org/10.1515/9783110421682-022>.
- Eska, Joseph F. (2007). "On basic configuration and movement within the Gaulish clause". In: *Gaulois et celtique continental. Études réunies par Pierre-Yves Lambert et Georges-Jean Pinault*. Ed. by Pierre-Yves Lambert and Georges-Jean Pinault. Genève: Librairie Droz S.A, pp. 215–229.
- Eythórsson, Thórhallur and Jóhanna Barðdal (2016). "Syntactic Reconstruction in Indo-European: State of the Art". In: *Veleia* 33, pp. 83–102. DOI: 10.1387/veleia.16823.
- Faarlund, Jan (1990). *Syntactic Change: Towards a Theory of Historical Syntax*. Trends in Linguistics: Studies and Monographs 50. Berlin and New York: Mouton de Gruyter.
- Ferraresi, Gisella and Maria Goldbach (2008). "Syntactic reconstruction: Methods and new insights". In: *Principles of Syntactic Reconstruction*. Ed. by Gisella Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 1–26. DOI: <https://doi.org/10.1075/cilt.302>.
- Fórizs, László (2016). "Apām Napāt, Dīrghatamas and Construction of the Brick Altar: Analysis of Ṛgveda 1.143". In: *Vedic Investigations (Papers of the 12th World Sanskrit Conference held in Helsinki, Finland, 13–18 July 2003, Vol. 1)*. Ed. by Petteri Koskikallio and Asko Parpola. Delhi: Motilal Banarsidass, pp. 97–126.
- Fortson, Benjamin W. (2010). *Indo-European Language and Culture: An Introduction*. Second Edition. Blackwell Textbooks in Linguistics, Vol. 19. Oxford: Wiley-Blackwell.



- Francfort, Henri-Paul (2001). "The archaeology of protohistoric Central Asia and the problems of identifying Indo-European and Uralic-speaking populations". In: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at Tvärminne Research Station of the University of Helsinki 8-10 January 1999*. Ed. by Christian Carpelan, Asko Parpola, and Petteri Koskikallio. Mémoires de la Société Finno-ougrienne 242. Helsinki: Société Finno-ougrienne, pp. 151–168.
- François, Alexandre (2014). "Trees, waves and linkages. Models of language diversification". In: *The Routledge Handbook of Historical Linguistics*. Ed. by Claire Bowerman and Bethwyn Evans. London and New York: Routledge, pp. 161–189. doi: <https://doi.org/10.4324/9781315794013>.
- Franklin, Amy, Anastasia Giannakidou, and Susan Goldin-Meadow (2011). "Negation, questions, and structure building in a homesign system". In: *Cognition* 118.3, pp. 398–416. doi: <https://doi.org/10.1016/j.cognition.2010.08.017>.
- Friedrich, Paul (1975). *Proto-Indo-European syntax: The order of meaningful elements*. Journal of Indo-European Studies; Monograph, No. 1. Washington: Institute for the Study of Man.
- Fukui, Naoki (1988). "Deriving the Differences between English and Japanese: A Case Study in Parametric Syntax". In: *English Linguistics* 5, pp. 249–270. doi: <https://doi.org/10.9793/elsj1984.5.249>.
- Garrett, Andrew (1990). "The Syntax of Anatolian Pronominal Clitics". PhD thesis. Cambridge, MA.: Harvard University.
- Geldner, Karl Friedrich (1951). *Der Rig-Veda. Aus dem Sanskrit ins Deutsche übersetzt und mit einem laufenden Kommentar versehen*. Cambridge, Mass.: Harvard University Press.
- Gershevitch, Ilya (1961). *A grammar of Manichean Sogdian*. Vol. 16. Publications of the Philological Society 16. Oxford: Basil Blackwell.
- Gholami, Saloumeh (2009). "Demonstratives determiners and pronouns in Bactrian". In: *From Daēnā to Dīn: Religion, Kultur und Sprache in der iranischen Welt: Festschrift für Philip Kreyenbroek zum 60. Geburtstag*. Ed. by Christine Allison, Anke Joisten-Pruschke, and Antje Wendtland. Wiesbaden: Harrassowitz, pp. 19–26.
- (2011). "Definite articles in Bactrian". In: *Topics in Iranian Linguistics*. Ed. by Agnes Korn, Geoffrey Haig, Simin Karimi, and Pollet Samvelian. Beiträge zur Iranistik, Band 34. Wiesbaden: Reichert, pp. 11–22.
- (2014). *Selected Features of Bactrian Grammar*. Göttinger Orientforschungen, III. Reihe: Iranica, Neue Folge 12. Wiesbaden: Harrassowitz.
- Ghomeshi, Jila, Ileana Paul, and Martina Wiltschko, eds. (2009). *Determiners: Universals and variation*. Linguistik Aktuell/Linguistics Today 147. Amsterdam: John Benjamins. doi: <https://doi.org/10.1075/la.147>.

- Gildea, Spike (2000). "On the Genesis of the Verb Phrase in Cariban Languages: Diversity through Reanalysis". In: *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*. Ed. by Spike Gildea. Typological Studies in Language 43. Amsterdam: John Benjamins, pp. 65–106. doi: <https://doi.org/10.1075/tsl.43>.
- Gildea, Spike, Eugenio R. Luján, and Jóhanna Barðdal (2020). "The Curious Case of Reconstruction in Syntax". In: *Reconstructing Syntax*. Ed. by Jóhanna Barðdal, Spike Gildea, and Eugenio R. Lujan. Leiden: Brill, pp. 1–44. doi: [https://doi.org/10.1163/9789004392007\\_002](https://doi.org/10.1163/9789004392007_002).
- Giorgi, Alessandra and Giuseppe Longobardi (1991). *The syntax of Noun Phrases: Configuration, parameters and empty categories*. Cambridge Studies in Linguistics 57. Cambridge: Cambridge University Press.
- Goldin-Meadow, Susan, D. Brentari, M. Coppola, L. Horton, and A. Senghas (2015). "Watching language grow in the manual modality: Nominals, predicates, and hand-shapes". In: *Cognition* 136, pp. 381–395. doi: <https://doi.org/10.1016/j.cognition.2014.11.029>.
- Goldin-Meadow, Susan, Savithry Namboodiripad, Carolyn Mylander, Aslı Özyürek, and Burcu Sancar (2015). "The Resilience of Structure Built Around the Predicate: Home-sign Gesture Systems in Turkish and American Deaf Children". In: *Journal of Cognition and Development* 16.1, pp. 55–80. doi: <https://doi.org/10.1080/15248372.2013.803970>.
- Goldstein, David (2016). *Classical Greek Syntax: Wackernagel's Law in Herodotus*. Brill's Studies in Indo-European Languages and Linguistics 16. Leiden and Boston: Brill.
- Greenberg, Joseph Harold (1966a). *Language Universals*. The Hague: Mouton de Gruyter.
- (1966b). "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: *Universals of Language: Reports of a Conference Held at Dobbs Ferry, New York, April 13-15, 1961*. Ed. by Joseph Harold Greenberg. Cambridge MA: MIT Press, pp. 73–113.
- (1966c). "Synchronic and diachronic universals in phonology". In: *Language* 42.2, pp. 508–17.
- Griffith, Ralph Thomas Hotchkin (1890). *The Hymns of the Rigveda: Translated with a popular commentary, Vol. II*. Benares: E.J. Lazarus.
- Grünberg, Aleksandr L. and Ivan M. Stéblin-Kamensky (1988). *La langue wakhi. 2. Essai grammatical et dictionnaire wakhi-français édité et traduit par Dominique Indjoudjian, suivi de Dictionnaire français-wakhi établi par Larissa Kydyrbaiéva*. Paris: Éditions de la Maison des Sciences de l'Homme.
- Guardiano, Cristina and Giuseppe Longobardi (2005). "Parametric comparison and language taxonomy". In: *Grammaticalization and Parametric Variation*. Ed. by Montserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca. Oxford: Oxford University Press, pp. 149–174.

