

# Grammaticalization and the Rise of Configurationality in Indo-Aryan

---

OXFORD STUDIES IN DIACHRONIC AND HISTORICAL LINGUISTICS

GENERAL EDITORS

Adam Ledgeway and Ian Roberts, University of Cambridge

ADVISORY EDITORS

Cynthia Allen, *Australian National University*; Ricardo Bermúdez-Otero, *University of Manchester*; Theresa Biberauer, *University of Cambridge*; Charlotte Galves, *University of Campinas*; Geoff Horrocks, *University of Cambridge*; Paul Kiparsky, *Stanford University*; Anthony Kroch, *University of Pennsylvania*; David Lightfoot, *Georgetown University*; Giuseppe Longobardi, *University of York*; David Willis, *University of Cambridge*

RECENTLY PUBLISHED IN THE SERIES

14

*Arabic Indefinites, Interrogatives, and Negators*

*A Linguistic History of Western Dialects*

David Wilmsen

15

*Syntax over Time*

*Lexical, Morphological, and Information-Structural Interactions*

*Edited by Theresa Biberauer and George Walkden*

16

*Syllable and Segment in Latin*

Ranjan Sen

17

*Participles in Rigvedic Sanskrit*

*The Syntax and Semantics of Adjectival Verb Forms*

John J. Lowe

18

*Verb Movement and Clause Structure in Old Romanian*

Virginia Hill and Gabriela Alboiu

19

*The Syntax of Old Romanian*

*Edited by Gabriela Pană Dindelegan*

20

*Grammaticalization and the Rise of Configurationality in Indo-Aryan*

Uta Reinöhl

21

*The Rise and Fall of Ergativity in Aramaic*

*Cycles of Alignment Change*

Eleanor Coghill

# Grammaticalization and the Rise of Configurationality in Indo-Aryan

---

UTA REINÖHL

OXFORD  
UNIVERSITY PRESS

# OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,  
United Kingdom

Oxford University Press is a department of the University of Oxford.  
It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide. Oxford is a registered trade mark of  
Oxford University Press in the UK and in certain other countries

© Uta Reinöhl 2016

The moral rights of the author have been asserted

First Edition published in 2016

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in  
a retrieval system, or transmitted, in any form or by any means, without the  
prior permission in writing of Oxford University Press, or as expressly permitted  
by law, by licence or under terms agreed with the appropriate reprographics  
rights organization. Enquiries concerning reproduction outside the scope of the  
above should be sent to the Rights Department, Oxford University Press, at the  
address above

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data  
Data available

Library of Congress Control Number: 2015950970

ISBN 978-0-19-873666-0

Printed in Great Britain by  
Clays Ltd, St Ives plc

Links to third party websites are provided by Oxford in good faith and  
for information only. Oxford disclaims any responsibility for the materials  
contained in any third party website referenced in this work.

## *Series preface*

Modern diachronic linguistics has important contacts with other subdisciplines, notably first-language acquisition, learnability theory, computational linguistics, sociolinguistics, and the traditional philological study of texts. It is now recognized in the wider field that diachronic linguistics can make a novel contribution to linguistic theory, to historical linguistics, and arguably to cognitive science more widely.

This series provides a forum for work in both diachronic and historical linguistics, including work on change in grammar, sound, and meaning within and across languages; synchronic studies of languages in the past; and descriptive histories of one or more languages. It is intended to reflect and encourage the links between these subjects and fields such as those mentioned above.

The goal of the series is to publish high-quality monographs and collections of papers in diachronic linguistics generally, i.e. studies focusing on change in linguistic structure, and/or change in grammars, which are also intended to make a contribution to linguistic theory, by developing and adopting a current theoretical model, by raising wider questions concerning the nature of language change, or by developing theoretical connections with other areas of linguistics and cognitive science as listed above. There is no bias towards a particular language or language family, or towards a particular theoretical framework; work in all theoretical frameworks, and work based on the descriptive tradition of language typology, as well as quantitatively based work using theoretical ideas, feature in the series.

Adam Ledgeway and Ian Roberts

*University of Cambridge*



# Acknowledgements

First and foremost, I thank Nikolaus Himmelmann for teaching me most of what I know about languages, for each and every critical remark, for his patience and for his trust. I could not have wished for a better supervisor. I am also immensely grateful to Hans Henrich Hock who agreed to be my second doctoral advisor at a very late stage in my writing of this thesis and for his many and most valuable comments. I also want to thank him for making the long trip from Urbana to Cologne to see me defend this dissertation.

I sincerely want to thank Volker Unterladstetter for sharing with me ideas and criticism alike, growing up together in linguistics, and particularly for proofreading this entire book, some chapters in more than one version. I also thank Felix Rau, Daniel Kölligan, and Birgit Hellwig for their very valuable comments on individual chapters and wish furthermore to express my sincere gratitude to two anonymous reviewers and the series editors for their very helpful comments and suggestions. I am very grateful too to Else Gellinek for polishing my English, to Amelie Bernzen, Robert Tegetenmeyer, and Sven Minnerup for corrections, and to Kolja Kreutz for always being ready to help. I would also particularly like to thank Adam Ledgeway and Ian Roberts, as well as Julia Steer at OUP, for including this study in their series.

I have greatly benefited over the years from several teachers of Sanskrit, Pali, and Hindi, among others Sebastian Hopp, Pratik Rumde, Sadananda Das, and my friends in Delhi. I could not have written this study without the unlimited access to the Indology department at Bonn where Sebastian Hopp generously allowed me to work in his office for over half a year. In the context of my research for this thesis, I am especially grateful to Eva De Clercq for helping me understand the Apabhrāmṣa data.

I am highly grateful for the support of scholarships from the *Studienstiftung des Deutschen Volkes*, the a.r.t.e.s. *Graduate School for the Humanities Cologne*, as well as the *Deutscher Akademischer Austauschdienst*. The DAAD blessed me with five blissful months in Delhi at the Jawaharlal Nehru University, the memory of which will stay with me for a long time. I am deeply thankful for positions that I have held at the General Linguistics department at Cologne.

My colleagues and friends at the General Linguistics department in Cologne have made this place a home for me, an environment in which it was a pleasure to write this book. In particular, these include Volker Unterladstetter, Ingeborg Fink, and Sarah Dolscheid. I thank them from my heart for innumerable happy hours. I also consider myself very lucky to have received advice from friends and mentors like Jörg Dötsch and Hana Filip.

I want to conclude by thanking those I can always come home to for giving me the foundation that allows me to tackle tasks such as writing a thesis: Saket Prakash, Fränz Conrad, and my family. In particular I wish to express my deepest gratitude to my parents for their constant support and unlimited generosity.



# *Abbreviations*

## **Glosses**

ABL	ablative
ACC	accusative
ADP	adposition
ADV	adverb
AOR	aorist
AUX	auxiliary
CONV	converb
DAT	dative
DEF	definite
DEM	demonstrative pronoun
DIR	direct case
DU	dual
EMPH	emphatic particle
ERG	ergative
EVID	evidential
EXCL	exclamatory element
F	feminine
FOC	focus particle
FUT	future tense
GEN	genitive
IMP	imperative
IMPF	imperfect
INDEF	indefinite pronoun
INF	infinitive
INJ	injunctive
INS	instrumental
INT	intensive
IPFV	imperfective
LOC	locative
LP	local particle
M	masculine

MID	middle
N	neuter
NEG	negation marker
NOM	nominative
NONPAST	non-past
OBL	oblique case
OPT	optative
PASS	passive
PFV	perfective
PL	plural
POSS	possessive
PPA	participle present active
PPM	participle present middle
PPP	participle perfective passive
PRF	perfect
PPRM	participle perfective medium
PRO	pronoun
PROG	progressive
PRS	present
PRT	discourse particle
PST	past
Q	question particle
QUOT	quotative particle
REFL	reflexive
REL	relative pronoun
SEQ	sequentializer
SG	singular
SUBJ	subjunctive
VOC	vocative

## Primary sources

### *Vedic Sanskrit:*

AiB	Aitareyabrāhmaṇa
AV	Atharvaveda (Saṃhitā)

ṚV	Rigveda <sup>1</sup> (Saṃhitā)
YV	Yajurveda (Saṃhitā)
ŚaB	Śatapathabrāhmaṇa
ŚāñB	Śāṅkhāyanabrāhmaṇa

*Pali:*

Aṅg	Aṅguttaranikāya
Apa	Apadāna
Buddha	Buddhavaṃsa
Cari	Cariyāpitaka
Dhamm	Dhammapada
Dīgh	Dīghanikāya
Iti	Itivuttaka
Jā	Jātaka
Majjh	Majjhimanikāya
Mili	Milindapañha
Nidd	Niddesa
Paṭi	Paṭisambhidāmagga
Peta	Petavatthu
Sam	Samyuttanikāya
Sutta	Suttanipāta
Thera	Theragāthā
Therī	Therīgāthā
Ud	Udāna
Vim	Vimānavatthu
Vina	Vinayapiṭaka

*Apabhramsha:*

BH	Bhavisattakaha
KA	Karakaṇḍacariu
PA	Paumacariu
SA	Sanatkumāracaritam

*Old Awadhi:*

RA	Rāmacaritamānasa
----	------------------

<sup>1</sup> The spelling *Rigveda*/Rigvedic rather than *Ṛgveda*/Ṛgvedic is used in the interest of discoverability.



# Introduction

The goal of this study is to offer insights into the way in which a non-configurational language may acquire a certain degree of configurationality over the course of time. ‘Non-configurational’ is used here as it was first defined in work by Kenneth Hale and other Australianists, describing a language that lacks phrase structure on any clausal level (besides certain further characteristics that are outlined below). To study the rise of configurational structure is—in essence—to study how obligatoriness and word order constraints can emerge as syntactic principles, shaping different types of expressions into phrasal units. The branch of languages chosen for this investigation, Indo-Aryan, provides a unique, and perhaps the only, opportunity to study this change in detail. In the three millennia of attested Indo-Aryan history, the grammatical system has developed from a stage without any phrasal organization in Vedic Sanskrit to the existence of postpositional phrases in modern Indo-Aryan languages such as Hindi.<sup>1</sup> In addition, the history of Indo-Aryan is attested without substantial gaps, so that it is possible to trace the development over successive historical periods.

The development from non-configurational to low-level configurational in Indo-Aryan will be argued to have been brought about as a by-product of the grammaticalization of certain function words (cp. Himmelmann 1997 on the rise of noun phrases as a by-product of the grammaticalization of articles). Vedic Sanskrit and other Old and Middle Indic varieties do not possess function words that organize expressions in a phrasal manner. By contrast, New Indo-Aryan languages display low-level configurationality, possessing postpositions that organize nominal expressions as phrases, that is, as syntagmatically fixed strings with slots for specific and in part obligatory elements.<sup>2</sup> I show how semantic changes, both prior to and during the

<sup>1</sup> In this study, I will make reference only to Hindi, even though most statements also extend to Urdu. The literature I refer to and the elicitations I gathered, however, all relate to Hindi as spoken in the Indian Hindi belt, so that this restriction seems appropriate. I will make special mention of Urdu only in such cases where authors relate to a form in existence before the sociolectal separation of Hindi and Urdu, or want to emphasize that a certain variant is found in today’s Urdu speaking areas.

<sup>2</sup> The term ‘postpositions’ as used here denotes New Indo-Aryan elements such as Hindi *ko*, *ne*, *mē*, *par* etc. that mark grammatical and semantic roles of nominal expressions. In historical perspective, these elements form a middle layer between the old inflectional remnants on the one hand, and younger ‘complex postpositions’ that consist of a possessive linker and typically an oblique noun on the other

grammaticalization of certain lexemes into the modern postpositions, entailed changes on the expression side, which eventually resulted in the formation of phrases.<sup>3</sup> The phrasal template, once in place, also expanded to nominal expressions that lack postpositions, even though no other function words (i.e. articles) have developed in the nominal domain. This study, then, is an exploration into the way in which changes on the micro level of individual grammaticalization phenomena bear on the macro level of a language's syntactic system.

In order to gain insights into the mechanisms of the rise of (low-level) configurationality, I trace in the course of this book the emergence of the modern Hindi simple postpositions from their earliest attestations in Vedic Sanskrit onwards—a scope that emerges from the research question. The beginnings of the development of the modern phrases are rooted in Vedic Sanskrit, but it is not before the transition from Middle to New Indo-Aryan that the postpositional phrases are in place in the way we know them from contemporary Indo-Aryan languages. Two of the postpositions can be traced back especially far, namely Hindi *mē* 'in' from Old Indic *madhye* 'in the middle' and *par* 'on' from Old Indic *upari* 'above'. A corpus of their attestations across four stages of Indo-Aryan, namely Vedic Sanskrit (around 1000 BC), Pali (around 500 BC), Apabhramsha (8th–13th centuries AD) and early Hindi (i.e. Old Awadhi from the 16th century), constitutes the empirical backbone of this study. This is complemented by selected evidence from further historical varieties as well as by their syntax and semantics in contemporary Hindi. The histories of the remaining simple postpositions are included as far as they can be reconstructed from the available sources. In analysing these data, I outline the different source constructions in which the respective lexemes grammaticalized, and the factors that turned these source constructions into phrasal units, as well as interactions between the various elements in the process of grammaticalizing. That interactions occurred is suggested by the fact that the postpositions go back to elements of several different syntactic classes, but today form a functionally and formally homogeneous paradigm.

The development of configurationality in Indo-Aryan is shown in this book to differ on a fundamental, structural level from that in more western branches of Indo-European. In fact, the study of Indo-Aryan reveals an origin of phrasality hitherto unattested not only in Indo-European, but cross-linguistically, as far as I am aware. While it is a circumscribed class of adverbial elements, the so-called local particles, which developed into the simple adpositions in western branches of the family, the Indo-Aryan elements are of an altogether different origin. I show how

hand, analogous to expressions such as English *on top of*, *in front of*, etc. I do not use the term 'postpositions' with reference to the local particles of Old Indo-Aryan (e.g. *ā*, *anu*, *pari* etc.), nor with reference to any other elements in the history of this branch of languages that may have been called by this name. A detailed introduction to the postpositions follows in Chapter 3.

<sup>3</sup> I will speak of *semantic* changes inducing phrasality, but note that this in part includes also certain pragmatic effects, as will be discussed in Chapter 7.

we are not only dealing with elements of etymological sources unrelated to the local particles, but that these elements in origin also had very different semantic and syntactic properties and consequently attest to an alternative path of grammaticalization. In western branches, the adverbial local particles were reanalysed as adpositions in combination with syntactically co-ranking local case forms. In Indo-Aryan, by contrast, nominal(ized) and participial heads with different types of dependents (especially adnominal genitives) are at the origin of the modern adpositions. Thus, on the one hand, we are dealing with an origin of phrasal structures in what I will refer to as ‘symmetrical groups’ of syntactically co-ranking and semantically co-referential elements in western branches, while on the other hand, we observe an origin in ‘asymmetrical groups’ of elements that are neither co-ranking nor co-referential, but are characterized by a semantic and formal head-dependent relationship in Indo-Aryan.

### 1.1 Contextualization

The syntax of nominal expressions in Vedic Sanskrit and in Hindi differs in fundamental aspects (a detailed outline is given in section 2.4). Vedic Sanskrit can be grouped with non-configurational languages of the dependent-marking type, similar to Australian Jiwari (Austin 2001), for instance.<sup>4</sup> With respect to the syntax of nominal expressions, word order is largely unconstrained and nouns, adjectives, demonstratives, and other elements, all case-marked, stand in relations of apposition,<sup>5</sup> no element being obligatory or clearly super- or subordinate. Vedic Sanskrit also generally lacks a developed apparatus of syntactic constraints relating to linear order or to the obligatoriness of elements. In Hindi, by contrast, complex nominal expressions are realized as contiguous strings (with a few, constrained exceptions) that adhere to a fixed template (simplified) DEM QUANT MOD NOM (POSTP) with specific internal morphosyntactic rules of government (if a postposition is present) and of agreement. Not every element in a nominal expression is case-marked;<sup>6</sup> postpositions, if present, occur at the right edge and have phrasal scope. The following two

<sup>4</sup> Subjects are marked by agreement on the verb, so that there is a degree of head-marking, too.

<sup>5</sup> I use ‘apposition’ for lack of a better term, designating combinations of nominal elements (in a wide sense, i.e. all elements that allow for case marking) that are co-referential, but where no clear relation of headedness can be identified, all elements having the same syntactic potential. This term is also found in other writings on non-configurationality, e.g. Austin and Bresnan (1996) on Australian non-configurational languages, or Luraghi (2010) on Indo-European. Note, however, that this usage of the term by far exceeds the traditional one, as the constructions in question involve a much greater variety of semantic types. Traditionally, constructions such as *King George* or *Paul, the steward* in English have been described as ‘appositions’ (cp. Bloomfield 1933: 186).

<sup>6</sup> For the purposes of the present study, I adopt a loose understanding of ‘case’, covering the expression of grammatical as well as semantic roles of nominal expressions, by morphological means or by adpositions. For details see sections 2.4.2 and 8.3.2.2.

examples illustrate these differences. Both contain complex nominal expressions, but whereas co-referential elements do not have to stand in adjacency and separately receive full morphological marking in Vedic (co-reference is indicated by subscripts), they form an integral syntagm with shared postpositional marking in Hindi:

(1) VEDIC SANSKRIT

<i>kāti<sub>i</sub></i>	<i>ayám<sub>j</sub></i>	<i>adyá</i>	<i>udgātā<sub>j</sub></i>	<i>asmin<sub>k</sub></i>
how_many	DEM.NOM.SG.M	today	Udgātṛ.NOM.SG.M	DEM.LOC.SG.M
<i>yajñē<sub>k</sub></i>	<i>stotriyāḥ<sub>i</sub></i>		<i>stoṣyati?</i>	
sacrifice.LOC.SG.M	stotriya.ACC.PL.M		sing.FUT.3SG	
'How many stotriyas will this Udgātṛ sing in this sacrifice today?'				
(ŚaB 14.4.6.12, cp. Schäufele 1991b: 76)				

(2) HINDI

<i>us</i>	<i>ādmī</i>	<i>ne</i>	<i>choṭe</i>	<i>bacce</i>	<i>ko</i>
DEM.OBL.SG	man.OBL.SG.M	ERG	small.OBL.SG.M	child.OBL.SG.M	DAT
<i>kitāb</i>	<i>bhejī.</i>				
book.DIR.SG.F	send.PFV.SG.F				
'That man sent the book to the little child.'					

Often, differences such as these have been framed by means of a binary 'parameter'—either a language is non-configurational, or it is configurational. With a somewhat different focus, languages like Vedic Sanskrit have been considered 'free word order' languages, while languages like Hindi are referred to as 'free phrase order' or 'scrambling' languages (Schäufele 1991a). What precisely defines 'non-configurational' or 'configurational' has been discussed controversially over the years, as is outlined in detail in section 2.2. For now, I focus on the preliminary distinction, that a non-configurational language lacks any type of phrasality, that is, elements do not have to appear in specific positions, even when co-referential, and do not have to be overtly expressed, whereas a configurational language has phrasal structures.

In this study, I conceive of configurationality as a historical phenomenon, subject to change and non-universal. For expository purposes, I speak of Vedic as being non-configurational and of Indo-Aryan as developing low-level configurationality in the course of its history, that is, phrasality on the level of nominal expressions (but not on the level of verbal expressions, as in strongly configurational English, for instance). The transition, however, will be shown to be non-discrete. Note that, in Indo-Aryan, not only do we witness the rise of a new phrasal category in a branch of languages, but we see the very initial rise of phrasality, that is, the emergence of syntactic word order constraints (rather than word order distinctions operating only on the level of information structure), as well as the obligatorification of elements. Thus, this is the special case of a non-configurational language becoming



configurational to a certain degree, rather than of an already configurational language becoming more configurational. Several other properties that have been connected to non-configurational languages, such as a weak noun–adjective distinction, will be shown to be affected by the rise of phrasal structures. Accordingly, I cast this study not only in the context of a rise of phrasality, but in the larger context of configurationality as a historical phenomenon.

Studies approaching configurationality or specifically phrase structure from a historical angle (whether conceiving of it as a gradual or as a discrete phenomenon) are few, but a handful of treatments do exist. This includes Vincent (1999) on the emergence and development of PPs in Latin and Romance, van de Velde (2010) on the formation of the determiner slot in Dutch, Himmelmann (1997) on the cross-linguistic rise of article systems, Hewson and Bubeník (2006) on the rise of adpositional phrases in the major branches of Indo-European, and Luraghi (2010) on the rise of configurationality in Indo-European. The present study owes a great deal to all of these works, especially to Himmelmann (1997), where the grammaticalization of articles is connected with the rise of nominal phrases:

The central hypothesis is that syntactic structure is just as much a result of grammaticalization processes as are grammatical elements, i.e. that in the course of such a process, not only do article-like elements emerge, but the characteristic syntactic categories and constituent structures emerge along with them. (Himmelmann 1997: 1, translation UR)<sup>7</sup>

At the same time, I aim to add insights in several domains that have not yet been treated in detail, as far as I am aware. First, I focus on the factors that ultimately trigger the development of constructions into phrasal units. I argue that these factors are semantic ones, and I propose a detailed scenario of how the semantics–syntax interface is crossed. Second, on an empirical level, it will become clear in the course of this book that the constructions that are at the origin of postpositional phrases in Indo-Aryan differ structurally not only from the origin of adpositional phrases in other branches of Indo-European, but from all origins of the phrasal types discussed in the above-mentioned studies, as they originate in asymmetrical, rather than symmetrical groups.

Let us take a closer look at the types of grammaticalization phenomena we are dealing with in this study. Consider the following examples, again contrasting Vedic Sanskrit with Hindi:

<sup>7</sup> ‘Die zentrale Hypothese besagt, daß syntaktische Struktur ebenso Ergebnis von Grammatikalisierungsprozessen ist wie grammatische Elemente, daß also in einem solchen Prozeß nicht nur die artikelähnlichen Elemente entstehen, sondern zugleich auch die für nominale Ausdrücke charakteristischen syntaktischen Kategorien und Konstituentenstrukturen.’

## (3) VEDIC SANSKRIT

*átha somopanáhanasya samutpāryāntān*  
*átha somopanáhanasya samutpārya ántān*  
 now Soma\_cloth.GEN.SG.N gather.CONV end.ACC.PL.M  
*uṣṇíṣeṇa vīgrathnāti... atha madhye 'ṅgūlyākāśaṃ karoti*  
 head\_band.INS.SG tie\_together.3SG now MADHYE finger\_hole.ACC.SG.M do.3SG  
 'Having gathered up the ends of the Soma-cloth, he ties them together with a  
 head-band... he then makes a finger-hole in the middle (of the knot).'

(ŚaB 3.3.2.19–20)

## (4) VEDIC SANSKRIT

*devasenānām abhibhañjatinām jáyantīnām*  
 army\_of\_gods.GEN.PL.F smash.PPA.GEN.PL.F conquer.PPA.GEN.PL.F  
*marúto yantu mādhye*  
 Marut.NOM.PL.M go.IMP.3PL MADHYE  
 'In the middle of the smashing conquering armies of the gods, the Maruts  
 should go.'

(AV 19.13.9)

## (5) HINDI

*hamār-e ghar ke pās nadī hai.*  
 POSS.1PL-OBL.SG.M house.OBL.SG POSS proximity river.DIR.SG.F be.3SG  
*us /\*Ø mẽ tairnā acchā lagtā hai.*  
 DEM.OBL.SG in swim.INF.DIR good.DIR.SG.M strike.IPFV.SG.M be.3SG  
 'There is a river near our house; it's nice to swim in it.'

In examples (3) and (4), we see an element *madhye*, the locative singular of *madhya*- 'middle' and the ancestor of the Hindi postposition *mē* 'in'. In Old Indic, it had the concrete lexical meaning 'in the middle' and could occur on its own whenever a spatio-temporal ground<sup>8</sup> could be inferred, or with genitive possessors that would specify the respective ground, as well as with other case forms, as we will see in Chapter 5. In example (5), we see an example of *mē* in Hindi which now means 'in' and requires an oblique which specifies the ground. How do we get from *madhye* to *mē* and to phrasality?

Grammaticalization phenomena such as this one (see section 2.1 for a detailed introduction), where a spatial noun develops into an adposition, have been amply

<sup>8</sup> Note that this usage of the term 'ground' deviates slightly from that widespread in the tradition of Leonard Talmy's (1972) work. Typically, this term is used to refer to entire spatial expressions locating an object (or 'figure') somewhere. For example, in an utterance *He is standing in the middle of the room*, *He* is the figure and *in the middle of the room* is the ground. By contrast, I will be predominantly interested in semantic relations internal to spatial expressions. Thus, *the room* will be called the ground of the spatial relation *in the middle*. Such alternative terms as for instance 'reference object' or 'frame of reference' are infelicitous, the former suggesting a concrete object, the latter being inadequate for a number of examples that will be discussed. For example, in an utterance *above a tree*, the tree does not fit the idea of a frame very well.

studied in the literature. Besides spatial nouns, we will be dealing with adverbs and participles grammaticalizing into postpositions in the course of this study. Semantic and formal changes involved in the grammaticalization of adpositions are described for instance in Kahr (1975, 1976), Lehmann (2002 [1982]), Svorou (1988), or Rubba (1994). Highly valuable information on the history of Hindi and other New Indo-Aryan simple postpositions can be found in classic works such as Beames (1970 [1875]), Kellogg (1972 [1875]), or Chatterji (1970 [1926]).

On a theoretical level, I approach grammaticalization following proposals by Bybee et al. (1994) and Himmelmann (1997, 2005), who stress that grammaticalization always pertains to a construction, rather than to a lexeme. For example, it is not the element *go* that has grammaticalized into a future marker of English, but *go* embedded in the construction *be \_ing to VP*. However, following this ‘construction-based’ approach—which is supported by a vast amount of evidence and generally accepted today, see section 2.1—poses a certain difficulty for the case at hand: Where does the construction come from when we are dealing with a non-configurational language? How could *madhye* and other elements develop into elements that obligatorily take dependents, when the latter used to be optional? Why has *madhye*, along with the other lexemes growing into postpositions in Indo-Aryan, grammaticalized in a specific position relative to its dependent, even though the reversed order would have been equally available in the prior linguistic stage? Such syntactic aspects of the development have rarely been focused on<sup>9</sup>—perhaps because most languages already have phrasal structures. I hold that it is not enough to observe that an already frequent pattern (e.g. a certain word order) becomes obligatory in the course of grammaticalization, a statement regularly encountered (see section 2.6 for discussion). Rather, the fixation of patterns, both pertaining to the obligatory overt expression of elements and to word order, that is the very development of a construction into a phrasal unit, must be accounted for, no matter how frequent a certain pattern had already been beforehand. As indicated, I will propose that semantic changes are at the origin of the formation of phrasal units. Metaphorical and metonymic changes induce the obligatorification of dependents (e.g. of adnominal genitives in the case of relational nouns such as *madhye*), as well as the fixation of word order, since the central shift from lexical to grammatical only takes place in a specific syntagm, as will be outlined in detail in Chapter 7.

In addition, the grammaticalization of the individual postpositions will be shown in Chapter 8 to have had an effect on the general syntax of nominal expressions. Grammaticalization in direct adjacency to referentially used nominal elements leads to the fixation of slots for nouns in contrast to adnominal elements, strengthening the noun–adjective distinction and giving rise to a hierarchical

<sup>9</sup> Lehmann (2002 [1982]) touches on these matters, if not from an explicitly syntactic perspective; see section 2.6 below.

phrase. This pattern, while triggered through the grammaticalization of the postpositions, will be shown to have expanded to such nominal expressions where no postposition is present. Further, new elements, such as ergative *ne* as well as the direct form (the syncretic successor of nominative and accusative), have become integrated into the paradigm in recent centuries, so that all nominal expressions are now structured on the basis of the postpositional phrasal template and only a few niches remain where remnants of the old inflectional case system retain functional contrastiveness.

Before I proceed to outlining the empirical foundations of this study, a note is in order on prior approaches to the historical syntax of case in Indo-European, including Indo-Aryan. Often, this field is treated in the context of a general change from predominantly synthetic strategies of case marking to predominantly analytical ones, which affected Indo-Aryan in the same way as most major branches of Indo-European. In the history of this family, the morphological marking of nominal and verbal categories has been replaced in many domains by analytical constructions, involving adpositions in the nominal domain and auxiliaries in the verbal domain. This reorganization has been subject to a large number of studies since the nineteenth century and remains in the focus of research as illustrated by the recent edited volume *From Case to Adposition* by Hewson and Bubeník from 2006. Both pull chain and push chain scenarios have been argued for in the past (see Harris 1978 for an early overview of arguments that have been repeated since then; Harris extends the issue to changes in word order). Today, pull chain scenarios, where phonological erosion is said to trigger the recruitment of new forms, are generally rejected, because of the logical flaw that one would have to assume a linguistic stage functionally deficient in some sense. Push chain scenarios, by contrast, continue to be argued for, where the grammaticalization of new function words makes the old morphological cases superfluous and so they erode further or even vanish completely (e.g. Vincent 1999; Hewson and Bubeník 2006). This is especially plausible in the domain of case marking, because the old morphological cases of many ancient Indo-European varieties were not superseded by the new adpositional constructions, but fed into them. After reanalysis of the local particles as adpositions, the functional load was on the adposition, making the case morphology on the nominal superfluous.<sup>10</sup> For instance, while Latin *ad* governs the accusative in *ad Romam*, French *Rome* in *à Rome* lacks inflectional marking, which does not impede the functional role of the whole expression. As pointed out, however, the situation in Indo-Aryan differs from other branches of Indo-European.<sup>11</sup> Instead of local case forms, (typically)

<sup>10</sup> Exceptional cases exist, where an adposition governs two different cases with different meanings, e.g. Latin *in* occurring with both the accusative and the ablative.

<sup>11</sup> A treatment of the changes in Indo-Aryan as paralleling those in other branches of Indo-European, where the modern postpositional phrases are said to also go back to combinations of adverbs and local case

genitive-marked possessors of nominal and participial heads form part of the source constructions, as in ex. (4). Accordingly, there were no local case forms involved which could have become redundant in constructions headed by postpositions, rendering the oft-cited explanation for the loss of morphological case in Indo-European obsolete for Indo-Aryan. This is another way in which a depiction of the historical developments in Indo-Aryan bears on the theory of phrase structure generally, and on the rise of adpositional phrases in Indo-European specifically, showing us alternative pathways of change. In general, the developments in other branches of Indo-European serve as an important foundation for this study, but also as a contrastive backdrop, which highlights the structural differences regarding ways in which phrase structure may emerge.

## 1.2 The corpus

The attested history of Indo-Aryan is an immensely rich source of data and simultaneously poses serious difficulties for a linguistic analysis. In this section, I will outline this source, point out its difficulties, and present my selection of historical varieties, which is guided by considerations of how to cope best with any challenges that may arise (section 1.2.1). Subsequently, I present the texts taken into consideration (section 1.2.2) and outline the data corpus that has been analysed, adding some remarks on its presentation in this study (section 1.2.3).

### 1.2.1 *General remarks*

The Indo-Aryan branch of languages offers a rich historical corpus that reaches into the second millennium BC and is attested across most historical stages. However, there are several difficulties for a linguistic evaluation. These difficulties are, first, that not all periods are equally well attested. The gap that is most problematic for the purposes of the present study is the scarce attestation of the transitory period between late Middle and early New Indo-Aryan. While Middle Indo-Aryan largely continues Old Indo-Aryan on a grammatical level, retaining a certain degree of word order freedom, but with significant reductions in the morphological system, New Indo-Aryan marks a clear break from the former system, with postpositional phrases in place and an elaborate new verbal system involving several types of multi-verb constructions:

The MIA/NIA hiatus is crucial, for while there is uneven documentation for some phases of MIA also, the history of MIA is essentially one of the slow erosion of OIA forms. The NIA languages, on the other hand, emerge into the light of day with radically transformed systems.

forms, can be found in Bubeník (2006a). I will outline in detail in Chapter 5 that this view must be rejected, an insight that emerges if data from Middle Indo-Aryan are also taken into account.

Many of the details of this process, certainly one of the aspects of the history of Indo-Aryan of greatest general interest for linguistics, seem unfortunately to be inaccessible. (Masica 1991: 54–5)

Beside such gaps in the attested sources, there are large bodies of text which do not directly reflect the spoken varieties of the respective period and therefore pose difficulties. For example, the majority of the Classical Sanskrit corpus was written at a time when the spoken varieties were already Middle Indic. Furthermore, even texts closer to the respective contemporary spoken varieties often borrow from and mimic Sanskrit as the learned language. Late Middle Indo-Aryan varieties also strongly borrow from the slightly earlier Middle Indic Prakrits, some of which had also become literary languages. Even if we do work with a text that is comparatively authentic, it is often difficult exactly to locate a certain variety in relation to other varieties. The Indo-Aryan branch of languages is notoriously difficult to disentangle and sub-branches remain contested (see Masica 1991: 446ff for an overview). By consequence, if we compare data from two historical stages, we generally cannot claim that a younger variety directly descends from an older one. This becomes even more problematic when we reach New Indo-Aryan. Hindi in particular descends from a melting pot of varieties (see McGregor 2001; Shapiro 2007). Besides taking into account the naturalness of a certain text, and questions of lineage, it was my aim to select varieties in view of their relevance for a study of the rise of the postpositional phrases. I only selected such varieties as mark a significant difference from the chronologically preceding and/or following one.

Taking into account all these factors, the varieties systematically drawn upon are Vedic Sanskrit, early Middle Indo-Aryan Pali,<sup>12</sup> late Middle Indo-Aryan Apabhramsha, and early New Indo-Aryan Old Awadhi, as well as contemporary Hindi.<sup>13</sup> First of all, some of these varieties can at least very roughly be regarded as representing a historical lineage, compare Masica:

Classical Sanskrit, Pali, Sauraseni Prakrit, Literary Apabhramsha, and various forms of Western Hindi (Khari Boli, Braj Bhasha) all do represent in a general way the language of the midland at various stages—but much qualification is necessary... (Masica 1991: 54)

While my selection overlaps with Masica's 'midland' varieties, it does not do so completely, for the following reasons. First of all, I begin with Vedic Sanskrit, the oldest attested variety, which will prove crucial for understanding the roots of the modern postpositions. Having selected Vedic Sanskrit, I do not systematically take

<sup>12</sup> As is common practice, I will transliterate examples according to the IAST standard, but use westernized spellings for language names.