- Guardiano, Cristina and Giuseppe Longobardi (2017). "Parameter Theory and Parametric Comparison". In: *The Oxford Handbook of Universal Grammar*. Ed. by Ian Roberts. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199573776.013.16.
- Haider, Hubert and Ronald Zwanziger (1984). "Relatively attributive: The 'ezāfe'-construction from Old Iranian to Modern Persian". In: *Historical Syntax*. Ed. by Jacek Fisiak. Trends in Linguistics, Studies and Monographs 23. Berlin, New York: De Gruyter Mouton, pp. 137–172. doi: <https://doi.org/10.1515/9783110824032.137>.
- Hajnal, Ivo (2018). "Graeco-Anatolian contacts in the Mycenaean period". In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 3. Berlin/Boston: de Gruyter, pp. 2037–2055. doi: <https://doi.org/10.1515/9783110542431-041>.
- Häkkinen, Jaakko (2012). "Early contacts between Uralic and Yukaghir". In: *Per Urales ad Orientem. Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*. Ed. by Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi, and Erika Sandman. Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 264. Helsinki: Société Finno-Ougrienne, pp. 91–101. URL: [https://www.sgr.fi/sust/sust264/sust264\\_hakkinen\\_j.pdf](https://www.sgr.fi/sust/sust264/sust264_hakkinen_j.pdf).
- Hale, Mark (1987a). "Notes on Wackernagel's Law in the language of the Rigveda". In: *Studies in Memory of Warren Cowgill (1929-1985): Papers from the Fourth East Coast Indo-European Conference, Cornell University, June 6-9, 1985*. Ed. by Calvert Watkins. Berlin, Boston: de Gruyter, pp. 38–50. doi: <https://doi.org/10.1515/9783110853223.38>.
- (1987b). "Studies in the comparative syntax of the oldest Indo-Iranian languages". PhD thesis. Cambridge, MA.: Harvard University.
- (1998). "Diachronic Syntax". In: *Syntax* 1, pp. 1–18. doi: <https://doi.org/10.1111/1467-9612.00001>.
- (2007). *Historical Linguistics: Theory and Method*. Blackwell Textbooks in Linguistics 21. Oxford, Malden, Mass.: Blackwell.
- (2014). "The Comparative Method: Theoretical Issues". In: *The Routledge Handbook of Historical Linguistics*. Ed. by Claire Bower and Bethwyn Evans. London: Routledge, pp. 146–160. doi: <https://doi.org/10.4324/9781315794013>.
- (2018). "The syntax of Indo-Iranian". In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Berlin/Boston: De Gruyter. Chap. 112, pp. 1942–1924. doi: <https://doi.org/10.1515/9783110542431-033>.
- Hardy, Heather K. (2005). "Alabama". In: *Native Languages of the Southeastern United States*. Ed. by Heather K. Hardy and Janine Scancarelli. Studies in the Anthropol-

- ogy of North American Indians. Lincoln and London: University of Nebraska Press, pp. 75–113.
- Harris, Alice C. (1985). *Diachronic Syntax: The Kartvelian Case*. Syntax and Semantics, Volume 18. New York [Leiden]: Academic Press [Brill]. doi: <https://doi.org/10.1163/9789004373143>.
- (2008). “Reconstruction in syntax: Reconstruction of patterns”. In: *Principles of syntactic reconstruction*. Ed. by Gisella Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 73–95. doi: <https://doi.org/10.1075/cilt.302>.
- Harris, Alice C. and Lyle Campbell (1995). *Historical syntax in cross-linguistic perspective*. Cambridge Studies in Linguistics 74. Cambridge: Cambridge University Press. doi: <https://doi.org/10.1017/CBO9780511620553>.
- Harrison, Sheldon P. (2003). “On the Limits of the Comparative Method”. In: *The Handbook of Historical Linguistics*. Ed. by Brian D. Joseph and Richard D. Janda. Blackwell Handbooks in Linguistics. Oxford: Blackwell, pp. 213–243. doi: <https://doi.org/10.1002/9780470756393.ch2>.
- Havers, Wilhelm (1911). *Untersuchungen zur Kasussyntax der indogermanischen Sprachen*. Untersuchungen zur indogermanischen Sprach- und Kulturwissenschaft 3. Berlin, New York: De Gruyter. doi: <https://doi.org/10.1515/9783111384566>.
- Hawkins, John A. (1983). *Word Order Universals*. New York: Academic Press. doi: <https://doi.org/10.1016/C2009-0-21896-4>.
- Henning, Walter Bruno (1945). “Sogdian Tales”. In: *Bulletin of the School of Oriental and African Studies* 11.3, pp. 465–487. doi: [10.1017/S0041977X00072244](https://doi.org/10.1017/S0041977X00072244).
- (1955). “The Khwarezmian Language”. In: *60. doğum yılı münasebetiyle Zeki Velidi Togan’a Armağan: symbolae in honorem Z. V. Togan*. Ed. by Tertip Heyeti. Istanbul: Maarif basımevi, pp. 421–436.
- Hermann, Eduard (1895). “Gab es im Indogermanischen Nebensätze?” In: *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 33.4, pp. 481–535. URL: <https://www.jstor.org/stable/40845941>.
- Heston, Wilma Louise (1976). “Selected Problems in fifth to tenth century Iranian Syntax”. PhD thesis. Philadelphia: University of Pennsylvania. URL: <https://www.proquest.com/dissertations-theses/selected-problems-fifth-tenth-century-iranian/docview/288323305/se-2?accountid=10218>.
- Hettrich, Heinrich (1984). “Zur historischen Syntax der nomina actionis im Ṛgveda: Der ‘doppelte Dativ’”. In: *Münchener Studien zur Sprachwissenschaft* 43, pp. 55–106.
- (1988). *Untersuchungen zur Hypotaxe im Vedischen*. Untersuchungen zur indogermanischen Sprach- und Kulturwissenschaft, N.F. Band 4. Berlin, Boston: De Gruyter. doi: <https://doi.org/10.1515/9783110892123>.