<sup>13</sup> Occasional references will also be made to data from Classical Sanskrit, Ashokan inscriptions (early Middle Indo-Aryan), Prakrits (middle stage of Middle Indo-Aryan), and further early New Indo-Aryan varieties.

into account Classical Sanskrit, as it is still relatively close to Vedic Sanskrit on a syntactic level and also underlies stronger normative constraints. Furthermore, we know that Vedic is the more direct ancestor with regard to certain grammatical domains of early Middle Indo-Aryan varieties. For instance, the pronominal system of Pali is rooted in Vedic forms (or closely related ones), rather than in Classical Sanskrit ones. However, Classical Sanskrit will be referred to a number of times, for instance in Chapters 4 and 5, where I discuss the development and eventual loss of local particles in Indo-Aryan, as well as the expansion of new periphrastic expressions of case. Pali, while generally retaining Old Indo-Aryan syntax, shows a notable increase in periphrastic constructions of case which take centre stage in this study. Therefore, Pali is the second historical stage that will be systematically examined. I do not take into account any Prakrit varieties from the middle period of Middle Indic in a systematic way, because they still remain relatively close to the syntactic system already represented by Pali, and because these varieties need possibly even more philological attention than others. Instead, I decided to select data as close as possible to the transitory period to New Indo-Aryan in the form of late Middle Indo-Aryan Apabhramsha. For the same reason, I included an early New Indo-Aryan variety in the corpus, the so-called Old Awadhi (see Masica 1991: 53 on its position in the lineage) as used in Tulsidas' *Rāmacaritamānasa*, one of the varieties that fed into modern Hindi, for which good sources and reference works are available (see McGregor 2001). I add insights by Strnad (2013) regarding the slightly older language of Kabir. In these varieties, the postpositional phrases are by and large established, with a certain variation remaining, for instance regarding word order. In Hindi, finally, the development has progressed almost too far as the postpositions have become clitics (see Butt and King 2004), which calls into question their status as phrasal heads. Nevertheless, the Hindi postpositions continue to show the relevant morphosyntactic features (scope over the NP string, government of the oblique, etc.). Accordingly, despite their clitic status, I will refer to the elements in question as postpositions and draw on Hindi for illustration of the completion of the rise of postpositional phrases (see section 8.3.2.2 on different terminological choices in the literature).

Before I offer a description of the data drawn upon, I want to briefly point out why this study can be carried out despite the difficulties I have mentioned. Its feasibility is assured, on the basis of the following consideration: while the changes I am going to explore are intimately tied to the histories of the individual simple postpositions of Hindi, these case studies are representative of developments in parallel varieties. The development of postpositions and postpositional phrases is a systematic one that is found in all the New Indo-Aryan languages spoken in South Asia (see the overview in Masica 1991: 230ff). Also, the equivalents in other New Indo-Aryan languages are often cognate with those in Hindi or, if not cognate, they descend from the same source types, that is from relational nouns, participles, and spatial adverbs. In many

cases, they are drawn from a circumscribed set of elements, which typically already began to be used as periphrastic expressions of local case relations from late Old Indic or early Middle Indic onwards (see, for example, the lists of elements in Fahs 1989). Therefore, it is not a hindrance that the language of the Apabhramsha texts, for instance, does not stand in a direct lineage from that of the Pali corpus, since the same structural development has occurred in all the parallel varieties. Thus, a diachronic comparison of strata in relative order, if not in direct lineage, can be drawn. Early on, Beames commented on this similarity between the New Indo-Aryan languages when it comes to the structure of nominal expressions headed by postpositions:

The structure of the modern noun is thus **strikingly homogeneous** in all the seven languages [i.e. certain New Indo-Aryan languages, UR]. To supply the distinctness of meaning which the mutilated case-endings no longer afforded, **case-affixes [i.e. postpositions], themselves for the most part shortened and simplified remnants of old adjectives, nouns, and participles, have been called in;** but though **this principle has been introduced into all the languages,** the precise words so adopted vary in each case, every language having forms of its own not used by the others. . . . Uniformity of principle is a far deeper lying bond and token of esoteric unity than mere surface similarity of individual words: the latter might have been borrowed; the former, being an inborn mental instinct, could not. (Beames 1970 [1875]: 299–300, emphasis added)

While I fully agree with Beames that the postpositional systems in New Indo-Aryan languages are structurally alike and go back to the same structural and in part to the same etymological sources, I hope that my examination of their Old and Middle Indic roots in Chapters 4 and 5 will lift the ‘esoteric’ (Beames) veil off these similarities.

To rely on structural homogeneity across the Indo-Aryan lineages, rooted in a common ancestor (and possibly reinforced through linguistic contact), has proven a fruitful approach in other studies. For instance, Hook in a paper on the Indo-Aryan compound verb assumes that ‘the direction of change is the same for all Indo-Aryan languages’ allowing for a ‘comparative study of the compound verb’ (1988: 38). Finally, in order to discern what is general and what is particular, I have striven to collect a relatively large number of examples from a variety of sources for this study.

### 1.2.2 *Texts*

Where possible, I took into account both metrical texts and prose. One might think that only prose offers good data for syntactic studies. However, prose is not available for all historical stages. Rather than excluding metrical works which would result in large gaps in the data, I try to turn the analysis of metrical texts into a virtue. A few times in this study, I exploit the assumption that metrical texts allow for a much greater variety of syntactic patterns, possibly even greater than what might actually



have been perceived as grammatical in conceptually oral discourse. Given this assumption, if a certain pattern does not even occur in large numbers of metrical texts, it may be safe to say that it was indeed not used or hardly used. In general, it is crucial to ensure that effects of genre are taken into consideration. In my presentation of the data, I point out whether an example is taken from a metrical or a prose portion in cases when this information may bear on the analysis.

My selection was also guided by the accessibility of sources, the existence of critical editions, and the availability of at least rough grammatical sketches. Whereas most Vedic Sanskrit and Pali sources are critically edited and well described—though only in part from a syntactic perspective—the comparable dearth of linguistic analyses is a challenge in working with later Middle and early New Indo-Aryan sources.

Of the Vedic Saṃhitās (the main metrical Mantra texts), the *Rigveda* (Aufrecht edition), as well as the younger White *Yajurveda* (Griffith edition) and *Atharvaveda* (Roth and Whitney edition) were included, dating from a period between the late second millennium BC and the early first millennium BC. The *Sāmaveda* was not considered as it consists almost entirely of verses taken from the *Rigveda* (cp. Mylius 2003: 35–7 for a short overview). In the following, I will refer to the portions from the Saṃhitās as Mantra Vedic (following Schäufele 1991b). Of the Brāhmaṇas, the somewhat younger prose commentaries on the Saṃhitās and Vedic ritual, I included the *Śāṃkhāyanabrāhmaṇa* (also known as *Kauṣītaki*) and *Aitareyabrāhmaṇa* belonging to the *Rigveda*, as well as the largest of the Brāhmaṇas, the *Śatapathabrāhmaṇa*, belonging to the *Yajurveda*.<sup>14</sup> Terminologically, I distinguish between Vedic Sanskrit, which includes Mantra Vedic as well as the Prose Vedic of the Brāhmaṇas, as opposed to the younger Classical Sanskrit. ‘Old Indo-Aryan’ (or the shorter ‘Old Indic’) serves as a cover term for Vedic Sanskrit as well as Classical Sanskrit varieties.

For Pali, we are in the lucky position of dealing with a large corpus which consists first and foremost of the oldest Buddhist canon, the *Tipiṭaka* (‘three baskets’, referring to its three parts) composed for the most part in the period between the First Council after the death of the Buddha in 483 BC and the Third Council under the emperor Ashoka in the third century BC (see e.g. Geiger 1916: 6). Pali is not a variety that was ever spoken in all probability, but it is based on some early Middle Indo-Aryan spoken variety (or varieties), whose exact identity is still being debated (cp. von Hinüber 2001). For this study, the texts of the middle one of the three baskets, the *Suttapiṭaka*, critically edited by the Pali Text Society, is drawn on, consisting of a large number of volumes, sufficient to fill several shelves. The parts

<sup>14</sup> There are only a couple of examples of *upari* in Brāhmaṇas other than the *Śatapathabrāhmaṇa*, which I excluded from consideration.

considered here include the *Dīghanikāya*, *Majjhimanikāya*, *Samyuttanikāya*, *Aṅguttaranikāya*, and *Khuddakanikāya*.<sup>15</sup>

For late Middle Indo-Aryan Apabhramsha—a cluster of varieties, rather than a homogeneous language (see Bubeník 1998)—I draw first of all on the *Bhavisattakaha* (tenth century?) which is the first Apabhramsha text that was made available in the West through Jacobi's edition from 1918. Further, I have included the *Paumacariu* (eighth century), the *Karakaṇḍacariu* (eleventh century), and the *Sanatkumāracaritam* (twelfth century) (see Bubeník 1998: 50–64 for an overview of extant Apabhramsha texts). Note that no prose work in Apabhramsha is known to date, apart from stray sentences. Nevertheless, because Apabhramsha (besides Avahattha) is closest to the transition to New Indo-Aryan and thus to the emergence of low-level configurationality, this stage plays an important role in this book. Also, it will become clear in the course of this study that questions of word order are only one of several issues that need be addressed in accounting for the changing syntax of nominal expressions between Vedic Sanskrit and Hindi, so that the factor of metricality does not bear so heavily after all.

For New Indo-Aryan, I selected Tulsidas' *Rāmacaritamānasa* written in a variety of Early Hindi, a text whose language is described in Saksena's (1971) grammar of Awadhi and for which a concordance exists. Data from Hindi are not evaluated quantitatively, as the rise of phrasality has been completed since the early New Indo-Aryan period. Hindi, however, serves as representative of this stage of completion, found similarly in other contemporary New Indo-Aryan languages. When discussing Hindi, I refer to citations in the literature, complemented by my own elicitations from Delhi-based speakers of Hindi (where no reference is given).

### 1.2.3 Data

As will be outlined in detail in Chapter 3, only Hindi *mē* 'in' and *par* 'on' can be traced to roots in Old Indic with certainty. The postpositions *ko* (DAT/ACC), *se* (INS/ABL), and *kā/-e/-ī* (POSS) have been linked to Old Indic forms, but there is no unanimity regarding the etymologies and they only appear in usages approximating case marking in late Middle Indo-Aryan. Ergative *ne* is a likely borrowing, appearing in the language roughly half a millennium ago. The seventh postposition *tak* 'until' will not be discussed in detail, as its history is even more uncertain (see proposals in Turner 1966 and Oberlies 1998) and it is functionally peripheral to the system of postpositions. In accordance with these circumstances regarding etymological reconstruction, I undertake a systematic corpus analysis of the attestations of precursor

<sup>15</sup> I have chosen the collection of the *Suttapiṭaka* as it figures in the Thai recension, and as made available through publications by the Pali Text Society (i.e. excluding the *Nettipakaraṇa*, the *Peṭakopadesa*, and the *Milindapañha* of the *Khuddakanikāya*).

forms of *mē* and *par* from Vedic onwards through all selected historical varieties. In addition, I take into account whatever can be known about the origin of the other postpositions. For the most part, it is only from Apabhramsha onwards that we can identify the forerunner constructions with certainty.

The corpus consists of a total of 844 attestations for *mē* and 216 attestations for *par*. In order to refer to forerunner forms of *mē* and *par* in their usages across texts and historical stages, I use small capital representations of their original Sanskrit forms for this purpose, *MADHYE* and *UPARI*. Since the attestations of *MADHYE* and *UPARI* constitute the empirical backbone of this study, I list all references of examples in the appendix.<sup>16</sup> In general, I collected the forms in their clausal contexts, and sometimes in larger passages, if the clause does not sufficiently disambiguate their usage. Of the examples attested, I only consider usages as free forms to be potential candidates for evolving into the modern postpositions, for example Skt. (Sanskrit) *samudrasya madhye* (ocean.GEN middle.LOC) ‘in the middle of the ocean’. Bound usages of the type *samudramadhye* (ocean\_middle.LOC) ‘in the ocean-middle’, where the potential gram forms the second member in a compound—frequent especially in Pali—are excluded from the analysis. This is due to the insight from grammaticalization research that morphologically unbound elements are the ones that may grammaticalize within certain constructions, and through frequency effects erode and possibly cliticize to their lexical host. Morphological fusion may be a late stage of the development, but does not present the starting point.

In my analysis, I have aimed to take into account the skewing of the data through the many repetitions of identical sentences or even entire paragraphs. Rigvedic verses, especially, are frequently quoted in other Vedic or younger texts. Also in the Pali canon repetitions, sometimes of entire paragraphs, are widespread. In order to deal with this problem, I only counted a clause once per published text. This is of course an artifice of editing practice; however, I believe that it reduces the many repetitions in an effective way, while still weakly representing highly frequent repetitions, if they appear spread over several edited texts. Citations from the *Rigveda* in other texts are not counted on a general basis due to their frequency and decidedly quotative character. In the appendix, where all attestations are listed, examples are annotated as to their usage in compounds or in repeated clauses (see the abbreviations given at the beginning of the appendix). Only references in bold print are examples that entered the analysis in this study.

Apart from data involving the former stages of the Hindi simple postpositions, various other types of data were collected. Attestations of the last remaining local particles in Pali were gathered for Chapter 4, as well as data on various relational

<sup>16</sup> For Old Awadhi, I only include instances of *MADHYE*, which were already studied in Reinöhl (2009).

nouns and adverbs besides MADHYE and UPARI in Vedic Sanskrit and Pali for Chapter 5. For Chapter 6, I collected data on the syntax of pronominal possessors across several historical stages in order to undertake a detailed analysis of the origins of the postpositional syntagmatics. The details of these additional collections of examples will be specified in the respective chapter.

The historical texts are available in different digital and non-digital formats. For Vedic Sanskrit, some of the texts are accessible in searchable digital form and the voluminous concordance of Viśva Bandhu (1935–1965) allows one to locate the relevant passages in remaining texts. The Pali *Tipiṭaka* is digitally available on several websites, with a very useful search engine implemented on the website of the Vipassana Research Institute.<sup>17</sup> However, the digital versions are hand-typed and in part rely on versions of the texts other than those chosen by the Pali Text Society for their critical editions. In all cases, it was necessary to go back to the critical editions to locate the examples, which often posed a challenge due to differences in textual divisions. Here, attestations are always cited as in the editions by the Pali Text Society. Rare exceptions to this practice, where an edition by the Pali Text Society has not yet been published, are specified. For Apabhramsha, the texts possess indices, which typically only give one or a few references for illustration. Further forms had to be searched for manually, which I did systematically for the *Bhavisattakaha*, *Karakaṇḍacariu*, and the *Sanatkumāracaritam*, as the comparatively old text editions do not yield good results of optical character recognition. For the Old Awadhi piece *Rāmacaritamānasa*, a word index (Kānta 1937) is available which—if not always corresponding to the numberings in the matching edition of the text—is very helpful and a main reason for selecting this text, as it allowed me to collect a large number of examples. In addition, the digital version on GRETIL<sup>18</sup> was sometimes drawn upon—even though it is based on a slightly different version of the text.

Apart from a few instances where I quote from secondary literature, I have adapted all examples, whether in Devanagari or Roman script, to the standard of the International Alphabet of Sanskrit Transliteration. Words written without a break in the original spelling will be separated, when this can be done neatly between two letters. However, when a sandhi needs resolution in order to separate words, I add a second line in between the example and the glossing, where I give the citation forms of the words separated. Only the first line will be written in italics. In some cases of examples cited from secondary literature, sandhis are already resolved. Readers unfamiliar with Indo-Aryan languages should expect a large number of sandhis, noted in the spelling. Accordingly, forms glossed alike may not in all cases be written alike. For instance, MADHYE in Sanskrit will at times be written as *madhya*, and UPARI as *upary*. Annotations of my own corpus data are my own, no annotated material

<sup>17</sup> The website is [www.palitext.org](http://www.palitext.org) (last accessed 5 October 2013).

<sup>18</sup> <http://gretil.sub.uni-goettingen.de/>.

being available, which also applies to many examples cited from secondary literature. In most cases, I refrain from a morpheme-by-morpheme annotation, avoiding the pitfalls posed by the highly inflectional character especially of the Old Indic varieties. Furthermore, I often do not specify all morphological categories in order to avoid overly complex annotations, but only mark categories crucial for the analysis besides additional ones which facilitate understanding. Most but not all of the texts have been translated into English, French, German, or Hindi, and Dutch in one case, languages that I could work with, besides other translations. In those cases where good translations are available, it was always exceedingly helpful to take them into account. The translations that I consulted are listed in the bibliography of primary sources directly under the matching primary source. Still, in the majority of cases, I offer translations that differ from existing ones in that mine are more literal, aiming at reflecting the original linguistic structure within the boundaries of English. Some of the nineteenth-century translations of Pali texts especially, while aesthetically valuable, are not very helpful from a strictly linguistic point of view, as they strive to offer a literary or even poetic rendering. In some cases, however, I do rely on translations from the literature, which will be referenced accordingly. This is the case for passages in Apabhramsha in particular, the variety or rather bundle of dialects least described, which employs vocabulary and morphological forms from various other varieties, and in part includes ‘deśī’ words (literally ‘country words’, i.e. words unattested in Sanskrit or not otherwise derivable from an Indo-Aryan source). As a result of this, not even specialists like Jacobi are able to discern the exact meaning at times. Unfortunately, particularly in the case of Apabhramsha, entire texts remain untranslated, such as the *Bhavisattakaha*. However, difficulties mostly pertain to the exact meanings of a word and sometimes to the exact identity of a morphological form—matters of syntax that are relevant for the research questions pursued here do not pose comparable difficulties.

### 1.3 Outline of the study

In Chapter 2, I introduce the two major domains that are brought together in this study: grammaticalization and configurationality. In order to lay the groundwork for the subsequent chapters, I put a special focus on configurationality as pertaining to the organization of nominal expressions. Introducing the empirical starting-point of this study, I contrast Vedic Sanskrit and Hindi from the perspective of configurationality, and I round off the chapter by outlining ways in which configurationality has been treated as a historical phenomenon, as well as ways in which grammaticalization may impact on configurationality. Having laid these foundations, I outline and evaluate what is known about the simple Hindi postpositions in Chapter 3, the empirical precondition for tracing the origins of phrase structure in this branch of languages.

In Chapter 4, before I can actually turn to exploring the rise of the postpositional phrases, I will discuss claims in the literature that link the New Indo-Aryan postpositions to the Proto-Indo-European local particles, in parallel with other branches of Indo-European, and despite a lack of etymological links. I will show that this scenario must be rejected on several grounds, which takes me to my own account of the origins of the postpositions and postpositional phrases.

I turn to my account of these origins in Chapters 5 to 8, the core of this study. The structure of these four chapters is thematically motivated, with a chronological correlate regarding the historical stages that will be the respective focus. In Chapter 5, I begin by addressing the question of the input components that fed into the grammaticalizing constructions, which is mostly based on Old Indo-Aryan and early Middle Indo-Aryan evidence. Here, it becomes clear that it is essential to include in the analysis data from Middle Indic and not only from Vedic Sanskrit, as is often done in similar studies. Only then does it become apparent that the origins of the modern postpositional phrases lie in asymmetrical groups, and not in symmetrical ones. In Chapter 6, I focus on questions of syntagmatic arrangement, taking into account data from all historical stages up to Apabhramsha. I argue that pronominal possessors play an important part in the development, foreshadowing this pattern from early Vedic times in a large number of cases. Here, I also remark on the alleged oddity that Indo-Aryan has postpositions rather than prepositions, differing from other branches of Indo-European. The obligatorification of overt genitive possessors is the focus of Chapter 7, where changes from Apabhramsha to Hindi are considered. I discuss the core shift from lexical to function words, and the concomitant emergence of the earliest of the modern phrases. In Chapter 8, finally, I turn to the way in which the various grammaticalizing elements have merged into a new functional category of Hindi. I explore interactions between the different elements in recent centuries and the way in which they have developed into a paradigm, also accommodating new members.

## Grammaticalization and configurationality

In this chapter, I introduce the two research domains that frame this study, grammaticalization and configurationality, and outline how the former bears on the latter in the sense that it can entail a change from non-configurational to configurational. I begin with a general introduction to grammaticalization (section 2.1), and then turn to configurationality or—more precisely—the lack thereof, non-configurationality (section 2.2). In accordance with the focus of the present study, I add a section specifically devoted to the role of nominal expressions in non-configurational languages, where I distinguish ‘true discontinuity’ from other phenomena (section 2.3). Having laid the theoretical foundations, I turn to Vedic Sanskrit and Hindi and outline their grammatical system from the perspective of configurationality and focusing on nominal expressions (section 2.4). This will give us the starting point as well as the endpoint of the changes that will be examined over the next chapters. Returning to the overarching research question, I outline prior approaches to configurationality as a historical phenomenon (section 2.5). I end with an exploration of ways in which grammaticalization processes may impact on configurationality, presenting three claims that will be substantiated in the course of this study (section 2.6), before I conclude in section 2.7.

### 2.1 Grammaticalization

Following nineteenth-century agglutination theory, Meillet coined the term ‘grammaticalization’ early in the twentieth century, which in his words designates ‘the attribution of a grammatical character to a word that used to be autonomous’ (Meillet 1975 [1912]: 131, transl. UR).<sup>1</sup> Meillet introduces this type of historical change with

<sup>1</sup> Von der Gabelentz (1901) also addressed the phenomenon at an early point in time. For the pre-history and early history of grammaticalization research see Lehmann (2002 [1982]). In Lehmann’s study and in Hopper and Traugott (2003), a general introduction and overview of topics can be found. Important observations and distinctions can be found in the first chapter of Bybee et al. (1994) and in the second chapter of Himmelmann (1997).

the example of French *suis* *be.1SG* 'am', which has developed from a locative-existential verb as in *je suis chez moi* 'I am at home' into a copula as in *je suis malade*, as well as into an auxiliary as in, for example, *je suis allé* 'I went' (Meillet 1975 [1912]: 131). Since Meillet, grammaticalization is typically defined in this way, that is, as the process by which a certain element becomes grammatical. This definition was expanded by Kuryłowicz (1964: 52) to also cover changes of an already grammatical word that progresses even further towards the grammatical pole, a variant of which can be found in many publications today (e.g. Lehmann 2002 [1982]: 8; Bybee et al. 1994: 4–5).

In this study, grammaticalization is conceived of as a cover term for a broad range of phenomena connected by the following characteristics. An element with an independent denotation (an *autosemanticon* in traditional terminology) develops into an element that expresses meaning in relation to other elements, that is it becomes a *synsemanticon* (what is now sometimes called 'primary grammaticalization'). Or, an element which is already synsemantic becomes even more synsemantic, that is it loses further autosemantic meaning components ('secondary grammaticalization').<sup>2</sup> In this study, I focus on the early stages in the development—primary grammaticalization—which will be shown to trigger the emergence of phrasality. An element undergoing grammaticalization will be referred to as a grammaticalizing element or *gram* for short, defined as in Bybee et al. (1994: 2), which can be a word, a clitic, a complex structure such as *be going to*, or even a stem change or reduplication. Note the important point that an element always grammaticalizes in a specific construction, about which more will be said below.

What unites changes in early stages and in late stages of grammaticalization alike, when considered from a distributional perspective, is context expansion (Himmelmann 2005). This refers to a widening of the number of lexical hosts a gram can combine with inside its constructional context, as well as to a widening of the syntactic and semantic–pragmatic contexts the entire construction can occur in. In studying a case of primary grammaticalization, to focus on the shift from autosemantic to synsemantic on the one hand, or on context expansion on the other hand, correlates with conceptualizations of grammaticalization as involving either discrete or continuous change. I believe that both perspectives can be productively applied to the phenomenon and add to our understanding.

On the one hand, for instance, the shift from a relational noun to an adposition, which will be examined here, involves the discrete reconceptualization of the element

<sup>2</sup> For example, a local adposition, grammatical as a synsemantic element, may lose meaning components of a concrete spatio-temporal reference. For instance, Latin *ad* 'at, to' developed into a marker of recipients and benefactives.



in question from having an independent denotation (e.g. denoting a body part or a certain spatial area) to expressing a meaning that locates another entity in a concrete or abstract space. Further, after a stage where both the lexical and the grammatical reading may be possible, the older lexical meaning will be blocked in certain contexts, also a discrete step towards grammatical meaning (cp. Diewald's shift from 'critical' to 'isolating' contexts, e.g. Diewald and Smirnova 2012).

On the other hand, evidence for the gradual nature of grammaticalization comes from the fact that even highly grammaticalized markers often retain a flavour of their lexical origin. English *will*, in general a future marker, in certain usages gives away its lexical origin, for instance when expressing volition as in the tag question *if you will* (cp. Bybee et al. 1994: 208). Non-discreteness also shows when we zoom in upon the threshold between autosemantic and synsemantic readings. For example, Eckardt (2006: 99ff) discusses examples from around 1550–1650 where either reading of *be going to* VP, literal or grammaticalized, is possible. In such cases, quantitative evaluations will give the impression of a continuum because the element in question will not shift at once to a new meaning, but will gradually show the new grammaticalized meaning more and more often. In general, the direct correlate of becoming (more) synsemantic is the increase in potential contexts of use and thus frequency, again a gradual change (e.g. Bybee 2003; Himmelmann 2005).

Thus, similar to the pair of sound laws and lexical diffusion, there is a conceptual level on which the threshold to synsemantic is discrete, while the empirical foundation suggests gradual change. In this study and in particular in Chapter 7, the emergence of postpositional phrases will be examined from both perspectives, in accordance with the aspect focused on. For example, I will strive to identify the precise contexts of use in which a metonymic, discrete shift from denoting a complex to a simple spatial relation could take place. At the same time, I will show how metaphorical shifts lead to an expansion of host class which spans the entirety of the attested history of Indo-Aryan, even starting before the elements in question approach a synsemantic status, and continuing through the ages until Hindi (cp. Bybee et al. 1994: 5).

A characteristic of grammaticalization, which plays an important role in this study, is that it is not elements in isolation, but elements embedded *within specific constructions* undergoing the change, that is, in combination with certain other autosemantic and/or synsemantic elements, which are arranged syntagmatically in a specific way (e.g. Compes et al. 1993; Bybee et al. 1994; Eckardt 2006): '[I]t may be more accurate to say that a construction with particular lexical items in it becomes grammaticized, instead of saying that a lexical item becomes grammaticized' (Bybee 2003: 602). For instance, English *go* did not as such develop from an autosemantic element denoting the event of moving by walking into a marker of prediction. Instead, while *go* could and still can be used in its 'lexical' or autosemantic sense, it developed into a marker of futurity within the construction *be \_ing to* VP. Note that

I speak in this study both of the grammaticalization of a *gram* as well as of the grammaticalization of a *construction*, depending on what I wish to focus on.<sup>3</sup>

Apart from the unifying characteristics that cases of grammaticalization involve changes to (more) synsemantic and pertain to constructions, the diversity of grammaticalization phenomena is wide. In particular, cases of grammaticalization differ (i) with regard to the particular types of semantic change involved in the process; (ii) with regard to the semantic, morphological, and syntactic relations between the grammaticalizing element and other elements inside and outside its constructional context; and (iii) with regard to different epiphenomenal effects of earlier and later stages of grammaticalization processes. I will briefly outline the variation found and specify the respective sub-types relevant for this study.

First of all, different types of semantic change can be observed in grammaticalization phenomena, primarily metaphor, metonymy, and pragmatic inference. In this study, the first two play a central role. Metaphorical changes are typical for early stages of grammaticalization where the lexical meaning of a string can be consciously manipulated by the speaker in such a way that a construction is used to stand in for another more general meaning. An example is body part expressions used metaphorically for spatial relations such as English *at the back of (the house)*, *at the foot of (the mountain)*, etc. After such an initial metaphorical transfer, a metonymic shift may follow from profiling the space denoted (e.g. the area behind something in *at the back of*) to profiling the spatial relation between the two objects connected (cp. Rubba 1994). An example that has run through all these stages is English *beside* which evolved from Old English *be sīdan* 'by the side'. *Sīde* at first denoted the flank of a human body or animal before it evolved into denoting the side of any object. After this metaphorical transfer, it was metonymically extended to profiling adjacency to the respective object. Similar shifts will be outlined in detail in Chapter 7. I will argue that they are the triggering factors turning the constructions in question into phrasal units.<sup>4</sup>

Second, the relation of the gram to other elements within the construction as well as outside can be of different kinds. The gram may constitute the head of the construction in one or several respects, syntactic, morphological, and/or semantic. For instance, in the case of the prepositional construction *in the house*, *in* determines the syntactic function of the construction, is subcategorized for a nominal expression,

<sup>3</sup> Constructions undergoing grammaticalization can be of very different complexity, consisting of one or several other autosemantic or synsemantic elements in addition to the gram.

<sup>4</sup> An example for a shift by pragmatic inference comes from *be going to* VP. 'Going' as such does not lend itself to a metaphorical or metonymic transfer to denoting intention or futurity. For example, German *gehen* 'to go' has no such implication in any usage, as pointed out by Eckardt (2006: 95). It is the specific construction *be going to* VP that is shown by Eckardt (2006) to have occurred in such contexts where a speaker is conveying that he is going somewhere *in order to* do something. It is the progressive form which conveys a sense of ongoing preparation for some future action. With time, the concrete meaning of *going* vanishes and only the meaning of intention remains.

and marks the construction as one expressing a spatial relation. In other cases, grammaticalizing elements do not accumulate properties of semantic and formal headedness in this way. Consider the development of French *pas* 'step' into a marker of negation. *Pas* does not determine the syntactic function of (written and formerly also spoken French) *ne V pas* and does not govern *ne*, but it is only *pas* that survives in contemporary spoken French as a marker of negation. Externally too, *pas* does not connect the subject with the verb complex. In this study, too, we will be dealing with different types of relation between the gram and other elements within and outside of the construction. Specifically, it will become clear that Indo-Aryan attests to constructions that have completely different internal semantic and morphosyntactic structures in comparison with those constructions that gave rise to adpositional phrases in other branches of the language family.

Third, grammaticalization processes entail somewhat different epiphenomenal effects in earlier stages in contrast with more advanced stages. Here, the transition from unbound to bound morphology is an important cut-off point (which may or may not coincide with 'primary' vs. 'secondary' grammaticalization). In particular, it appears that several 'processes' that have been regarded as necessary or optional components of grammaticalization phenomena in general only pertain either to the stage before cliticization, or to the stage after cliticization. For example, Hopper's (1991) 'principle' of decategorialization seems to be relevant mainly for the period until the gram reaches cliticity. On the other hand, phonetic alterations at the boundary of host and gram may become relevant in later stages of grammaticalization, in the course of or after cliticization. In this study, I focus on grammaticalization phenomena that lead to the emergence of grams which organize their constructional context in a phrasal way; that is, which pose constraints of selection and syntagmatic arrangement on other elements of the construction. I do not consider advanced stages of grammaticalization where the gram has become a clitic or even an affix, as I only focus on the rise, not the break-down of phrasality.

## 2.2 Non-configurationality: A brief history

In this study,<sup>5</sup> non-configurationality is conceived of as describing a language that shows the characteristics identified by Ken Hale (1983) for Warlpiri: (I) free constituent order, (II) discontinuous nominal expressions, and (III) null anaphora.<sup>6</sup>

<sup>5</sup> For another recent overview of non-configurationality see Murasugi (2014). Nordlinger (2014) surveys the phenomenon in Australian languages. A recent implementation in a generative theory (i.e. in the Minimalist framework) is É. Kiss (2008) on Hungarian.

<sup>6</sup> Hale speaks of 'free word order' for criterion I and others of 'discontinuous NPs' for criterion II. I use 'free constituent order' for criterion I because what Hale means by free word order is the placement of S and O arguments in relation to the verb. Free placement down to the level of the components of nominal expressions, i.e. 'free word order' in a literal sense, only comes into play with criterion II, discontinuous

Consider Hale's oft-quoted examples below, which illustrate a largely free permutation of arguments without a change in logical content (exx. (6)–(8)), nominal elements forming a semantic unit being placed apart (ex. (9)), and the dropping of arguments (ex. (10)) (Hale 1983: 6–7):

- (6) *Ngarrka-ngku ka wawirri panti-rni.*  
 man-ERG AUX kangaroo spear-NONPAST  
 'The man is spearing the kangaroo.'
- (7) *Wawirri ka panti-rni ngarrka-ngku.*  
 'The man is spearing the kangaroo.'
- (8) *Panti-rni ka ngarrka-ngku wawirri.*  
 'The man is spearing the kangaroo.'
- (9) *Wawirri kapi-rna panti-rni yalumpu.*  
 kangaroo AUX spear-NONPAST that  
 '(I) will spear that kangaroo.'
- (10) *Panti-rni ka.*  
 spear-NONPAST AUX  
 '(He/she) is spearing (him/her/it).'

Since S and O arguments can be placed in any order relative to the verb, complex nominal expressions do not have to form a syntagmatic unit, and arguments can be dropped at will, Hale concluded that there is no evidence for configurationality in Warlpiri, that is for the hierarchical organization of expressions into phrases determined by the principles of dominance and precedence. The claimed lack of a VP in particular sparked a controversial debate in generative circles in the early 1980s when Hale's work began to be discussed widely.<sup>7</sup> Since the asymmetry between subjects and objects had been defined by means of the VP, as the distinction between its external and internal argument respectively, one of the basic building blocks of the then current Principles and Parameters framework was fundamentally called into question. The problem was briefly addressed by Chomsky (1981: 127–35), which

nominal expressions; with regard to criterion II, I prefer to speak of nominal 'expressions' rather than 'phrases,' because discontinuity violates one of the defining characteristics of phrases in my understanding, that of forming an integral syntagm.

<sup>7</sup> The evidence underlying the claimed lack of subject–object asymmetries has been known for a longer time. Already before Hale, who is generally cited as having identified the problem, other authors had diagnosed a lack of subject–object asymmetries in some languages (e.g. Hinds 1974 on Japanese and Staal 1967 on Vedic Sanskrit, see section 2.4.1). A very early and very insightful exploration of free order, one domain in which the existence of the VP may be called into question, is Weil (1844), and in more recent times e.g. Blake (1983) and Mithun (1987). I will not review these contributions in detail here because 'free word order' is often used to refer to the position of constituents (i.e. various orders of S, O, and V), but not so much to discontinuous nominal expressions, which are the focus of this study.

resulted in a considerable amount of attention being drawn to the phenomenon. Searching for a theoretical solution, Hale (1983) proposed a Configurationality Parameter that divides languages into two types, configurational and non-configurational. According to Hale, Warlpiri showed asymmetries between subject and object only on the level of lexical structure (LS), but not on the level of phrase structure (PS). This meant abolishing the so-called Projection Principle, that is, the idea that asymmetries at LS are mirrored on PS. However, to preserve asymmetries in ‘non-configurational’ Warlpiri on levels other than phrase structure was an important device for Hale (1983) to explain subject–object asymmetries in domains such as person marking, binding, and control relations. For example, in reflexive-reciprocal constructions, a marker on the object specifies that it is controlled by the subject with an interpretation of reflexive or alternatively reciprocal if plural. The selection of the controller is based on the subject relation, that is ergative if present, otherwise absolutive, as illustrated by the following examples with the argument frames *erg-abs* and *abs-dat* (Hale 1983: 21):

- (11) *Kurdu-jarra-rlu ka-pala-nyanu paka-rni.*  
 child-DU-ERG PRS-33SUBJ-REFL<sup>8</sup> strike-NONPAST  
 ‘The two children are striking themselves/each other.’
- (12) *Karnta ka-nyanu yarnka-mi (jurru-ku).*  
 woman PRS-REFL grab-NONPAST (head-DAT)  
 ‘The woman is grabbing herself (by the head).’