- Hettrich, Heinrich (2007). *Materialien zu einer Kasussyntax des Rgveda*. Würzburg. URL: <https://www.phil.uni-wuerzburg.de/fileadmin/04080400/Materialien.pdf>.
- Hewson, John and Vit Bubenik (2006). *From Case to Adposition: The development of configurational syntax in Indo-European languages*. Current Issues in Linguistic Theory 280. Amsterdam: John Benjamins. doi: <https://doi.org/10.1075/cilt.280>.
- Hiebert, Fredrik Talmage (1994). *Origins of the bronze age oasis civilization in central Asia*. American School of Prehistoric Research Bulletin 42. Cambridge, Mass.: Harvard University, Peabody Museum of Archaeology and Ethnology.
- (2002a). “Bronze Age interaction between the Eurasian steppe and Central Asia”. In: *Ancient interactions: East and West in Eurasia*. Ed. by Katie Boyle, Colin Renfrew, and Marsha Levine. Cambridge: McDonald Institute for Archaeological Research, pp. 237–248.
  - (2002b). “The Context of the Anau Seal”. In: *Sino-Platonic Papers* 124, pp. 1–34. URL: [http://sino-platonic.org/complete/spp124\\_anau\\_niya\\_seals.pdf](http://sino-platonic.org/complete/spp124_anau_niya_seals.pdf).
- Hill, Eugen (2022). “The dative and instrumental dual in East Baltic”. In: *Vilnius University Open Series, Baltistikos platybėse*, pp. 145–160. doi: 10.15388/Baltistikos\_platybese.2022.7.
- Hintze, Almut (2015). “Zarathustras Time and Homeland: Linguistic perspectives”. In: *The Wiley Blackwell Companion to Zoroastrianism*. Ed. by Michael Stausberg and Yuhan Sohrab-Dinshaw Vevaind. Oxford: Oxford University Press, pp. 31–38.
- Hirt, Hermann (1934). *Syntax I. Syntaktische Verwendung der Kasus und der Verbalformen*. Indogermanische Grammatik VI. Heidelberg: Carl Winter.
- (1937). *Syntax II. Die Lehre vom einfachen und zusammengesetzten Satz*. Indogermanische Grammatik VII. Heidelberg: Carl Winter.
- Hock, Hans Henrich (2013). “Proto-Indo-European verb-finality”. In: *Journal of Historical Linguistics* 3, pp. 49–76.
- Hock, Hans Henrich and Brian D. Joseph (2009). *Language History, Language Change, and Language Relationship. An Introduction to Historical and Comparative Linguistics*. Second revised edition. Berlin/New York: Mouton de Gruyter.
- Hoffmann, Karl and Bernhard Forssmann (2004). *Avestische Laut- und Flexionslehre*. 2nd ed. Vol. 84. Innsbrucker Beiträge zur Sprachwissenschaft. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Holland, Gary (1986). “Relativization and word order in Old Irish and Vedic Sanskrit”. In: *Journal of Indo-European Studies* 24, pp. 323–331.
- (2003). “Syntactic Reconstruction”. In: *The International Encyclopedia of Linguistics, 2nd edition*. Ed. by William Frawley. Vol. 3. Oxford: Oxford University Press, pp. 448–450.

- Holmberg, Anders and David Odden (2008). "The noun phrase in Hawrami". In: *Aspects of Iranian Linguistics*. Ed. by Simin Karimi, Vida Samiian, and Donald L. Stilo. Cambridge Scholars Publishing, pp. 129–152.
- Holopainen, Sampsa (2019). "Indo-Iranian borrowings in Uralic. Critical overview of the sound substitutions and the distribution criterion". PhD thesis. Helsinki: University of Helsinki. URL: <https://helda.helsinki.fi/bitstream/handle/10138/307582/INDO-IRA.pdf?sequence=1&isAllowed=y>.
- Huang, C.-T. J. (2015). "On syntactic analyticity and parametric theory". In: *Chinese syntax in a cross-linguistic perspective*. Ed. by A. Li, A. Simpson, and W.-T. Dylan Tsai. Oxford: Oxford University Press, pp. 1–50.
- Humbach, Helmut (1989). "Choresmian". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 193–203.
- Jacquiod, Bernard (2017). "The syntax of Greek". In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 1. Berlin/Boston: De Gruyter, pp. 682–695. doi: <https://doi.org/10.1515/9783110261288-042>.
- Jamison, Stephanie W. (1991). *The Ravenous Hyenas and the Wounded Sun: Myth and Ritual in Ancient India*. Ithaca, NY / London: Cornell University Press.
- (1993). "Determining the Synchronic Syntax of a Dead Language". In: *Historical Linguistics 1989. Papers from the 9th international Conference on Historical Linguistics, Rutgers University, 14-18 August 1989*. Ed. by Henk Aertsen and Robert J. Jeffers. Amsterdam/Philadelphia: John Benjamins, pp. 211–220.
- Janhunen, Juha (2009). "Proto-Uralic—what, where, and when?" In: *The Quasquicentennial of the Finno-Ugric Society*. Ed. by Yussi Ylikoski. Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 258. Société Finno-Ougrienne, pp. 57–78. URL: [https://www.sgr.fi/sust/sust258/sust258\\_janhunen.pdf](https://www.sgr.fi/sust/sust258/sust258_janhunen.pdf).
- (2020). "Towards an understanding of the lexical history of the Yeniseic languages". In: *International Journal of Eurasian Linguistics* 2.1, pp. 166–171. doi: <https://doi.org/10.1163/25898833-12340029>.
- Jeffers, Robert J. (1976). "Syntactic Change and Syntactic Reconstruction. Current Progress in Historical Linguistics". In: *Proceedings of the Second International Conference on Historical Linguistics, Tucson, Arizona, 12–16 January 1976*. Ed. by Jr. William M. Christie. Amsterdam, New York: North Holland, Elsevier, pp. 1–16.
- Jolly, Julius (1872). *Ein Kapitel vergleichender Syntax: Der Conjunctiv und Optativ und die Nebensätze im Zend und Altpersischen in Vergleich mit dem Sanskrit und Griechischen*. München: Theodor Ackermann.

- Jügel, Thomas (2014). "On the linguistic history of Kurdish". In: *Kurdish Studies* 2.2, pp. 123–142. URL: <https://kurdishstudies.net/menu-script/index.php/KS/article/view/61>.
- (2017). "The syntax of Iranian". In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 1. Berlin/Boston: De Gruyter, pp. 549–566. DOI: <https://doi.org/10.1515/9783110261288-036>.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Keenan, Edward L. and Denis Paperno (2012). *Handbook of Quantifiers in Natural Language*. Studies in Linguistics and Philosophy, Vol. 90. Dordrecht: Springer. DOI: <https://doi.org/10.1007/978-94-007-2681-9>.
- Kellens, Jean (1989). "Avestique". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 32–55.
- Kent, Roland G. (1953). *Old Persian Grammar, Texts, Lexicon. Second Edition, Revised*. American Oriental Series 33. New Haven, CT: American Oriental Society.
- Keydana, Götz (1997). *Absolute Konstruktionen in altindogermanischen Sprachen*. Historische Sprachforschung, Ergänzungsheft 40. Göttingen: Vandenhoeck & Ruprecht.
- (2018). "The syntax of Proto-Indo-European". In: *Handbook of comparative and historical Indo-European linguistics*. Ed. by Brian Joseph Jared Klein and Matthias Fritz. Vol. 3. Berlin, Boston: De Gruyter Mouton, pp. 2195–2228. DOI: <https://doi.org/10.1515/9783110542431-044>.
- Kikusawa, Ritsuko (2003). "The Development of Some Indonesian Pronominal Systems". In: *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13-17 August 2001*. Ed. by Barry J. Blake, Kate Burridge, and Jo Taylor. Amsterdam: John Benjamins, pp. 237–268.
- (2020). "Conducting Syntactic Reconstruction of Languages with no Written Records". In: *Reconstructing Syntax*. Ed. by Jóhanna Barðdal, Spike Gildea, and Eugenio R. Lujan. Leiden: Brill, pp. 108–161. DOI: [https://doi.org/10.1163/9789004392007\\_004](https://doi.org/10.1163/9789004392007_004).
- Kiparsky, Paul (2014). "New perspectives in historical linguistics". In: *The Routledge Handbook of Historical Linguistics*. Ed. by Claire Bowern and Bethwyn Evans. New York: Routledge, pp. 64–102. DOI: <https://doi.org/10.4324/9781315794013>.
- Konow, Sten (1941). *Khotansakische Grammatik. Mit Bibliographie, Lesestücken und Wörterverzeichnis*. Porta Linguarum Orientalium 22. Leipzig: Harrassowitz.
- Koopman, Hilda Judith (1984). "The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar". PhD thesis. Tilburg University. URL: <https://pure.uvt.nl/ws/portalfiles/portal/1204885/3955255.pdf>.

- Korn, Agnes (2016). "A partial tree of Central Iranian: A new look at Iranian subphyla". In: *Indogermanische Forschungen* 121.1, pp. 401–434. doi: <https://doi.org/10.1515/if-2016-0021>.
- (2019). "Isoglosses and subdivisions of Iranian". In: *Journal of Historical Linguistics* 9.2, pp. 239–281. doi: <https://doi.org/10.1075/jhl.17010.kor>.
- Kroonen, Guus, Gojko Barjamovic, and Michaël Peyrot (2018). *Linguistic supplement to Damgaard et al. 2018: Early Indo-European languages, Anatolian, Tocharian and Indo-Iranian*. DOI: 10.5281/zenodo.1240524. URL: <https://doi.org/10.5281/zenodo.1240524>.
- Kuhn, Adalbert (1853). "Ueber die durch nasale erweiterten Verbalstämme". In: *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 2.6, pp. 455–471. URL: <https://www.jstor.org/stable/40845266>.
- Kümmel, Martin Joachim (2018). "The morphology of Indo-Iranian". In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. De Gruyter. Chap. 111, pp. 1888–1924. doi: 10.1515/9783110542431-032.
- Kuz'mina, Elena E. (2007). *The Origin of the Indo-Iranians*. Ed. by James P. Mallory. Leiden Indo-European Etymological Dictionary Series, 3. Leiden: Brill. doi: <https://doi.org/10.1163/ej.9789004160545.i-763>.
- Kuznetsov, Pavel F. (2006). "The emergence of Bronze Age chariots in eastern Europe". In: *Antiquity* 80.309, pp. 638–645. doi: 10.1017/S0003598X00094096.
- Lazard, Gilbert (1992). *A grammar of contemporary Persian*. Translated by Shirley A. Lyon. Persian studies series 14. Costa Mesa, CA and New York: Mazda Publishers.
- Le Feuvre, Claire, Daniel Petit, and Georges-Jean Pinault, eds. (2017). *Verbal Adjectives and Participles in Indo-European languages/Adjectifs verbaux et participes dans les langues indo-européennes. Proceedings of the conference of the Society for Indo-European Studies (Indogermanische Gesellschaft), Paris, 24th to 26th September 2014*. Bremen: Hempen.
- Ledgeway, Adam and Ian Roberts (2017). "Principles and Parameters". In: *The Cambridge Handbook of Historical Syntax*. Ed. by Adam Ledgeway and Ian Roberts. Cambridge Handbooks in Languages and Linguistics. Cambridge: Cambridge University Press, pp. 581–628. doi: 10.1017/9781107279070.028.
- Lehmann, Winfred Philipp (1973). "A Structural Principle of Language and its Implications". In: *Language* 49.1, pp. 47–66. doi: <https://doi.org/10.2307/412102>.
- (1974). *Proto-Indo-European Syntax*. Austin and London: University of Texas Press.
- (1976). "From topic to subject in Indo-European". In: *Subject and Topic*. Ed. by Charles N. Li. London, New York: Academic Press, pp. 447–456.



- Li, Charles N. (1973). *Word Order and Word Order Change*. Austin: University of Texas Press.
- Lightfoot, David W. (1979). *Principles of Diachronic Syntax*. Cambridge Studies in Linguistics 23. New York and London: Cambridge University Press.
- (1980). “On reconstructing a proto-syntax”. In: *Linguistic Reconstruction and Indo-European Syntax*. Ed. by Paolo Ramat. Current Issues in Linguistic Theory 19. Amsterdam: John Benjamins, pp. 27–45. doi: <https://doi.org/10.1075/cilt.19>.
  - (1991). *How to set parameters: Arguments from language change*. Cambridge MA, London: MIT Press.
  - (1998). *The development of language: acquisition, change, and evolution*. Oxford: Blackwell.
  - (2002). “Myths and the Prehistory of Grammars”. In: *Journal of Linguistics* 38.1, pp. 113–136. doi: [10.1017/S0022226701001268](https://doi.org/10.1017/S0022226701001268).
  - (2006). *How new languages emerge*. Cambridge: Cambridge University Press. doi: <https://doi.org/10.1017/CBO9780511616204>.
- Livšić, Vladimir Aronovič (2003). “Three silver bowls from the Isakovka burial-ground no. 1 with Khwarezmian and Parthian inscriptions”. In: *Ancient Civilizations from Scythia to Siberia. An International Journal of Comparative Studies in History and Archaeology* 9.1–2, pp. 147–172. doi: <https://doi.org/10.1163/157005703322114874>.
- Longobardi, Giuseppe (2001a). “How Comparative is Semantics? A Unified Parametric Theory of Bare Nouns and Proper Names”. In: *Natural Language Semantics* 9, pp. 335–369. doi: <https://doi.org/10.1023/A:1014861111123>.
- (2001b). “The Structure of DPs: Some Principles, Parameters and Problems”. In: *The Handbook of Contemporary Syntactic Theory*. Ed. by Mark Baltin and Chris Collins. Oxford: Blackwell, pp. 562–603. doi: <https://doi.org/10.1002/9780470756416.ch18>.
  - (2003). “On parameters and parameter theory”. In: *Syntaxtheorien: Modelle, Methoden, Motive*. Ed. by Elisabeth Stark and Ulrich Wandruszka. Tübingen: Narr, pp. 273–290.
  - (2005). “A minimalist program for parametric linguistics?” In: *Organizing Grammar: Linguistic studies in Honor of Henk van Riemsdijk*. Ed. by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster. Berlin, New York: De Gruyter Mouton, pp. 407–414. doi: <https://doi.org/10.1515/9783110892994.407>.
  - (2008). “Reference to individuals, person, and the variety of mapping parameters”. In: *Essays on Nominal Determination: From morphology to discourse management*. Ed. by Henrik Høeg Müller and Alex Klinge. Studies in Language Companion Series