Several strands of research take off from this early work by Hale. First, his study was followed by a wave of publications examining diverse languages in the light of non-configurationality, Australian languages, but also certain languages from North America, South Asia, and others (see e.g. the articles in Marácz and Muysken 1989). The view in several of these publications was that, while there may be cases like Warlpiri with a far-reaching ‘flat’ syntax with regard to phrase structure, all languages show asymmetries in at least some domains. As an important counter-proposal, Heath (1986) argued in a paper on Nunggubuyu, that here, indeed, we do encounter a language that is not only non-configurational on a phrase-structural level with a general absence of function words that might constrain expressions into phrasal units, but also lacks asymmetries in other domains such as binding and control.

The discussion on how to reconcile languages like English with languages like Warlpiri on a theoretical level also continued. An alternative proposal to Hale (1983), which retains configurationality on the level of phrase structure, is associated with Jelinek (1984) (but see Austin 2001: 307 for references to earlier publications). Jelinek

<sup>8</sup> ‘33’ = 3rd person dual.

proposes the Pronominal Argument Hypothesis (PAH), which crucially relies on an auxiliary complex in 2nd position in Warlpiri, the only position that is fixed, as can be gleaned from examples (6)–(10). This complex consists in an auxiliary stem which combines with affixes that cross-reference subject and object arguments (null morphemes cross-reference 3rd person singular arguments). Jelinek argues that this auxiliary complex in 2nd position (P<sub>2</sub>) contains the entire core syntactic structure of the clause, that is, the grammatical features of the verbal component and of its arguments. The bound pronominals are argued to be full-fledged arguments, while lexical NPs, if present, function as adjuncts adjoined to them. This account offers a proposal whereby Hale's three non-configurationality criteria are licensed while the projection principle is retained: since the arguments are already expressed in the auxiliary complex, free nominals may be placed in any position irrespective of their syntactic function (criterion I); several nominals relating to one grammatical relation do not underlie a constraint to appear in juxtaposition, being separate adjuncts (criterion II); and by virtue of being adjuncts, these may freely be dropped (criterion III).

Jelinek's PAH proposal has had a strong impact on the literature (e.g. Speas 1990; Baker 1996) and Hale himself subscribed to it at one point (see Hale 1989). However, the growing amount of research into other Australian languages and languages of other continents soon brought to the fore languages that showed all the defining characteristics, but were not or only partially head-marking, that is, lacked an auxiliary complex that could be argued to build up a configurational clause structure.<sup>9</sup> For example, Austin (2001) describes how Jiwari, another Australian language, qualifies for non-configurationality in every respect, but is thoroughly dependent-marking. Even quite apart from a language like Jiwari, Austin and Bresnan (1996) show, based on Simpson (1991), that the PAH fails to capture a number of phenomena in Warlpiri itself.<sup>10</sup> Austin and Bresnan (1996: 259ff) and others (e.g. Nordlinger 1998) have argued that pronominal morphology is an areal feature of Australian languages that cross-cuts languages with different constellations of non-configurational and configurational characteristics.<sup>11</sup>

Based on these empirical insights, another theoretical path was taken by authors writing in the framework of Lexical Functional Grammar (LFG), developed, among other things, in order to accommodate the facts of non-configurational languages (see Bresnan 2001: 5ff). In a spirit reminiscent of Hale's (1983) differentiation

<sup>9</sup> Golumbia (2004) offers a critique of attempts at salvaging configurationality as universal, arguing that this reveals a colonialist perspective on 'non-modern' languages.

<sup>10</sup> The PAH does not account for the semantic interpretation of nominal expressions as definite or indefinite, and does not explain lexically governed case forms, the occurrence of null pronominals in clauses that lack auxiliaries, or the occurrence of arguments unregistered in the auxiliary complex (for details see Austin and Bresnan 1996: 234–43).

<sup>11</sup> Applications of the PAH to other languages, where the hypothesis cannot account for a number of phenomena either, are e.g. Bowern (2005) on Bardi, LeSourd (2006) on Maliseet-Passamaquoddy, or Davis and Matthewson (2009) on Salish languages (see Murasugi 2014).

between PS and LS, LFG teases apart different types of domain, for example constituent structure (CS, corresponding to what others call phrase structure), functional structure (FS) specifying grammatical relations (independently of phrasal organization), and argument structure (AS). Crucially, there is no such thing as a projection principle, mismatches between the layers being allowed. An application of LFG to head-marking Warlpiri can be found in Austin and Bresnan (1996), and to dependent-marking Wambaya in Nordlinger (1998), who proposes a ‘constructive case’ analysis, based on the idea that in some languages, case morphemes rather than hierarchical positions are involved in the construal of grammatical relations.

Besides a disentanglement of syntactic domains, for example, in LFG, several authors have striven to distinguish more carefully among Hale’s criteria as well as other criteria that have been proposed as being linked to non-configurationality.<sup>12</sup> In particular, several authors have noted that many languages show free constituent order and null anaphora—as such already calling into question the existence of a VP—while only some languages also show discontinuous nominal expressions. Today, a distinction is made between languages that are non-configurational in Hale’s (1983) sense, that is including discontinuity of nominal expressions, and languages which ‘only’ show free constituent order and null anaphora.<sup>13</sup> I will refer to the former as ‘non-configurational’ and to the latter as ‘low-level configurational’.<sup>14</sup> With regard to the relation between the three criteria, it appears that discontinuity of nominal expressions implies the other two criteria, while the reverse is not the case (cp. remarks in Baker 2001 on possible interrelations).

## 2.3 Non-configurationality and nominal expressions

If a language appears to be non-configurational, a number of issues remain to be clarified. In particular, it is important to examine whether discontinuity of nominal expressions actually exists as a phenomenon of its own kind, or whether it is instead (i) an epiphenomenon of syntactic, prosodic, or other types of constraint that have nothing to do with the placement of nominal expressions per se, or (ii) only constitutes

<sup>12</sup> In contrast with his earlier work, Hale (1989) proposes separating the question of word order from the question of non-configurationality, conceived of as pertaining to subject–object asymmetries. With reference mainly to Navajo, Hale argues that it is null anaphora on the one hand, and certain binding asymmetries on the other, which are central for ‘the relation between PS and LS’, i.e. the projection principle, which he deems crucial for the issue of non-configurationality.

<sup>13</sup> In generative theories, the latter have played the more central role due to the importance of subject–object asymmetries.

<sup>14</sup> Alternative terms that can be found are ‘strong’ vs. ‘weak’ non-configurationality (Webelhuth 1984–85: 206), ‘free word order’ vs. ‘free phrase order’ languages (Schäufele 1991a), or non-configurational (narrowly speaking) vs. ‘scrambling’ (Ross 1967; see also Stucky 1981). The third type, strong configurationality, describes languages such as English, which show a certain amount of evidence of a phrasal organization even up to the VP.

'apparent discontinuity', that is, where the elements placed apart do not in fact share a syntactic role. Since I claim that we start off in Indo-Aryan with potentially discontinuous nominal expressions which only gradually develop into phrasal units, and since this is the criterion, of the three given by Hale, which sets non-configurational languages apart from low-level configurational ones, this point requires clarification.

First, I outline Wackernagel (P2) effects, which have been claimed to trigger discontinuity as independent syntactic or prosodic (depending on the nature of the respective P2) mechanisms (section 2.3.1). Second, I discuss types of apparent discontinuity (section 2.3.2). Having established that 'true discontinuity' exists, I turn to phenomena that have been assumed to correlate with it, namely (i) a lack of obligatory function words in the nominal domain, and (ii) a weak or absent noun–adjective distinction (section 2.3.3). These correlates are of particular importance for the present study, as I claim that the development of function words in the nominal domain entails the rise of phrasal organization, which in turn results in a sharpening of the noun–adjective distinction.

### 2.3.1 *The Wackernagel position*

Several of the languages that have been considered non-configurational have a Wackernagel position (Wackernagel 1892), somewhat infamously Warlpiri.<sup>15</sup> Depending on language and clause type, a special slot in 2nd position may be optionally or obligatorily filled by different kinds of clitic (sometimes also accented elements as in Vedic) (see Halpern and Zwicky 1996; Anderson 1993). The Wackernagel position has played an important role in discussions of non-configurationality, and it has been used to argue for the existence of phrasality in different ways.<sup>16</sup>

Evidence for phrasality is found in cases where P2 delimits the initial position to a single word, but exceptionally allows for complex constituents, as in Warlpiri (in other cases, the constraint may be defined in terms of morphemes, or normally allow

<sup>15</sup> Nevis (1988: 114–15, fn. 4) gives a long list of languages that are non-configurational or low-level configurational languages and at the same time 'exhibit at least one Wackernagel-type clitic'.

<sup>16</sup> Here, I only discuss ways in which the Wackernagel position may be involved in actual cases of discontinuity (or continuity) of nominal expressions. Note, however, that the theoretical conceptualization of elements in P2 may bear on the question of discontinuity in a general way. Discontinuity presupposes that a nominal expression is conceived of as a syntactic unit on some level of analysis. In some frameworks, there is no space for this, however. For example, in Jelinek's account (see also Austin and Bresnan 1996: 230), the affixes on the verb constitute the arguments, with the lexical nominal expressions being adjuncts to them. There is no level on which the bound pronominals and the disparate adjuncts are merged into a unit—hence, there cannot be any discontinuity either.

Above and beyond the categorization of pronominals in P2, the mere fact that there is a fixed syntactic position has been taken as evidence by some authors that the language in question is not actually syntactically 'flat', at least not with respect to word order. For Warlpiri, Austin and Bresnan (1996: 225) therefore posit a part-hierarchical phrase structure tree with hierarchical nodes for the first and second positions at the left edge of the clause and a flat tree after the Wackernagel position. Since I consider phrasality as central, however, the mere fixation of a P2 does not as such bear on the question of configurationality.



for entire constituents; see Halpern and Zwicky 1996 and Simpson and Mushin 2008 on the clause-initial position in a number of Australian languages). Compare the following example, which suggests a weak degree of phrasality of nominal expressions:

- (13) *Wawirri yalumpu kapi-rna panti-rni.*  
 kangaroo that AUX spear-NONPAST  
 ‘(I) will spear that kangaroo.’ (Hale 1983: 6)

Otherwise, however, the P2 has been argued to trigger discontinuity in a variety of ways, which I will refer to as ‘P2 effects’. These come in two different forms. First, pronominal elements co-referential with other nominal elements may be placed inside P2. In such a case, discontinuity between the pronominal in P2 and other co-referential nominal (in a broad syntactic sense, i.e. any element that may form part of a nominal expression) elements in the clause arises automatically.<sup>17</sup> Second, P2 may involve a constraint of always allowing only a single word into initial position. A word placed in initial position preceding P2, when part of a complex nominal expression, is automatically placed at a distance. Both these effects will be shown to exist in Vedic. Importantly, if all potential cases of discontinuity in Vedic were to involve one or both of these two types of P2 effect, this could be taken as an argument against Vedic being non-configurational with discontinuity only being epiphenomenal. I argue below that Vedic does show cases of discontinuity unconnected to P2 effects.

### 2.3.2 *True and apparent discontinuity*

When dealing with a language that appears to be non-configurational, it is important to distinguish between so-called ‘apparent’ and ‘true’ discontinuity (see Rijkhoff 2002: 255–9).<sup>18</sup> Apparent discontinuity is found in cases where co-referential elements in a clause do not share a syntactic role,<sup>19</sup> and therefore should not be considered a single nominal expression. This is a phenomenon by no means restricted to non-configurational languages. The special challenge in non-configurational languages is that general word order freedom and the oft-found weak noun–adjective distinction can lead to ambiguity. Hale (1983) addressed this difficulty, discussing ‘merged’ and ‘unmerged’ interpretations of nominal expressions in Warlpiri. Since Hale, however, comparatively little has been published on this topic (but see a few

<sup>17</sup> This is also the case if a co-referential nominal element directly precedes or follows a pronoun in P2, as the elements are placed in adjacency for independent syntactic or prosodic reasons, but not because they are semantically connected (see section 6.1.2 on such cases in my corpus).

<sup>18</sup> Apparent discontinuity may co-occur with a P2 effect. For expository purposes, I discuss these cases separately, but see fn. 31 below on Vedic.

<sup>19</sup> In the absence of hierarchical structure, sharing a syntactic role can be conceived of in terms of dependency relations.

references in Schultze-Berndt and Simard 2012: 1026). I focus here on Schultze-Berndt and Simard (2012) on Jaminjung, based in part on an earlier study by McGregor (1997) on Gooniyandi.

For Jaminjung, Schultze-Berndt and Simard (2012) assume two types of nominal expression, continuous and discontinuous, consisting of co-referential elements that share a syntactic role and appear within a single prosodic phrase. Whenever co-referential nominal elements are spread over several prosodic phrases, it appears that the elements in question do not share a syntactic role. Thus, prosodic boundaries appear to delimit domains for syntactic constituents. Examples where nominal elements are spread over several prosodic phrases and do not share a syntactic role—cases of apparent discontinuity—include secondary predications, afterthoughts, and part-whole constructions.<sup>20</sup> Consider the following examples illustrating an apparently discontinuous pattern involving an afterthought (ex. (14)) in contrast to an example of true discontinuity (ex. (15)). The former shows nominal elements spread over prosodic phrases, while the latter shows a discontinuous nominal expression within a single prosodic phrase, the boundaries being marked by back-slashes. As argued by Schultze-Berndt and Simard, the entire expression in (15) is focal, but *mulanggirr* in addition bears a contrastive focus accent (indicated by ^), *wirib* being already activated in discourse.<sup>21</sup>

- (14) *Buyud=biyang jabul=ni burr-angu=rrgu=rndi \*  
 sand=SEQ shovel=ERG/INS 3PL>3SG-get/handle.PST=1SG.OBL=EVID  
*buj-mawu buyud \*  
 bush-DWELLER sand  
 ‘They got sand for me with a shovel, the bush kind of sand.’  
 (Schultze-Berndt and Simard 2012: 1027)
- (15) ^*mulanggirr ngantha-ma-ya wirib \*  
 fierce 2SG>3SG-have-PRS dog  
 ‘You have a **dangerous** dog!’  
 (Schultze-Berndt and Simard 2012: 1035, emphasis in original)

The difference between continuous and true discontinuous nominal expressions marks a difference on the level of information structure. As Schultze-Berndt and Simard (2012: 1040–1) point out, discontinuity is often found when one element bears contrastive focus and the other element(s) is/are given, as in (15).

In the configurationality debate, cases of apparent discontinuity are sometimes used to argue that all languages show discontinuity and that there is accordingly nothing

<sup>20</sup> These are body part expressions, where the possessor bears the role of an argument, and the body part is linked to the possessor meronymically on a semantic level (see Schultze-Berndt and Simard 2012: 1031).

<sup>21</sup> For earlier associations of discontinuous nominal expressions with contrastive focus, see references in Schultze-Berndt and Simard (2012: 1038–9).

special about alleged cases of non-configurationality. An oft-quoted type discussed by Schultze-Berndt and Simard (2012) is so-called ‘split topicalizations’ as illustrated by the following German example:

- (16) *Zeitungen* liest er nur eine, die taz.  
 newspaper.PL read.PRS.3SG 3SG.M only one.SG.F DEF.SG.F [name]  
 ‘(As for) newspapers, he only reads one—the taz [newspaper name].’  
 (Fanselow and Ćavar 2002: 96, quoted in Schultze-Berndt  
 and Simard 2012: 1047)

Here, *Zeitungen* is a contrastive topic, whereas *eine* is focal, as argued by Schultze-Berndt and Simard (2012: 1047). Thus, we are not in fact dealing with a single nominal expression on a syntactic level; it is a case of apparent discontinuity. In contrast, example (15) is a single (discontinuous) nominal expression.

In the absence of prosodic evidence, it is even more difficult to disentangle apparent and true discontinuity. Perhaps unsurprisingly, the issue has hardly been explored systematically for Vedic and even renowned scholars at times differ in their interpretations of what is a complex nominal expression and what is not, as illustrated by the following alternative translations that can be found (from Bhat 1994: 176–7):

- (17) *sarvam khalv idam brahma*  
 all.NOM.SG.N indeed this.NOM.SG.N Brahma.NOM.SG.N  
 ‘All this is Brahma.’ vs. ‘The Brahma is all this.’

In section 2.4.1.2, I propose criteria that help to identify candidates for true discontinuity in Vedic despite prosodic evidence being unavailable.

### 2.3.3 Correlates of discontinuous nominal expressions

Assuming that true discontinuity exists, which may not be attributed solely to P2 effects and which can be delineated from cases of apparent discontinuity, two characteristics are often cited as correlates of discontinuous nominal expressions: the lack of obligatory function words in the nominal domain, and a weak or absent noun–adjective distinction. With regard to the first correlate and focusing on articles, Himmelmann notes:

[This implies] that languages in which nominal syntagms are generally ‘loosely’ structured (as for example in Latin or Nunggubuyu) do not have strongly adnominally grammaticalized D-elements (such as a definite article). To my knowledge, this claim is empirically correct. (Himmelmann 1997: 156–7, translation UR)<sup>22</sup>

<sup>22</sup> ‘[Dies impliziert] dass es in Sprachen, in denen nominale Syntagmen generell “locker” strukturiert sind (wie z.B. im Lateinischen und Nunggubuyu), keine stark adnominal grammatikalisierten D-Elemente (wie z.B. Definitartikel) gibt. Meines Wissens ist ... diese Behauptung empirisch korrekt.’