99. Amsterdam: John Benjamins, pp. 189–211. doi: <https://doi.org/10.1075/slcs.99.11lon>.
- Longobardi, Giuseppe and Cristina Guardiano (2009). “Evidence for syntax as a signal of historical relatedness”. In: *Lingua* 119.11, pp. 1679–1706. doi: <https://doi.org/10.1016/j.lingua.2008.09.012>.
- (2017). “Phylogenetic reconstruction in syntax: the Parametric Comparison Method”. In: *The Cambridge Handbook of Historical Syntax*. Ed. by Adam Ledgeway and Ian Roberts. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, pp. 241–272. doi: [10.1017/9781107279070.013](https://doi.org/10.1017/9781107279070.013).
- Longobardi, Giuseppe, Cristina Guardiano, Giuseppina Silvestri, Alessio Boattini, and Andrea Ceolin (2013). “Toward a syntactic phylogeny of modern Indo-European languages”. In: *Journal of Historical Linguistics* 3.1, pp. 122–152. doi: <https://doi.org/10.1075/jhl.3.1.07lon>.
- Longobardi, Giuseppe and Giuseppina Silvestri (2012). “The Structure of NPs: Some Insights on Case, Empty Categories and Poverty of Stimulus”. In: *The Bloomsbury Companion to Syntax*. Ed. by Silvia Luraghi and Claudia Parodi. London, New Delhi, New York, Sydney: Bloomsbury, pp. 88–117.
- Lorimer, David Lockhart Robertson (1958). *The Wakhi Language. Volume 1: Introduction, Phonetics, Grammar and Texts, Volume 2: Vocabulary and Index*. London: School of Oriental and African Studies.
- Lubotsky, Alexander (2001). “The Indo-Iranian substratum”. In: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at Tvärminne Research Station of the University of Helsinki 8-10 January 1999*. Ed. by Christian Carpelan, Asko Parpola, and Petteri Koskikallio. Mémoires de la Société Finno-ougrienne 242. Helsinki: Société Finno-ougrienne, pp. 301–31.
- Lundquist, Jesse and Anthony D. Yates (2018). “The morphology of Proto-Indo-European”. In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 3. Berlin/Boston: De Gruyter, pp. 2079–2195. doi: <https://doi.org/10.1515/9783110542431-044>.
- MacKenzie, David Neil (1990). *The Khwarezmian Element in the Qunyat al-munya*. London: School of Oriental and African Studies, University of London.
- Malandra, William W. (1991). “Chariot”. In: *Encyclopædia Iranica*. Last Updated: October 13, 2011. Vol. V. 4, pp. 377–380. url: <http://www.iranicaonline.org/articles/chariot-av>.
- Mallory, James Patrick (1989). *In Search of the Indo-Europeans. Language, Archaeology and Myth*. London and New York: Thames and Hudson.
- Mallory, James Patrick and Douglas Q. Adams, eds. (1997). *Encyclopedia of Indo-European Culture*. London and Chicago: Fitzroy-Dearborn.

- Mallory, James Patrick and Victor H. Mair (2000). *The Tarim Mummies. Ancient China and the Mystery of the Earliest Peoples from the West*. (First paperback edition 2008). London and New York: Thames and Hudson.
- Maloletko, Aleksej Mixajlovič (1992). *Paleotoponimika*. Tomsk: Tomskij gosudarstvennyj universitet.
- Marchenko, Z., S. Svyatko, V. Molodin, A. Grishin, and M. Rykun (2017). "Radiocarbon Chronology of Complexes With Seima-Turbino Type Objects (Bronze Age) in Southwestern Siberia". In: *Radiocarbon* 59.5, pp. 1381–1397. doi: 10.1017/RDC.2017.24.
- Masson, Vadim Mikhaïlovich (1986). "Archeology. V: Pre-Islamic Central Asia". In: *Encyclopædia Iranica*. Last Updated: August 11, 2011. Vol. II. 3, pp. 308–317. URL: <http://www.iranicaonline.org/articles/archeology-v>.
- Matić, Dejan (2003). "Topic, focus, and discourse structure. Ancient Greek Word Order". In: *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 27.3, pp. 573–633. doi: <https://doi.org/10.1075/sl.27.3.05mat>.
- Mayrhofer, Manfred (1966). *Die Indo-Arier im alten Vorderasien. Mit einer analytischen Bibliographie*. Wiesbaden: Harrassowitz.
- (1982). "Welches Material aus dem Indo-Arischen von Mitanni verbleibt für eine selektive Darstellung?" In: *Investigationes philologicae et comparativae: Gedenkschrift für Heinz Kronasser*. Ed. by Erich Neu. Wiesbaden: Harrassowitz, pp. 72–90.
  - (1989). "Vorgeschichte der iranischen Sprachen; Uriranisch". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Reichert, pp. 4–24.
  - (1992–2001). *Etymologisches Wörterbuch des Altindoarischen*. Vol. I–III. Heidelberg: Carl Winter Universitätsverlag.
- Meier-Brügger, Michael (2010). *Indogermanische Sprachwissenschaft*. 9., durchgesehene und ergänzte Auflage. Unter Mitarbeit von Matthias Fritz und Manfred Mayrhofer. Berlin/New York: De Gruyter.
- Meiser, Gerhard (2010). *Historische Laut- und Formenlehre der lateinischen Sprache. Dritte Auflage*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Miller, D. Gary (1975). "Indo-European: VSO, SOV, SVO, or all three?" In: *Lingua* 37.1, pp. 31–52. doi: [https://doi.org/10.1016/0024-3841\(75\)90003-0](https://doi.org/10.1016/0024-3841(75)90003-0).
- Mithun, Marianne (1999). *The Languages of Native North America*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- Morgenstierne, Georg (1973). *Indo-Iranian Frontier Languages, Volume II: Iranian Pamir languages (Yidgha-Munji, Sanglechi-Ishkashmi and Wakhi)*. 2. ed. rev. and with new material. Serie B: Skrifter XXXV. Oslo: Universitetsforlaget.

- Narasimhan, Vagheesh M. et al. (2019). "The formation of human populations in South and Central Asia". In: *Science* 365.6457. doi: <https://doi.org/10.1126/science.aat7487>.
- Newmeyer, F. J. (2004). "Against a parameter-setting approach to typological variation". In: *Linguistic Variation Yearbook* 4.1, pp. 181–234. doi: <https://doi.org/10.1075/livy.4.06new>.
- Nichols, Johanna (1986). "Head-Marking and Dependent-Marking Grammar". In: *Language* 62.1, pp. 56–119. doi: <https://doi.org/10.2307/415601>.
- (1996). "The Comparative Method as Heuristic". In: *The Comparative Method Re-viewed: Regularity and Irregularity in Language Change*. Ed. by Mark Durie and Malcolm Ross. New York, Oxford: Oxford University Press, pp. 39–71.
- Oberlies, Thomas (2012). *Der Rigveda und seine Religion*. Berlin: Verlag der Weltreligionen.
- Olander, Thomas (2019). "The Indo-European homeland: Introducing the problem". In: *Tracing the Indo-Europeans. New evidence from archaeology and historical linguistics*. Ed. by Birgit Anette Olsen, Thomas Olander, and Kristian Kristiansen. Oxford / Philadelphia: Oxbow Books, pp. 7–34.
- Paperno, Denis and Edward L. Keenan (2012). *Handbook of Quantifiers in Natural Language: Volume II*. Studies in Linguistics and Philosophy, Vol. 97. Dordrecht: Springer. doi: <https://doi.org/10.1007/978-3-319-44330-0>.
- Parpola, Asko (2002). "From the dialects of Old Indo-Aryan to Proto-Indo-Aryan and Proto-Iranian". In: *Indo-Iranian Languages and Peoples*. Ed. by Nicholas Sims-Williams. Proceedings of the British Academy 116. Oxford: Oxford University Press, pp. 43–102. doi: <https://doi.org/10.5871/bacad/9780197262856.001.0001>.
- Paul, Ludwig (2019). "Persian". In: *The Languages and Linguistics of Western Asia: An Areal Perspective*. Ed. by Geoffrey Haig and Geoffrey Khan. The World of Linguistics 6. Berlin, Boston: De Gruyter Mouton, pp. 569–624. doi: <https://doi.org/10.1515/9783110421682-017>.
- Paxalina, Tatjana N. (1969). *Pamirskie jazyki. Jazyki Narodov Azii i Afriki*. Moskva: Akademija Nauk SSSR.
- (1975). *Vaxanskij jazyk. Jazyki Vostočnogo Gindukusha*. Moskva: Akademija Nauk SSSR.
- Payne, John (1989). "Pāmir Languages". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 417–444.
- Pereltsvaig, Asya and Martin W. Lewis (2015). *The Indo-European Controversy: Facts and Fallacies in Historical Linguistics*. Cambridge: Cambridge University Press.
- Pinault, Georges-Jean (2006). "Further links between the Indo-Iranian substratum and the BMAC language". In: *Themes and Tasks in Old and Middle Indo-Aryan Linguistics*.

- Ed. by Bertil Tikkanen and Heinrich Hettrich. Papers of the 12th World Sanskrit Conference Vol. 5. Delhi: Motilal Banarsidass Publishers, pp. 167–196.
- Pires, Acrisio and Sarah Grey Thomason (2008). “How Much Syntactic Reconstruction is Possible?” In: *Principles of Syntactic Reconstruction*. Ed. by Gisela Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 27–72. doi: <https://doi.org/10.1075/cilt.302>.
- Plank, Frans, ed. (2003). *Noun phrase Structure in the Languages of Europe*. Empirical Approaches to Language Typology, Eurotyp 20-7. Berlin, New York: De Gruyter Mouton. doi: <https://doi.org/10.1515/9783110197075>.
- Pollock, Jean-Yves (1989). “Verb Movement, Universal Grammar and the Structure of IP”. In: *Linguistic Inquiry* 20.3, pp. 365–424. URL: <https://www.jstor.org/stable/4178634>.
- Probert, Philomen (2015). *Early Greek Relative Clauses*. Oxford: Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780198713821.001.0001>.
- Rastorgueva, Vera S. and E. K. Molčanova (1981a). “Parfjanskij Jazyk”. In: *Sredneiranskije Jazyki*. Vol. II. Osnovy Iranskogo Jazykoznanija. Moskva: Nauka, pp. 147–232.
- (1981b). “Srednepersidskij Jazyk”. In: *Sredneiranskije Jazyki*. Vol. II. Osnovy Iranskogo Jazykoznanija. Moskva: Nauka, pp. 6–146.
- Reichelt, Hans (1978). *Awestisches Elementarbuch. 3., unveränderte Auflage (Nachdruck der 1. Auflage Heidelberg 1909)*. 3rd ed. Indogermanische Bibliothek: Erste Reihe, Grammatiken: Band 5. Heidelberg: Winter.
- Riesberg, Sonja, Kurt Malcher, and Nikolaus P. Himmelmann (2019). “How universal is agent-first? Evidence from symmetrical voice languages”. In: *Language* 95.3, pp. 523–561. doi: <http://doi.org/10.1353/lan.2019.0055>.
- Riolfi, Alessandro (2017). “A history of the concept of parameter in Generative Grammar”. PhD thesis. Verona: Università degli Studi di Verona. URL: <https://core.ac.uk/download/pdf/217568085.pdf>.
- Rizzi, Luigi (1982). *Issues in Italian Syntax*. Studies in Generative Grammar 11. Berlin, Boston: De Gruyter Mouton. doi: <https://doi.org/10.1515/9783110883718>.
- Roberts, Ian (1998). “Review of Alice Harris and Lyle Campbell, *Historical Syntax in Cross-Linguistic Perspective* (Cambridge: Cambridge University Press, 1995)”. In: *Romance Philology* 51.3, pp. 363–370. URL: <https://www.jstor.org/stable/44941034>.
- (2007). *Diachronic Syntax*. Oxford: Oxford University Press.
- (2012). “Macroparameters and Minimalism: A Programme for Comparative Research”. In: *Parameter Theory and Linguistic Change*. Ed. by Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sândalo, and Juanito Avelar. Oxford Studies

- in *Diachronic and Historical Linguistics*. Oxford: Oxford University Press, pp. 320–335.
- (2021). *Diachronic Syntax*. Second Edition. Oxford: Oxford University Press.
- Robson, Barbara and Habibullah Tegey (2009). “Pashto”. In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 721–772.
- Ross, Malcolm D. (1988). *Proto-Oceanic and the Austronesian languages of Western Melanesia*. Pacific Linguistics C-98. Canberra: The Australian National University.
- Rumsey, Alan (1987). “The chimera of Proto-Indo-European ergativity: Lessons for historical syntax”. In: *Lingua* 71, pp. 297–318. doi: [https://doi.org/10.1016/0024-3841\(87\)90076-3](https://doi.org/10.1016/0024-3841(87)90076-3).
- Salzmann, Martin (2020). “The NP vs. DP debate. Why previous arguments are inconclusive and what a good argument could look like. Evidence from agreement with hybrid nouns”. In: *Glossa: A Journal of General Linguistics* 5.1, pp. 1–46. doi: <https://doi.org/10.5334/gjgl.1123>.
- Samvelian, Pollet (2018). “Specific Features of Persian Syntax: The Ezafe construction, differential object marking, and complex predicates”. In: *The Oxford Handbook of Persian Linguistics*. Ed. by Anousha Sedighi and Pounesh Shabani-Jadidi. Oxford: Oxford University Press. doi: [10.1093/oxfordhb/9780198736745.013.9](https://doi.org/10.1093/oxfordhb/9780198736745.013.9).
- Schlerath, Bernfried (1968). *Awesta-Wörterbuch: Vorarbeiten II. Konkordanz*. Wiesbaden: Harrassowitz.
- Schmidt, Gernot (1978). *Stammbildung und Flexion der indogermanischen Personalpronomina*. Wiesbaden: Harrassowitz.
- Schmitt, Rüdiger (1967). *Dichtung und Dichtersprache in indogermanischer Zeit*. Wiesbaden: Harrassowitz.
- (1986). *Das grammatische Geschlecht der Substantive des Altpersischen: die Beschreibung einer Sprache in Theorie und Praxis*. Ed. by Rüdiger Schmitt and Prods Oktor Skjærvø. Münchener Studien zur Sprachwissenschaft, Beiheft 13, Neue Folge. Kitzinger, pp. 311–326.
- (1989a). “Altpersisch”. In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 56–85.
- ed. (1989b). *Compendium Linguarum Iranicarum*. Wiesbaden: Dr. Ludwig Reichert.
- (2008). “Old Persian”. In: *The Ancient Languages of Asia and the Americas*. Ed. by Roger D. Woodard. Cambridge: Cambridge University Press, pp. 76–100. doi: [10.1017/CBO9780511486838.008](https://doi.org/10.1017/CBO9780511486838.008).
- (2018). “The lexicon of Indo-Iranian”. In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 3. Berlin/Boston: De Gruyter. Chap. 113, pp. 1942–1959. doi: <https://doi.org/10.1515/9783110542431-034>.