Indeed, non-configurational languages that have been studied, such as Warlpiri (Austin and Bresnan 1996: 234), Nunggubuyu (Heath 1986: 378) and Jaminjung (Schultze-Berndt and Simard 2012: 1019), lack obligatory function words in the nominal domain. These two findings are connected as follows: an obligatory function word such as an article or a (strongly grammaticalized) adposition opens up an adjacent syntactic slot to be filled by a nominal element. When such function words are lacking, there is nothing to constrain nominal expressions into patterns of rigid phrasality.

Turning to the second correlate, Baker summarizes prior claims regarding a connection between discontinuous nominal expressions and a weak noun–adjective distinction:

Indeed, Bhat (1994) and Baker (2000) claim that discontinuous constituents are possible only in languages with no more than a weak N/A contrast; in addition to the Australian languages, other languages of this type include Quechua (Lefebvre and Muysken 1988: 16265), Yimas (Foley 1991: 18091, 36976), Sanskrit (and perhaps Latin?) (Bhat 1994) and languages of the Klamath/Sahaptian family (Barker 1964: 33839). (Baker 2001: 1438)

A weak noun–adjective distinction typically involves the same morphological marking for all elements, especially regarding case. Receiving full morphological marking, all nominals are equipped to serve in the same syntactic functions and may occur on their own, as will be shown in section 2.4.1.3 for Vedic. While elements may typically be used preferably in referring or preferably in modifying function, and certain word formation processes may reflect this distinction, there is no sharp formal delineation into two syntactic categories with regard to distribution. By contrast, in a language with a sharp noun–adjective distinction such as English, an adjective cannot occur on its own in referring function and needs a dummy pronoun: *He bought a red one* (= a shirt).

In this study, I put these two correlates—a lack of function words and a weak noun–adjective distinction—to a diachronic test. It will become clear that they are connected to the rise of phrase structure, as well as to each other: the rise of postpositions will be shown to lead not only to the phrasal organization of nominal expressions, but also to a strengthening of the noun–adjective distinction, as will be shown in Chapter 8.

## 2.4 What is different between Vedic Sanskrit and Hindi?

In the following, I compare Vedic Sanskrit and Hindi with regard to their configurationality or lack thereof in order to work out the initial and final state of the language change I examine over the next chapters. First, I illustrate that Vedic Sanskrit is a non-configurational language in virtually all respects that have been discussed (section 2.4.1). Hindi, while still allowing for free constituent order and

null arguments, has developed low-level configurationality in the form of postpositional phrases, accompanied by a rise of obligatory function words and a sharpened noun–adjective distinction (section 2.4.2). In my outlines, I devote more space to Vedic because of the long-lasting controversy around the status of non-configurationality, outlining not only its syntax of nominal expressions, but also other aspects that bear on the matter. For Hindi, as a low-level configurational language less subject to discussion, I focus on nominal expressions.

#### 2.4.1 Vedic Sanskrit

While Vedic Sanskrit has at times been called non-configurational (see the quotation from Baker in section 2.3.3), this has not actually been substantiated in detail, as far as I am aware. Therefore, before I turn to the topic of nominal expressions which is of most interest for the present study (section 2.4.1.2), I first outline how Vedic Sanskrit indeed qualifies for non-configurationality (section 2.4.1.1). Finally, I add a section on the correlates of non-configurationality (section 2.4.1.3).

**2.4.1.1 Non-configurationality in Vedic Sanskrit** Numerous authors have mentioned the free word order of Vedic Sanskrit.<sup>23</sup> Since the linguistic research on Vedic has concentrated on the oldest available texts, it is typically material from the *Rigveda* (or other Mantra texts) that is cited as evidence. This practice has been criticized, since the variation may be due to the genre of the Mantra texts, a highly stylized and ritualized language, constrained into tight metrical patterns (cp. Hock 1997). In order to avoid this criticism, I cite from Vedic prose in this and in the next section, with only few complementary examples from Mantra Vedic. In this section, I test Hale’s criteria I and III, and comment on certain other phenomena that have been linked to non-configurationality. I discuss Hale’s criterion II in the next section, where I focus on nominal expressions.

Testing criterion I, it is of special interest whether we find, besides the ‘basic’ SOV order of Vedic, also OSV and VSO, the crucial patterns calling into question the existence of a verb phrase with the O argument as internal to it. Compare the following passage from the *Śatapathabrāhmaṇa* (cp. Delbrück 1878: 28):

- (18) *brāhmaṇá vai vayám smo rājanyābandhur asaú*  
 Bramin.NOM.PL.M PRT 1PL be.1PL friend\_of\_a\_prince DEM.NOM.3SG  
 ‘We (are) Brahmins, he (is) (some) connection of a prince

<sup>23</sup> There are several studies that address the free word order of Vedic Sanskrit. Apart from Delbrück (2009 [1888]) and Delbrück (1878), Staal (1967) and Schäufele (1991b) especially have made important contributions, see also the papers in Hock (1991), as well as Gillon (1996) or Gillon and Shaer (2005).

<i>yády</i>	<i>amúm</i>	<i>vayám</i>	<i>jáyema</i>
if	DEM.ACC.3SG	NOM.1PL	defeat.OPT.1PL
	O	S	V

if we defeat that one

<i>kám</i>	<i>ajaiṣméti</i>	<i>brūyāmātha</i>
kám	ajaiṣma íti	brūyāma átha
Q.ACC.SG.M	defeat.AOR.1PL QUOT	say.OPT.1PL now
O	V	

(we) will say: “Whom did (we) defeat?” Now,

<i>yády</i>	<i>asáv</i>	<i>asmán</i>	<i>jáyed</i>
if	DEM.NOM.SG	ACC.1PL	defeat.OPT.3SG
	S	O	V

if that one defeats us

<i>brāhmaṇán</i>	<i>rājanyābandhur</i>	<i>ajaiṣīd</i>	<i>íti</i>
Brahmin.ACC.PL.M	friend_of_a_prince.NOM.SG.M	defeat.AOR.3SG	QUOT
O	S	V	

“the Brahmins, the friend of the prince defeated”

*no brūyuh*  
1PL say.OPT.3PL  
(they) will tell us.’

(ŚaB 11.6.2.5)

In this passage, some Brahmins are contemplating who is going to defeat whom in a conflict between them and some other person, the friend of some prince, and how this will be perceived by others. The Brahmins, clearly the more powerful players here, are expected to win out against this friend. If, however, the latter should be successful in defeating the Brahmins, the incident will be noted by others as something unexpected. Twice we find an OVS pattern, specifying the respective defeatee, contrasting the two possibilities. However, we find an SOV sentence when the expected scenario, the Brahmins defeating the princely friend, appears in a backgrounded *if*-clause, without a special focus on either of the participants. This passage illustrates a typical case where the O-argument is drawn into initial position when focal and illustrates how the adjacency of verb and object may be disrupted.<sup>24</sup>

With regard to Hale’s criterion no. III, both null subjects and null objects are common phenomena in Vedic Sanskrit. As pointed out by Luraghi (2010), while the subject is cross-referenced on the verb through agreement, the object is not indicated at all when unexpressed. In defiance of Jelinek’s PAH hypothesis, this is nevertheless

<sup>24</sup> VSO patterns (interpreted strictly, i.e. without a single nominal or pronominal element indexing O placed between V and S) appear to be less frequent in Vedic prose, but see Klein (1991) and Viti (2009) on their usage in the *Rigveda*.

grammatical. In example (18), we already saw a clause with a missing subject, *kām ajaiṣma* ‘Whom did we defeat?’. In my corpus collected for this study, a large number of clauses do not have an overt subject, which is often dropped as the activated topic. The example below describes the ritual actions of the Adhvaryu priest, who performs a sacrifice for the benefit of someone else (the ‘sacrificer’) by means of a Soma (a psychedelic substance) ritual. The Adhvaryu, the sacrificer and the Soma are introduced earlier in the text, but then hardly mentioned explicitly any more over many paragraphs. Instead, one infers from context who is the S or O argument depending on the particular action narrated. For example, there is a passage which describes how a piece of cloth is used in the Soma ritual. From context, the contemporary audience would gather that this task is carried out by the Adhvaryu priest, even though he is not even pronominally expressed:

(19) (= ex. (3))

*átha somopanáhanasya samutpāryāntān*

*átha somopanáhanasya samutpārya ántān*

now Soma\_cloth.GEN.SG.N gather.CONV end.ACC.PL.M

*uṣṇīṣeṇa vígrathnāti ... atha madhye 'ngúlyākāśam*

head\_band.INS.SG tie\_together.3SG now MADHYE finger\_hole.ACC.SG.M

*karoti*

do.3SG

‘Having gathered up the ends of the Soma-cloth, (he) ties them together with a head-band..., (he) then makes a finger-hole in the middle (of the knot).’  
(ŚaB 3.3.2.19–20)

A few paragraphs earlier, we find a case where an object remains unexpressed. The sacrificer is described as first chanting mantras, then sitting down and touching the Soma, which is implied and non-overt:

(20) *athétya prāṅ upaviśati*

*atha itya prāṅ upaviśati*

now go.CONV eastwards sit\_down.3SG

‘Having gone (there), (he) sits down (behind the Soma) with his face towards the east.

*so 'bhimṛśati*

DEM.NOM.SG.M touch.3SG

**He touches** (the Soma)

*āsmākò 'śīti*

āsmāko 'śi iti

POSS.1PL be.2SG QUOT

‘You (oh Soma) are ours’ (he speaks).’

(ŚaB 3.3.2.6–7)

Such examples, in particular null subjects, occur with great frequency in Vedic prose. Long passages are devoted to the description of ritual, with detailed and repetitive narration of actions, where the participants and actions are well known.<sup>25</sup>

While structural issues of phrasal organization are the focus of this study, I now briefly sketch how Vedic behaves with respect to some further characteristics that have been linked to non-configurationality. A characteristic mentioned by Hale (1983) and related to the non-expression of arguments is a lack of expletives that would fill an obligatory subject position in a configurational language such as English, for example *it is raining*. For Vedic, Delbrück (2009 [1888]: 3–6) gives an overview of ‘subject-less verbs’ including weather verbs and other types that occur without a subject. No dummy pronoun is used in Vedic, as there is no constraint that the slot must be filled. This is further evidence for the lack of phrasality and hence for the non-configurationality of Vedic. But there are a few areas in which the question of subject–object asymmetries does not have a straightforward answer. Regarding alignment, basic verbal constructions show a nominative–accusative pattern, but there is also a variety of constructions involving non-canonical subjects, and it has been argued that the *ta*-construction already shows ergative–absolutive alignment in Old Indic (e.g. Hock 1986), which may be regarded as attesting to an alternative type of subject–object asymmetry. By contrast, reflexive constructions and infinitival control appear to at least in part hinge on semantic, rather than syntactic role.<sup>26</sup> In specific types of control construction, again, we find evidence for a subject relation. Delbrück (2009 [1888]) and Hock (1986) outline how converb constructions (‘absolutives’ in the tradition of Western grammar-writing on Sanskrit) are controlled by subjects in early Vedic,<sup>27</sup> which only in younger stages develops into agent control (Hock 1986). Compare the following examples with an active (ex. (21)) and a passive (ex. (22)) matrix verb in passages from the *Rigveda* (see Hock 1986).

- (21) *eṣā*                      *syā*                      *nāvyam*                      *āyur*                      *dādhānā*  
 DEM.NOM.SG.F    DEM.NOM.SG.F    new.ACC.SG.N    life.ACC.SG.N    put.PPRM.NOM.SG.F  
*gūḍhvī*                      *tāmo*                      *jyōtiṣoṣā*                      *abodhi*.  
 hide.CONV    darkness.ACC.SG.N    light.INS.SG.N    awaken.AOR.3SG  
 ‘This one (the goddess of the dawn), offering new vitality, having hidden away  
 darkness, awakes with light.’ (RV 7.80.2)

<sup>25</sup> For null arguments in Mantra Vedic see Keydana and Luraghi (2012).

<sup>26</sup> Besides the possible reflexive readings of the middle voice, *svā* typically indicates a possessor and only sometimes allows for a reflexive reading, *tanū* appears in a number of ambiguous passages where it may literally mean ‘body’ rather than be used as a reflexive, and the same goes for *ātmān*, which in Vedic typically literally means ‘breath, spirit, soul, self’, and only gradually develops into a reflexive (see Hock 2006 and Kulikov 2007 for critical assessments). For infinitival control, see Keydana (2013).

<sup>27</sup> Delbrück (2009 [1888]: 406–7) notes a few examples where a nominative subject is lacking and where we find control by the ‘logical subject’, i.e. an experiencer in the cases listed.



- (22) *sārvam parigṛhya sūyū iti*  
 all.ACC.SG.N give\_up.CONV consecrate.PASS.1SG QUOT  
 ‘Having encompassed everything, may (I) be consecrated.’  
 (ŚaB 5.2.3.1, cp. Hock 1986: 22)

In sum, Vedic Sanskrit fulfils Hale’s criteria I and III, lacks expletives, and there are several domains in which subject–object asymmetries are not found or only in part. Only in a small pocket of the grammatical system, for example in the control of converb constructions, it appears that we can identify an asymmetry. Future research may reveal further domains in which the absence or presence of asymmetries or other phenomena connected to non-configurationality can be observed.

2.4.1.2 *Non-configurationality and nominal expressions in Vedic Sanskrit* In the literature, it is often asserted that Vedic has discontinuity. While there is certainly an enormous number of examples of clauses where several nominal (in a wide sense, i.e. nouns, adjectives, pronouns, quantifiers, besides other elements that can occur as parts of a nominal expression) elements have the same morphological marking for case, gender, and number, but are discontinuous, it is important to show that this involves true discontinuity. As far as I am aware, this has not yet been demonstrated. In fact, authors claiming that Vedic is non-configurational or has free order sometimes choose examples that may be explained away as clitic insertion (into P<sub>2</sub> or otherwise) or as apparent discontinuity. I first discuss P<sub>2</sub> effects<sup>28</sup> and then turn to apparent discontinuity.

Pronominal elements that are co-referential with nominals can be inserted into P<sub>2</sub>, resulting in discontinuity as in example (23) (a detailed outline of the internal structure of P<sub>2</sub> and its historical development will be given in Chapter 6).<sup>29</sup> Also, P<sub>2</sub> elements can appear to ‘split’ expressions as in example (24). Compare:

- (23) *yābhyah<sub>i</sub> eva tāni<sub>j</sub> devatābhyah<sub>i</sub> havīmṣi<sub>j</sub>*  
 REL.DAT.PL.F PRT DEM.NOM.PL.N god.DAT.PL.F oblation.NOM.PL.N  
*grhītāni bhavanti*  
 take.PPP.NOM.PL.N be.3PL  
 ‘for which deities **those oblations** are taken’ (AB 7.2.3)

<sup>28</sup> The insertion of emphatic clitics into nominal expressions in places other than P<sub>2</sub> might also be discussed (for illustration, see the Pali ex. (135) in Chapter 6). However, since clitic insertion remains possible even in Hindi, which shows a full-fledged postpositional phrase (cp. Sharma 1999), I disregard such cases.

<sup>29</sup> Vedic has a P<sub>2</sub> position allowing for different pronominal elements, discourse particles, conjunctions and other elements, which may be combined. It is particularly complex and frequent in Vedic prose, where it consists of up to five internal slots as described by Hock (1982) and others (see Chapter 6 for details and in particular fn. 17 regarding the question of whether P<sub>2</sub> is constrained syntactically or prosodically).

- (24) *madhye ha saṁvatsarasya svargo lokah*  
 MADHYE PRT year.GEN.SG.M sky.GEN.SG.M world.NOM.SG.M  
 ‘(For one half year, he worships the fire with the Vātsapra,) for the world of  
 heaven is in the middle of the year.’ (ŚaB 9.3.1.10)

The insertion of pronominals into P2 may indeed be regarded as linked to the behaviour of pronominal clitics, rather than to the syntax of nominal expressions. However, what about P2 allegedly splitting nominal expressions?<sup>30</sup> I propose that this is not the correct analysis for Vedic, since a nominal expression does not necessarily form an immediate frame around P2, as illustrated by example (25). In such cases, therefore, we are potentially dealing with true discontinuity:

- (25) *māno ha vai devā manuśyāya ājānanti*  
 mind.ACC.SG.N PRT PRT god.NOM.PL.M man.GEN.SG.M know.3PL  
 ‘The gods know the mind of man.’  
 (ŚaB 1.1.1.7, adapted from SchāufeLe 1991b: 78)

I come to cases of apparent discontinuity.<sup>31</sup> Similar to patterns in Jaminjung illustrated in section 2.3.2, this involves afterthoughts (ex. (26)).<sup>32</sup> In the absence of prosodic evidence, I classify afterthoughts as added nominal material outside of the frame of the finite clause established by the S or A argument (if overtly expressed) and the finite verb. We also find secondary predicates as in example (27). Here, *āpramūrā* ‘not foolish’ shows the same morphological marking as *te* ‘they’ (anaphoric with the ‘preservers of wealth’), describing them in the context of the event expressed. For the possibility of an attributive reading see the discussion in Krisch (2005: 308).