- Seiler, Hansjakob (1960). *Relativsatz, Attribut und Apposition: I. Teil: Relativsatz, Attribut und Apposition als Problem der allgemeinen Syntax; II. Teil: Die Relativkonstruktionen im Awestischen (mit Ausblicken auf andere indogermanische Sprachen)*. Wiesbaden: Harrassowitz.
- Sims-Williams, Nicholas (1979). "On the plural and dual in Sogdian". In: *Bulletin of the School of Oriental and African Studies* 42.2, pp. 337–346. doi: <https://doi.org/10.1017/S0041977X00145896>.
- (1982). "The double system of nominal inflexion in Sogdian". In: *Transactions of the Philological Society* 80.1, pp. 67–76. doi: <https://doi.org/10.1111/j.1467-968X.1982.tb01194.x>.
  - (1989). "Sogdian". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 173–192.
  - (1996). "Eastern Iranian Languages". In: *Encyclopædia Iranica*. Vol. 7. 6, pp. 649–652. URL: <https://www.iranicaonline.org/articles/eastern-iranian-languages>.
  - (2004). "The Bactrian Inscription of Rabatak: A New Reading". In: *Bulletin of the Asia Institute. New Series* 18, pp. 53–68. URL: <https://www.jstor.org/stable/24049141>.
  - (2007). *Bactrian Documents from Northern Afghanistan II: Letters and Buddhist Texts*. Studies in the Khalili Collection III, Corpus Inscriptionum Iranicarum II, Vol. III: Bactrian. Oxford: The Nour Foundation in Association with Azimuth Editions and Oxford University Press.
  - (2012). *Bactrian Documents from Northern Afghanistan I: Legal and Economic Documents (revised edition)*. Studies in the Khalili Collection III, Corpus Inscriptionum Iranicarum II, Vol. VI: Bactrian. Oxford: The Nour Foundation in Association with Azimuth Editions.
- Sims-Williams, Nicholas and Harry Falk (2014). "Kushan Dynasty II: Inscriptions of the Kushans". In: *Encyclopædia Iranica, online edition*. Costa Mesa, California. URL: <https://www.iranicaonline.org/articles/kushan-02-inscriptions>.
- Skjærvø, Prods Oktor (1989a). "Modern East Iranian Languages". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert Verlag, pp. 370–383.
- (1989b). "Pashto". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 384–410.
  - (2003). *An Introduction to Young Avestan*. 4th ed.
  - (2009a). "Middle West Iranian". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 196–278.
  - (2009b). "Old Iranian". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 43–195.

- (2017). “The morphology of Iranian”. In: *Handbook of Comparative and Historical Indo-European Linguistics*. Ed. by Jared Klein, Brian Joseph, and Matthias Fritz. Vol. 1. Berlin/Boston: De Gruyter, pp. 503–549. doi: <https://doi.org/10.1515/9783110261288-034>.
- Sleeman, Petra (2011). “Verbal and adjectival participles: Position and internal structure”. In: *Lingua* 121.10, pp. 1569–1587. doi: <https://doi.org/10.1016/j.lingua.2011.05.001>.
- Sokolov, Sergej Nikolaevič (1979a). “Drevnepersidskij Jazyk”. In: *Drevneiranske Jazyki*. Ed. by Vera S. Rastorgueva. Vol. 1. Osnovy Iranskogo Jazykoznanija. Moskva: Nauka, pp. 234–271.
- (1979b). “Jazyk Avesty”. In: *Drevneiranske Jazyki*. Ed. by Vera S. Rastorgueva. Vol. 1. Osnovy Iranskogo Jazykoznanija. Moskva: Nauka, pp. 129–233.
- Sommer, Ferdinand (1931). *Vergleichende Syntax der Schulsprachen (Deutsch, Englisch, Französisch, Griechisch, Lateinisch), mit besonderer Berücksichtigung des Deutschen*. 3. Auflage. Leipzig: Teubner.
- Speyer, Jacob Samuel (1896). *Vedische und Sanskrit-Syntax*. Grundriss der indo-arischen Philologie und Altertumskunde 6,1. Berlin, Boston: De Gruyter Mouton. doi: <https://doi.org/10.1515/9783111581842>.
- Steblin-Kamenskij, Ivan M. (1981). “Baktrijskij jazyk”. In: *Sredneiranske Jazyki*. Vol. II. Osnovy Iranskogo Jazykoznanija. Moskva: Nauka, pp. 314–346.
- (1999). *Ėtimologičeskij slovar' vaxanskogo jazyka*. Sankt-Peterburg: Vostokoved.
- Stilo, Donald L. (2005). “Iranian as Buffer Zone between the Universal Typologies of Turkic and Semitic”. In: *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. Ed. by Éva Ágnes Csató, Bo Isaksson, and Carina Jahani. London - New York: Routledge Curzon, pp. 35–63.
- (2012). “Intersection Zones, Overlapping Isoglosses, and ‘Fade-out/Fade-in’ Phenomena in Central Iran”. In: *Iranian Languages and Culture: Essays in Honor of Gernot Ludwig Windfuhr*. Ed. by Behrad Aghaei and M. R. Ghanoonparvar. Costa Mesa: Mazda Publishers, pp. 3–33.
- Sundermann, Werner (1989a). “Mittelpersisch”. In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 138–164.
- (1989b). “Parthisch”. In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 114–137.
- Szemerényi, Oswald (1990). *Einführung in die vergleichende Sprachwissenschaft*. 4., durchgesehene Auflage 1990. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Tegey, Habibullah and Barbara Robson (1996). *A Reference Grammar of Pashto*. Washington, DC: Center for Applied Linguistics.
- Thomason, Sarah Grey (2001). *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press. doi: <https://doi.org/10.1515/9781474473125>.



- Thomason, Sarah Grey and Terrence Kaufman (1988). *Language Contact, Creolization, and Genetic Inheritance*. Berkeley: University of California Press. doi: <https://doi.org/10.1525/9780520912793>.
- Thordarson, Fridrik (1989). "Ossetic". In: *Compendium Linguarum Iranicarum*. Ed. by Rüdiger Schmitt. Wiesbaden: Dr. Ludwig Reichert, pp. 456–479.
- Thurneysen, Rudolf (1892). "Zur Stellung des Verbuns im Altfranzösischen". In: *Zeitschrift für romanische Philologie* 16, pp. 289–307. doi: <https://doi.org/10.1515/zrph.1892.16.1-4.289>.
- Tichy, Eva (2009). *Indogermanistisches Grundwissen für Studierende sprachwissenschaftlicher Disziplinen*. 3., vollständig überarbeitete Auflage. Bremen: Hempen Verlag.
- Tremblay, Xavier (2005). "Bildeten die iranischen Sprachen ursprünglich eine genetische Familie oder einen Sprachbund innerhalb des indo-iranischen Zweiges? Beiträge zur vergleichenden Grammatik der iranischen Sprachen V". In: *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.-23. September 2000, Halle an der Saale*. Ed. by Olav Hackstein and Gerhard Meiser. Wiesbaden: Reichert, pp. 673–688.
- Truckenbrodt, Hubert (2007). "The syntax-phonology interface". In: *The Cambridge handbook of phonology*. Ed. by Paul de Lacy. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press, pp. 435–456. doi: 10.1017/CBO9780511486371.019.
- Uhlenbeck, Christianus C. (1901). "Agens und Patiens im Kasussystem der indogermanischen Sprachen". In: *Indogermanische Forschungen* 12, pp. 170–171.
- Vaillant, A. (1936). "L'ergatif indo-européen". In: *Bulletin de la Société de Linguistique de Paris* 37, pp. 93–108.
- Vajda, Edward J. (2019). "Yeniseian and Dene toponyms". In: *Language and Toponymy in Alaska and Beyond: Papers in Honor of James Kari*. Ed. by Gary Holton and Thomas Thornton. Language Documentation and Conservation Special Publication No. 17. Honolulu: University of Hawai'i Press, pp. 174–190. url: <https://scholarspace.manoa.hawaii.edu/handle/10125/24847>.
- Vendryes, Joseph (1918). "Les correspondances de vocabulaire entre l'indo-iranien et l'italo-celtique". In: *Mémoires de la Société de linguistique de Paris* 20, pp. 265–285.
- Vine, Brent (1998). "Remarks on the Archaic Latin "Garigliano Bowl" inscription". In: *Zeitschrift für Papyrologie und Epigraphik* 121, pp. 257–262. doi: <https://www.jstor.org/stable/20190222>.
- Viti, Carlotta (2013). "Forms and functions of subordination in Indo-European". In: *Historische Sprachforschung* 126, pp. 89–117. doi: <https://www.jstor.org/stable/43857937>.