- (26) *indhe ha vā etad adhvaryuḥ idhmenāgniṁ*  
 indhe ha vā etad adhvaryuḥ idhmena agniṁ  
 ignite.3SG PRT PRT DEM.ACC.SG.N Adhvaryu.NOM.SG.M fuel.INS.SG.N fire.ACC.SG.M  
 ‘The Adhvaryu lights it, that fire, with a kindling.’  
 (ŚaB 1.3.5.1, adapted from SchāufeLe 1991b: 44)

- (27) *té hí vásvo vásavānās*  
 DEM.NOM.PL.M because wealth.GEN.SG.N preserver\_of\_wealth.NOM.PL.M  
*té āpramūrā māvobhiḥ / vratā*  
 DEM.NOM.PL.M non\_foolish.NOM.PL.M greatness.INS.PL.N norm.ACC.PL.N

<sup>30</sup> A ‘splitting’ analysis is assumed by Taylor (1990) for Ancient Greek (see Taylor 1990: 72–3, with further references).

<sup>31</sup> A combination of P2 and apparent discontinuity is particularly frequent when nominal predicates are placed initially, as in ex. (17).

<sup>32</sup> Cp. Gonda’s (1959) ‘amplified sentences’. Afterthoughts may or may not be co-referential with other elements in the clause, and, when co-referential, may or may not show the same morphological marking.



(30) (= ex. (25))

*māno*                      *ha*   *vaí*   *devā*                      *manuṣyāśya*   *ājānanti*  
 mind.ACC.SG.N   PRT   PRT   god.NOM.PL.M   man.GEN.SG.M   know.3PL  
 ‘The gods know **the mind of man**.’

(ŚaB 1.1.1.7, adapted from Schäufele 1991b: 78)

In sum, even when excluding cases such as pronominal insertion (the first type of P2 effect), and cases of apparent discontinuity, we are still left with cases of discontinuity, that is true discontinuity. First, this may involve the second type of P2 effect where one element is drawn into the clause-initial position and other co-referential elements follow later in the clause, since a ‘splitting’ account of P2 elements can be rejected. Second, it may involve clauses where a P2 effect is not involved in discontinuity and where the finite verb can be taken as an important demarcation point at the right clausal edge, which allows identifying and excluding afterthoughts, for example. Third, it may involve any complex nominal expressions containing an adnominal genitive standing apart (exempting pronominal possessors inserted in P2).

2.4.1.3 *Correlates* Vedic Sanskrit confirms the hypothesis that true discontinuity correlates with an absence of obligatory function words in the nominal domain, as well as with a weak noun–adjective distinction. With regard to the first correlate, we find no obligatory function words such as articles or adpositions that would impose phrasal constraints on nominal expressions. The elements that come closest to such categories, demonstratives and local particles (which are semantically basic and highly frequent spatial adverbs), have an independent denotation and may occur on their own. Compare example (31), which shows two demonstratives used pronominally (see also (26)), and example (32), which illustrates the usage of a local particle on its own:<sup>34</sup>

(31) (Vedic Sanskrit)

*saḥ*                      *enam*                      *svargam*                      *lokam*  
 DEM.NOM.SG.M   DEM.ACC.SG.M   heaven.ACC.SG.M   world.ACC.SG.M  
*abhivahati*                      *kṣatram*                      *ca*   *balam*                      *ca*  
 carry\_towards.3SG   dominion.ACC.SG.N   and   strength.ACC.SG.N   and

possessor phrases headed by the postposition *kā/-e/-ī* may be placed apart (see the examples in section 2.4.2). What is important for this study, however, is that the old genitives encoding possessors, while enjoying freedom of placement in Old Indic, over time come to stand obligatorily adjacent to relational nouns—the origins of several of the modern postpositional constructions, as will be shown in Chapters 5 and 6. The new layer of possessor phrases in Hindi is a recent, secondary development.

<sup>34</sup> *Ādhi* in ex. (32) does not modify the locative, which is used in its usual function here. An example from the *Rigveda* instead of a prose example is chosen because the local particles vanish as free forms soon after early Vedic, see Chapter 4.

*rāṣṭram*                      *ca*      *viśam*                      *ca*  
 empire.ACC.SG.N      and      tribe.ACC.SG.F      and  
 ‘He carries him toward the heavenly world, to dominion, to strength, to  
 empire, to tribe.’                      (AB 8.24.7, adapted from Schäufele 1991b: 83)

- (32) *eṣā*                      *sūryeṇa*                      *hāsate*                      *pávamāno*                      *ádhi*  
 DEM.NOM.SG.M      sun.INS.SG.M      race.MID.3SG      purify.PPM.NOM.SG.M      above  
*dyávi*  
 sky.LOC.SG.M  
 ‘This one, purifying himself, races with the sun, **above**, in the sky.’  
 (RV 9.27.5, adapted from Hettrich 1991)

Turning to the second correlate, the noun–adjective distinction is notoriously weak in Vedic Sanskrit. There are mainly two distinctions between the two categories, namely that adjectives show gender agreement, and are more frequently used in comparative and superlative formations. However, this categorization is not strict and many elements, not only those denoting qualities, inflect for more than one gender, and even proper names allow for comparative and superlative forms, for example *índratama*- ‘Indra-most’ (see Delbrück 2009 [1888]: 193–4). In addition, two nouns that would usually be used referentially each may co-occur, where one modifies the other, depending on the context. Derivational morphemes exist, but are not necessary, as can be seen in example (31), where *svarga*- ‘heaven’ and *loka*- ‘world’ construe together. Importantly for this study, nominal elements are on a par in that they all receive the full set of morphological values (i.e. case, gender, number), irrespective of whether there is a co-referential element beside them or not. This is a difference from several of the Australian languages that have been described as non-configurational, for example Nunggubuyu (Heath 1986) or Jaminjung (Schultze-Berndt and Simard 2012: 1021), where case markers may be omitted when co-referential elements stand in juxtaposition.<sup>35</sup> Also, all nominal elements may occur without a referential ‘head’, even if they denote qualities, as in (33):

- (33) *tābhiḥ*                      *jvalantībhiḥ*                      *cīpyamānābhiḥ*                      *upauteti*                      *rājānam*  
 DEM.INS.PL.F      flame.PPA.INS.PL.F      shine.PPA.INS.PL.F      approach.3SG      king.ACC.SG.M  
 ‘With those **flaming, shining** (weapons) (he) approaches the king.’  
 (AB 8.24.6, adapted from Schäufele 1991b: 57)

<sup>35</sup> In Nunggubuyu, while it is possibly the most strongly non-configurational language that has been described (see section 2.2), Heath (1986) outlines how juxtaposition of co-referential nominals may have an effect on marking. Juxtaposed nominal elements may all separately receive case marking, but they may also share a marker which appears on either one of the elements. Furthermore, demonstratives, when juxtaposed to the nominal they are co-referential with, optionally drop an anaphoric suffix. While these are not obligatory syntactic rules, but only optional possibilities of omission, more of which are outlined by Heath, this attests to weak phrasality effects; a phenomenon not found in Vedic, however.

In accordance with the syntactically co-ranking status of nominal elements, there is no fixed relative order of modifying and referential elements (if this distinction can be drawn in the example given). Delbrück (2009 [1888]: 19–20) observes that modifying elements typically precede referential ones. However, this is only a tendency and the opposite is attested. In the example below, the string *bṛhadúkṣaḥ marútaḥ viśvāvedasaḥ*, with the proper name in the middle and one modifier on each side, shows that either order can be found in a single nominal expression.

- (34) *bṛhadúkṣaḥ*                      *marútaḥ*                      *viśvāvedasaḥ*  
 richly\_raining.NOM.PL.M    Marut.NOM.PL.M    owning\_all\_treasure.NOM.PL.M  
*prāvepayanti*    *párvatān*                      *ádābhyāḥ*  
 shake.3PL                      mountain.ACC.PL.M    invulnerable.NOM.PL.M  
 ‘The richly-raining Maruts, owning all treasure, shake the mountains, the invulnerable ones.’

(RV 3.26.4, adapted from Schäufele 1991b: 58)

In sum, the characteristics outlined in this and in the preceding sections locate Vedic Sanskrit close to the pole of thorough non-configurationality. In fact, it can be considered more strongly non-configurational than head-marking languages such as Warlpiri, since even O arguments can be dropped, even though there is no cross-reference on the verb (cp. Luraghi 2010). Moreover, even if nominal elements sharing a syntactic role stand in juxtaposition, case marking is not omitted. With regard to domains other than phrase structure, there are only small pockets in which we may detect subject–object asymmetries.

#### 2.4.2 Hindi

Hindi continues to fulfil Hale’s criteria I and III, as it allows for free constituent order (exx. (35)–(38)<sup>36</sup> from Mohanan 1994: 11–12) and null anaphora including the omission of the O argument (see ex. (39), from Simpson et al. 2013: 106):

- (35) *Ilā ne Anū ko hār bhejā*  
 Ilā ERG Anu DAT necklace send.PRF  
 ‘Ila sent Anu a/the necklace.’  
 (36) *Ilā ne hār Anū ko bhejā*  
 ‘Ila sent Anu the/\*a necklace.’  
 (37) *hār Ilā ne Anū ko bhejā*  
 ‘Ila sent Anu the/\*a necklace.’

<sup>36</sup> While equivalent regarding the basic proposition being expressed, there are differences in definiteness and, naturally, in information structure, as indicated in the translations.

- (38) *Ilā ne bhejā Anū ko hār*  
 ‘(It was) Ilā (who) sent Anu the/a necklace.’
- (39) (i) *Amit apnī premikā ko pyār kartā hai*  
 Amit his girlfriend ACC love do.IPFV.SG.M be.3SG  
 ‘Amit<sub>i</sub> loves his<sub>i</sub> girlfriend.’
- (ii) *Ravi bhi pyār kartā hai*  
 Ravi also love do.IPFV.SG.M be.3SG  
 ‘Ravi<sub>i</sub> loves (his<sub>i</sub> girlfriend) also.’

While Hindi retains these characteristics of non-configurationality, the organization of nominal expressions is fundamentally altered in comparison with Vedic. Nominal expressions, whether marked by a postposition (which not only marks adjuncts, but also certain arguments)<sup>37</sup> or not, show a fixed phrasal template (in reduced form) DEM QUANT MOD NOM (POSTP). In contrast to Vedic Sanskrit, Hindi allows neither for discontinuity (apart from specific cases: see below) nor free permutation within nominal expressions:

- (40) HINDI  
*Rajesh ne choṭe bacce ko kitāb bhejī*  
 Rajesh ERG small.OBL.SG.M child.OBL.SG.M DAT book.DIR.SG.F send.PFV.SG.F  
 ‘Rajesh sent the book to the little child.’  
 \**Rajesh ne bacce choṭe ko kitāb bhejī*  
 \**Rajesh ne choṭe ko bacce kitāb bhejī*  
 \**choṭe Rajesh ne bacce ko kitāb bhejī* etc.

Apart from the nominal head (which can also be a nominalized adjective, a pronoun, etc.), none of the elements which may form part of the postpositional template occurs in every nominal expression. A postposition, however, requires a dependent, an obligatoriness constraint that has no equivalent in Vedic.<sup>38</sup> Certain elements may occur detached from the nominal expression they belong to syntactically, such as

<sup>37</sup> The seven simple postpositions, which will be introduced in detail in Chapter 3, mark arguments, adnominal possessors, and certain spatial adjuncts. In addition, we find complex postpositions, consisting of the possessive postposition and a noun, as well as compound postpositions, consisting of two simple postpositions (see Chapter 8 for details).

<sup>38</sup> While in earlier historical stages, relational nouns such as *MADHYE* clearly were the nominal head taking a genitive as dependent, the relations of headedness have changed in Hindi. Since the oblique form has few independent functions and almost only occurs when triggered by a postposition (see section 8.3.2.2), one may adopt the stance that the oblique-marked noun is not a dependent of the postposition any more, but a nominal head with a case marker (i.e. the postposition) attached to its right. However, in view of the general diachronic setting of this study and in order to avoid terminological confusion, I continue to refer to oblique-marked nominals as being dependents of postpositions. The few functions that the bare oblique does retain support this approach (also adopted by Spencer 2005, for instance).

quantifiers (cp. Raza and Ahmed 2011) or possessor phrases.<sup>39</sup> While the former are well known for their word order freedom cross-linguistically, the latter by themselves constitute postpositional phrases marked by the postposition *kā/-e/-ī* (agreeing with the possessee in gender, number, and morphological case), which explains their syntactic freedom:

- (41) *us kā, mujhe lagtā hai ki*  
 DEM.OBL.SG POSS DAT.1SG seem.IPFV.SG.M be.3SG COMP  
*pati hī bevakūf hai*  
 husband.DIR.SG.M EMPH stupid be.3SG  
 ‘It seems to me that **her husband** is stupid.’ (adapted from Gambhir 1981: 224)
- (42) *tumhē merā kārd mil gayā thā, dīvālī kā?*  
 DAT.2SG POSS.1SG.M card meet go.PFV.SG.M be.PST.SG.M Diwali POSS  
 ‘Did you receive **my Diwali card**?’ (adapted from Gambhir 1981: 321)

Thus, the essential difference between Vedic Sanskrit and Hindi is that nominal expressions were not phrasally organized in the former, but are so in the latter. Note that this does not mean that all elements of nominal expressions always form an integral syntagm in Hindi (as illustrated by the possessor phrases above), or that there is no permutability whatsoever (see Gambhir 1981 for discussion), but that there is a template that underlies a great number of complex nominal expressions, which also involves certain obligatoriness constraints, such as that a postposition always requires a nominal dependent.

Besides the rise of postpositions, there have been changes regarding the noun–adjective distinction. In Hindi, there are separate slots for referential elements (directly adjacent to the postposition) in contrast to modifying elements, as well as slots for further elements, such as demonstratives and quantifiers. As pointed out above, while the referential slot is typically filled by nouns, it can also host ‘nominalized’ adjectives, demonstratives, and certain other forms (e.g. oblique-marked infinitives). A postposition governs the oblique case of its head noun (if distinctive), as well as of a modifying masculine singular *a*-stem adjective (ex. (44)). Other types of adnominal elements do not show distinctive forms for direct and oblique (exx. (46), (48), (50)). Nominal expressions without postpositions stand in the direct case with the exception of certain adverbial usages (exx. (43), (45), (47), (49)).<sup>40</sup> (Note that the convention is to gloss modifying adjectives as ‘direct’ and ‘oblique’, depending on the absence or presence of a postposition, irrespective of form.)

<sup>39</sup> This point was brought up by one reviewer.

<sup>40</sup> A note on terminology is in order here. I will at times refer to the direct–oblique distinction as well as to the simple postpositions as marking case. The direct–oblique distinction still in part indicates a case distinction, albeit in very small pockets of the grammar. Otherwise, the direct contrasts with the simple postpositions in marking arguments (see the discussion in section 8.3.2.2).



- (43) *choṭ-ā*                      *kamr-ā*  
 small-DIR.SG.M    room-DIR.SG.M  
 ‘a/the small room’
- (44) *choṭ-e*                      *kamr-e*                      *mē*  
 small-OBL.SG.M    room-OBL.SG.M    in  
 ‘in a/the small room’
- (45) *choṭ-e*                      *kamr-e*  
 small-DIR.PL.M    room-DIR.PL.M  
 ‘small rooms’
- (46) *choṭ-e*                      *kamr-ō*                      *mē*  
 small-OBL.PL.M    room-OBL.PL.M    in  
 ‘in (the) small rooms’
- (47) *choṭ-ī*                      *almār-ī*  
 small-DIR.SG.F    cupboard-DIR.SG.F  
 ‘a/the small cupboard’
- (48) *choṭ-ī*                      *almār-ī*                      *mē*  
 small-OBL.SG.F    cupboard-OBL.SG.F    in  
 ‘in a/the small cupboard’
- (49) *choṭ-ī*                      *almār-iyā*  
 small-DIR.PL.F    cupboard-DIR.PL.F  
 ‘(the) small cupboards’
- (50) *choṭ-ī*                      *almār-iyō*                      *mē*  
 small-OBL.PL.F    cupboard-OBL.PL.F    in  
 ‘in (the) small cupboards’

In sum, Hindi has developed phrasal nominal expressions, obligatory function words in the nominal domain, and shows a strengthened noun–adjective distinction, marked by distributional correlates with the referential and the modifier slot respectively. Having outlined the starting and the end points of the developments to be surveyed, I now turn to a more general perspective on configurationality as a historical phenomenon.

## 2.5 Configurationality as a historical phenomenon

The fact that Latin and Ancient Greek only possess incipient phrasal structures, and different ones at that (Latin having prepositional phrases and Ancient Greek developing nominal expressions involving articles), together with the fact that Vedic

shows a lack of such structures, has long been interpreted as indicating that these must have been recent developments.<sup>41</sup> Several proposals have been offered as to how Ancient Greek and Latin and other languages acquired these structures and developed them further, expanding to other syntactic domains, in the course of their histories. There is unanimity with regard to the general outlines of the scenario, whether we are dealing with local particles turning into prepositions in Latin or with demonstratives turning into articles in Ancient Greek.<sup>42</sup> Several authors have described how syntactically independent and co-ranking elements with a shared reference, for example local particles and local case forms, or demonstratives and nominals (typically in core cases), frequently co-occurred in a sentence. They would often stand adjacent to each other in accordance with Behaghel's principle that what belongs together semantically also stands together. Such syntagmatic strings of elements that are co-referential, but are not constrained by phrase-structural principles, are referred to as *groups* (Lehmann 1991; Himmelmänn 1997), a term I adopt in this study. In the case of syntactically co-ranking elements such as those just described, we are dealing with symmetrical groups. At some point, elements would co-occur in such a symmetrical group so frequently that the string is reanalysed as a single syntactic unit, that is as a phrase. On the expression side, the local particles and demonstratives develop from autosemantic elements into synsemantic elements and cease to occur on their own. Vice versa, the nominal elements come to require these operators in certain usages. For example, while the Latin accusative could still be used on its own to encode a direct object, it ceases to occur on its own in its local case function of encoding goals, a usage in which it came to require the preposition *ad*, a former local particle.