- (2014). “Reconstructing Syntactic Variation in Proto-Indo-European”. In: *Indo-European Linguistics* 2.1, pp. 73–111. doi: <https://doi.org/10.1163/22125892-00201004>.
- (2015). “Historical syntax: problems, materials, methods, hypotheses”. In: *Perspectives on Historical Syntax*. Ed. by Carlotta Viti. Amsterdam: John Benjamins, pp. 3–31.
- von Mengden, Ferdinand (2008). “Reconstructing Complex Structures: A Typological Perspective”. In: *Principles of Syntactic Reconstruction*. Ed. by Gisella Ferraresi and Maria Goldbach. Current Issues in Linguistic Theory 302. Amsterdam: John Benjamins, pp. 97–119. doi: <https://doi.org/10.1075/cilt.302>.
- Wackernagel, Jacob (1892). “Ueber ein Gesetz der indogermanischen Wortstellung”. In: *Indogermanische Forschungen* 1, pp. 333–436. doi: <https://doi.org/10.1515/9783110242430.333>.
- (2009). *Lectures on Syntax. With special reference to Greek, Latin, and Germanic. Edited with notes and bibliography by David Langslow*. Oxford: Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780198153023.001.0001>.
- Walkden, George (2013). “The correspondence problem in syntactic reconstruction”. In: *Diachronica* 30.1, pp. 95–122. doi: <https://doi.org/10.1075/dia.30.1.04wal>.
- (2014). *Syntactic Reconstruction and Proto-Germanic*. Oxford: Oxford University Press.
- Watkins, Calvert (1964). “Preliminaries to the Reconstruction of Indo-European Sentence Structure”. In: *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., August 27–31, 1962*. Ed. by Horace G. Lunt. Berlin, Boston: De Gruyter Mouton, pp. 1035–1045. doi: <https://doi.org/10.1515/9783112317419-090>.
- (1976). “Towards Proto-Indo-European Syntax: Problems and Pseudo-Problems”. In: *Papers from the Parasession on Diachronic Syntax, April 22, 1976*. Ed. by Sanford B. Steever, Carol A. Walker, and Salikoko S. Mufwene. Chicago: Chicago Linguistic Society, pp. 305–326.
- (1995). *How to Kill a Dragon. Aspects of Indo-European Poetics*. New York; Oxford: Oxford University Press.
- Weiss, Michael (2020). *Outline of the Historical and Comparative Grammar of Latin. Second Edition*. Ann Arbor/New York: Beech Stave Press.
- Wemhoff, Matthias and Herlinde Koelbl, eds. (2018). *Margiana: Ein Königreich der Bronzezeit in Turkmenistan*. Petersberg: Michael Imhof Verlag.
- Wendtland, Antje (2011). *Die Entwicklung von Demonstrativpronomen zu Artikeln im Soghdischen*. Göttinger Orientforschungen, Iranica, Neue Folge 8. Wiesbaden: Harrassowitz.

- West, Martin L. (2011). *Old Avestan Syntax and Stylistics*. Abhandlungen der Akademie der Wissenschaften zu Göttingen, Neue Folge, Band 13. Berlin, Boston: De Gruyter. doi: <https://doi.org/10.1515/9783110253092>.
- Willi, Andreas (2018). *Origins of the Greek Verb*. Cambridge: Cambridge University Press. doi: <https://doi.org/10.1017/9781108164207>.
- Willim, Ewa (2000). "On the grammar of Polish nominals". In: *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Ed. by Roger Martin, David Michaels, and Juan Uriagereka. Cambridge: MIT Press, pp. 319–346.
- Willis, David (2011). "Reconstructing last week's weather: Syntactic reconstruction and Brythonic free relatives". In: *Journal of Linguistics* 47, pp. 407–446. doi: 10.1017/S0022226710000381.
- Windfuhr, Gernot (2009a). "Dialectology and Topics". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 5–42.
- ed. (2009b). *The Iranian Languages*. London/New York: Routledge.
- Windfuhr, Gernot and John R. Perry (2009). "Persian and Tajik". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 416–544.
- Winter, Werner (1984). "Reconstructional comparative linguistics and the reconstruction of the syntax of Undocumented stages in the development of languages and language families". In: *Historical Syntax*. Ed. by Jacek Fisiak. Trends in Linguistics. Studies and Monographs 23. Berlin, New York: De Gruyter Mouton, pp. 613–625. doi: <https://doi.org/10.1515/9783110824032.613>.
- Witzel, Michael (1999). "Substrate Languages in Old Indo-Aryan (Ṛgvedic, Middle and Late Vedic)". In: *Electronic Journal of Vedic Studies* 5.1, pp. 1–67. url: <https://crossasia-journals.ub.uni-heidelberg.de/index.php/ejvs/article/viewFile/828/806>.
- (2001). "Autochthonous Aryans? The Evidence from Old Indian and Iranian Texts". In: *Electronic Journal of Vedic Studies* 7.3, pp. 1–118. doi: 10.11588/ejvs.2001.3.830.
- (2003). "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia". In: *Sino-Platonic Papers* 129, pp. 1–70. url: [http://www.sino-platonic.org/complete/spp129\\_prehistoric\\_central\\_asia\\_linguistics.pdf](http://www.sino-platonic.org/complete/spp129_prehistoric_central_asia_linguistics.pdf).
- Witzlack-Makarevich, Alena and Ilja A. Seržant (2018). "Differential argument marking: Patterns of variation". In: *Diachrony of differential argument marking*. Ed. by Ilja A. Seržant and Alena Witzlack-Makarevich. Studies in Diversity Linguistics 19. Berlin: Language Science Press, pp. 1–40. doi: <http://dx.doi.org/10.5281/zenodo.1228243>.

- Wood, Sandra K., Diane C. Lillo-Martin, and Joseph C. Hill (2018). "Homesign systems". In: *Sign Languages: Structures and Contexts*. Ed. by Joseph C. Hill, Diane C. Lillo-Martin, and Sandra K. Wood. London: Routledge.
- Xromov, Al'bert L. (1987). *Ĵagnobskij Ĵazyk*. Osnovy iranskogo jazykoznanija, novoiran-skie jazyki: vostočnaja gruppa. Moskva: Nauka, pp. 644–701.
- Yakubovich, Ilya (2020). "Persian ezāfe as a contact-induced feature". In: *Voprosy jazykoz-nanija / Topics in the Study of Language* 5, pp. 91–114. doi: <https://doi.org/10.31857/0373-658X.2020.5.91-114>.
- Yoshida, Yutaka (2009). "Sogdian". In: *The Iranian Languages*. Ed. by Gernot Windfuhr. Routledge Family Series. London & New York: Routledge, pp. 279–335.
- Zekī Walidī, Ahmed (1930). "Ĥwāresmische Sätze in einem arabischen *Fiqh*-Werke". In: *Islamica* 3, pp. 190–213. url: [https://resolver.sub.uni-goettingen.de/purl?PPN864009399\\_0003](https://resolver.sub.uni-goettingen.de/purl?PPN864009399_0003).