There is a rich amount of evidence that this scenario adequately accounts for the rise of prepositional phrases in many branches of Indo-European, as already sketched by Meillet and Vendryes (1927: 497) and as recently studied by Hewson and Bubeník

<sup>41</sup> Regarding the ultimate source of Proto-Indo-European, even though languages such as Latin or Ancient Greek possess a certain phrasal structure, it appears justified to assume that PIE itself was non-configurational. First, Vedic attests to full-fledged non-configurationality lacking any syntactic phrasality, even more so than some of the Australian languages commonly cited as non-configurational. Second, the fact that different branches of Indo-European develop different types of phrasal structures first suggests that these are not shared heritage. Third, correlates of non-configurationality, such as a weak noun–adjective distinction, are shared by many of the ancient varieties, which does suggest shared heritage (see Luraghi 2010: 219 for references). Therefore, even though Hittite constituent order, for instance, is quite fixed, as one reviewer points out, the cumulative evidence strongly suggests that PIE, too, was non-configurational.

<sup>42</sup> Outside of Indo-European, changes from non-configurationality (or only incipient phrasal structures) to low-level configurationality do not seem to have been documented. Within Indo-European, studies dealing with the phenomenon are Ledgeway (2012) on Italic, Taylor (1990) on different stages of Ancient Greek, Hewson and Bubeník (2006) concentrating on the rise of adpositions in all major branches of the Indo-European family, Rögnvaldsson (1995) on Icelandic, and Luraghi (2010) offering a general overview of the rise of configurationality in the language family.

(2006) and Luraghi (2010), and it equally accounts for the development of groups of demonstratives and nominals into noun phrases (see Himmelmann 1997 for several unrelated languages). I will propose in this study, however, that there is an alternative pathway to phrasality, attested by the rise of postpositional phrases in Indo-Aryan. Contrary to what has been claimed (see for instance Hewson and Bubeník 2006), the difference between Indo-Aryan and other branches of Indo-European is not a superficial difference of preposed vs. postposed adpositions. Rather, the Indo-Aryan postpositional phrases do not develop from symmetrical groups of syntactically unconnected, co-ranking elements, but instead from asymmetrical groups that consist of a nominal or verbal head and a dependent.

Returning for a moment to the well-known origin in symmetrical groups, changes such as the reanalysis of combinations of local particles and local case forms, or demonstratives and nominal elements into phrasal units have been interpreted in a number of different ways on a theoretical level. The major proposals that have been put forward are summarized in Ledgeway (2012) in the context of his study of syntactic changes from Latin to modern Romance languages. Ledgeway outlines and discusses major theoretical stances on the loss of morphological marking and the rise of phrasal constructions, such as the change from syntheticity to analyticity, the rise of constituent structure and/or functional structure, change in the head directionality parameter, and change from dependent-marking to head-marking. Ledgeway evaluates each approach in turn with regard to its adequacy in explaining the historical evidence. He states that certain configurational structures already existed in Latin, namely the prepositional phrase and, less developed, the complementizer phrase.<sup>43</sup> In addition, there are studies arguing that demonstratives were developing characteristics of a specifier in Latin (see Giusti and Oniga 2007; Giusti and Iovino 2011; Iovino 2011 on demonstratives and the status of nominal expressions). The further development of these incipient phrasal structures as well as the new creation of phrases is connected by Ledgeway to a change in head directionality.

In this study, I take a somewhat different approach to the rise of postpositional phrases in Indo-Aryan, different because the evidence suggests another pathway to phrasality, as will be laid out in detail in Chapters 4 and 5, and different because I focus on the developments of the individual postpositions, and only towards the end of this study turn to the more abstract level of the emergence of a new class of function words. This approach is motivated by the data: as shown in the next chapter, the simple postpositions of Hindi descend from diverse word classes, that is from elements with diverse semantic and morphosyntactic properties. Accordingly, the pathways of their grammaticalizations differ markedly between them, and they converge into a single new part of speech only at a relatively late stage, as will be

<sup>43</sup> Giampolo Salvi, according to the presentation of his arguments in Ledgeway (2012: 154, 192), argues that subordinators placed in P2 in Latin project a CP.

shown in Chapter 8. The variation in their original semantic and morphosyntactic properties, as well as the ensuing differences in their development into members of the modern postpositional paradigm, are such that they cannot be captured by a generalized change, whether conceived of as pertaining to a category or to a structural parameter. Instead, it is the grammaticalizations of the individual postpositions along different trajectories which bring about the emergence of the new word class and, as a consequence, of the first phrasal category in this branch of languages. As pointed out, these trajectories will be shown to differ on a structural level from those attested in other branches of the language family, originating in asymmetrical groups and not in symmetrical groups. The contrast between the developments in Italic and other western branches, on the one hand, and the developments in Indo-Aryan, on the other hand, will highlight how the historical syntax of Indo-European as well as the origins of configurationality in general are much more varied than previously assumed.

## 2.6 Grammaticalization and the development of phrasal structures

The central hypothesis of this study is that the grammaticalization of postpositions entails the rise of phrasal organization of nominal expressions in modern Indo-Aryan languages, paralleling Himmelmann's (1997) claim that the grammaticalization of articles can be a source of phrasal organization in nominal expressions. In the literature on grammaticalization, remarks are found that pertain to the emergence of phrasal organization, but this link is usually not made explicit. A major hindrance in this regard is that authors often adopt an element-based approach to grammaticalization, rather than a construction-based one, which in my view precludes an actual understanding of the origins of phrase structure in non-configurational languages.<sup>44</sup> Consider, for instance, Lehmann's (2002 [1982]) comments on the obligatorification of dependents, as well as on the fixation of order. First, with regard to selectional restrictions, Lehmann describes a 'narrowing down' of combinations:

Further grammaticalization of the adpositional phrase involves a **narrowing down of the choice of the case** for the NP. Some Latin prepositions are like *sub* in allowing either the ablative or the accusative; but most of them invariably require just one of these cases. (Lehmann 2002 [1982]: 82, emphasis added)

<sup>44</sup> In the majority of cases of grammaticalization that have been studied, there already are certain phrasal structures in place. For example, the earliest examples ambiguous between a lexical and a grammaticalized interpretation for *be going to* VP identified by Eckardt (2006: 98–100) already show the syntagm still found today, only allowing for general movement operations such as auxiliary inversion, e.g. *Stout resolved mates! Are you now going to dispatch the deed?* (Shakespeare, *Richard III*: I.iii.341), *Sir we want a service, and are going to get a Master* (Drama Corpus: E. Sharpham, *The Fleire*, 1606: II).

Second, Lehmann outlines the reduction of 'syntagmatic variability' in the course of grammaticalization:

**The syntagmatic variability of a sign is the ease with which it can be shifted around in its context. In the case of a grammaticalized sign, this concerns mainly its positional mutability with respect to those constituents with which it enters into construction. Syntagmatic variability decreases with increasing grammaticalization.**

The grammaticalization of adverbs to adpositions provides an example. An adverb which specifies an aspect of a local NP may often be juxtaposed to it on either side and sometimes even be separated from it. **The more intimate its connection with the NP becomes, the more its position vis-à-vis the latter becomes fixed; it develops either into a preposition or into a postposition.** (Lehmann 2002 [1982]: 140–1; emphasis added)

These quotations make clear that Lehmann here does not assume a construction-based view (or at least not a construction-based view that includes fixed order), but the construction only emerges in the course of grammaticalization. I propose, by contrast, that it is essential to maintain the construction-based view, even when faced with a non-configurational language whose lack of obligatoriness constraints and free word order may at first suggest otherwise. Whether non-configurational or not, it has been shown that the entrenchment of certain fixed syntagms is a prerequisite for grammaticalization. Processes of ritualization and habituation (Haiman 1994) can turn a construction into 'a single processing unit' (Bybee 2003: 603). This allows the relations between the elements in a construction to alter in such a way that a lexical element can be reanalysed as synsemantic (see Eckardt 2006: 115 for a similar view cast in formal semantic terms). Relating to word order, the importance of a syntagmatically fixed construction is stressed in Hopper's account of the formation of the French future:

**A prerequisite of grammaticalization is the fixing of a habitual order, a preference for one out of several possibilities.** The predecessor of French [*je*] *chanterai une chanson*, 'I will sing a song' would have been a Latin phrase *habeo canticulum cantare*, 'I have a song to sing, I have to sing a song', whose component words could occur freely in any order (*canticulum cantare habeo*, etc.). Even in Old and Middle French, some flexibility was possible between *cantar ayo* and *ayo cantar*, 'I have to sing, I will sing' (4, 34, 35). **But at the stage preceding the grammaticalization of *cantar ayo* as *chanterai* the order of the infinitive *cantar* and the auxiliary *ayo* has become fixed, and no variation is possible.** (Hopper 1996: 223, emphasis added)

Thus, rather than including data of the source element in all its environments as a part of grammaticalization, we must ask ourselves how the input construction for grammaticalization gets singled out from other possible syntagms. Accordingly, the central question addressed in this study is how grammaticalization phenomena can give rise to phrase structure, introducing syntactic constraints of obligatoriness and linearity for the first time. In contrast with Hopper's account, I will propose in

Chapter 7 that it is not so much the case that fixation of order precedes grammaticalization—rather, grammaticalization only takes place in a specific syntagm. I will show how semantic shifts first lead to the obligatorification of dependents, and then to the selection of a specific syntagm in which grammaticalization takes place.

The overarching hypothesis of this study, that the grammaticalization of postpositions entails the emergence of phrase structure, can be elaborated into the following three principle claims:

1. *The rise of the modern postpositional phrases hinges on the grammaticalizations of the individual postpositions.* Their pathways of change differ according to their heterogeneous origins and only from the dawn of the New Indo-Aryan period onwards do the individual elements show evidence for paradigmaticization into the integrated syntactic class we find today. After an outline of the respective etymologies in Chapter 3, the individuality of the grammaticalization paths will be outlined in Chapters 5, 6, and 7, and paradigmaticization in recent centuries will be studied in Chapter 8.
2. *Postpositional phrases in Indo-Aryan develop from asymmetrical constructions.* This adds an alternative to the well-attested pathway of a development of phrasality from symmetrical constructions consisting of syntactically independent and co-ranking, while semantically co-referential, elements. The structural differences between both pathways will be worked out in Chapter 5, and the morphosyntactic and semantic details of an origin in asymmetrical constructions will be outlined in Chapters 6 and 7.
3. *The formation of syntactic phrases is triggered by semantic changes involving metaphorical as well as metonymic transfers from the literal meaning of the constructions in question.* This regards the aspects of obligatoriness and juxtaposition. The precise way in which metaphorical and metonymic changes entail the obligatorification of the nominal dependent will be studied in Chapter 7. With regard to the specific postnominal placement, I argue that it has an important source in constructions involving pronominal possessors, from which it spreads to other constructions, a development which will be outlined in Chapter 6 as well as Chapter 7.

## 2.7 Summary

In this chapter, I have brought together the two major research areas that frame this study: grammaticalization and configurationality. I first defined grammaticalization as the change from autosemantic to (more) synsemantic, and then outlined aspects of grammaticalization phenomena central to this study. Focusing on early stages in grammaticalization processes, I outlined how I conceptualize grammaticalization on two levels, both as discrete and as continuous, and I will adopt either perspective in

this study, depending on the respective sub-process discussed. I then spent some time discussing non-configurationality both with respect to the varied history of its research and especially relating to the status of discontinuous nominal expressions. In particular, I focused on how to delimit true discontinuity, excluding certain subtypes of P2 effects (i.e. pronominal insertion into P2), and excluding cases of apparent discontinuity. Based on the general overview of non-configurationality and the role of nominal expressions, I compared the syntax of Vedic Sanskrit and Hindi. Vedic Sanskrit shows non-configurationality to a far-reaching degree, even greater than Warlpiri in certain respects. By contrast, Hindi, in the same way as other New Indo-Aryan languages, has developed postpositional phrases, but still allows for free constituent order and null arguments, so that it is best characterized as being low-level configurational. Turning to the general theme of approaching configurationality as a historical phenomenon, I outlined prior proposals on the rise of phrasal structure in non-configurational languages, all of which involve an origin in symmetrical constructions. Finally, I fleshed out in greater detail the overarching hypothesis of this study, that the grammaticalization of postposition entails the rise of postpositional phrases in Indo-Aryan. The following three principle claims will be substantiated in the course of this study: first, the individual trajectories of the different postpositions bring about the modern phrasal constructions, with certain effects of paradigm formation in recent centuries; second, the New Indo-Aryan postpositional phrases arise from asymmetric groups, consisting of head and dependent; and third, semantic changes are the ones to entail the formation of syntactic phrases.





## The diverse origins of the Hindi simple postpositions

In the following sections, I present and evaluate the etymologies that have been proposed for the simple postpositions of Hindi.<sup>1</sup> These etymologies guide the way to the historical evidence of the emergence of the first phrasal structures. While several details remain unclear, it is possible to reconstruct at least the source lexeme and in part the morphological form for some of the postpositions. However, there is as yet no agreement on several of the etymologies in the literature and I will therefore spend some time arguing for certain trajectories. Accordingly, the next few pages contain a certain amount of philological detail. The essence of the following sections can be gathered from the short summary in section 3.7. Before I take up the individual postpositions in turn, I want to point out some general factors in the reconstruction of their etymologies.

As is well known, function words do not abide by general sound changes as their usage and frequency differs from that of ordinary lexemes which impacts on their phonetic development. So, while there are in-depth studies of sound change in Indo-Aryan (see Masica 1991, ch. 7 for an overview), idiosyncratic trajectories like those of the postpositions continue to defy even the specialists. Turner summarizes the situation as follows:

[The MIA and NIA] postpositions are the descendants of full words, and although they have all come into use as such in the comparatively well-documented period between Sanskrit and the modern languages, their phonetic change has been so different from that of normal words, and so intensive, that it is in many cases difficult or impossible to do more than guess at their origin. (Turner 1960: 23)

In addition to the limits of general sound changes when it comes to function words, texts written in Middle Indo-Aryan varieties often display a great amount of

<sup>1</sup> Hans Henrich Hock has pointed out to me a certain variation regarding the Hindi complex postposition *ke pās* 'near'. In older or non-standard varieties, this form would also connect directly to the oblique form without the possessive linker, and could therefore also be discussed here.

variation that is in part due to non-standardized orthography (some editions are already partially cleared of variant forms) but also due to the usage of forms from a host of varieties. The texts in Apabhramsha and Old Awadhi in particular include many Sanskritisms and Prakritisms, and it is not always clear whether or not such similar forms should be subsumed as variants of a single lexeme, a difficulty which also pertains to some of the elements to be discussed in this chapter. Nevertheless, several scholars since the nineteenth century have made attempts at reconstruction and some of these are today generally endorsed by the academic community. In the following, I will present and evaluate the reconstructions of the postpositions, roughly in the order in which they appear historically in usages approaching case marking, starting with *mē*, followed by *par*, *se*, *ko*, *ne*, and separately at the end adnominal *kā/-e/-ī*.

### 3.1 *Mē* ‘in’

Nowadays, there is widespread consensus to derive *mē* ‘in’ from *madhye*, locative singular of the Old Indo-Aryan lexeme *madhya-* ‘middle’ (e.g. Hoernle 1880: 241; Beames 1970 [1875]: 292ff; Kellogg 1972 [1875]: 132; Oberlies 1998: 12), a cognate of Greek μέσος, Latin *medius*, and English *middle* (Pokorny 1959: 706–7). Another meaning attested for *madhya-* in Old Indo-Aryan and Pali is ‘waist’, which is also found for some of its cognates in other Indo-European languages and is reconstructed already for PIE (Pokorny 1959: 706, Wodtke et al. 2008: 465). Whether *madhya-* started off as a body part term and later developed into a spatial term, or whether it was metaphorically used for ‘waist’, in both usages it represents a prototypical source of an adposition (cp. Kahr 1975, 1976).

There is one particular aspect which has posed problems for the reconstruction of *mē*, and that is its word-final nasalization which is difficult to derive in a straightforward manner from *madhye*. For this reason, some authors discuss whether *mē* instead goes back to accusative *madhyam* (homonymous with the nominative form) which has a final nasal that could have triggered the development (Kellogg 1972 [1875]: 132–3). As much of this study is based on a detailed analysis of the syntactic history especially of *mē* and *par*, it is vital to ensure that we are dealing with the correct historical predecessors leading the way to the relevant historical data. In what follows, therefore, arguments for *madhye* or *madhyam* as the correct precursor will be discussed.<sup>2</sup>

Semantically, both *madhye* (LOC) ‘in the middle’ and *madhyam* (ACC) ‘(in)to the middle’ seem logically possible sources at first glance, as *mē* today expresses both

<sup>2</sup> I will not discuss proposals that have obvious problems, among which is Weber’s (1877: 519) suggestion to derive *mē* from either Persian-Arabic *min* or the Skt. pronominal locative singular affix *-smin*. The reader is referred to Kellogg for a critique (1972 [1875]: 133).

inessive as well as directional meaning. However, while the Old Indo-Aryan locative may mark not only locations but also goals when construed with verbs of movement (usually when an intrusion into a goal is expressed), the accusative when used spatially typically only marks goals and not locations (see Kellogg 1972 [1875]: 418; Delbrück 2009 [1888]: 121). *Mē* thus shows the semantic range of an old locative rather than of an accusative. Compare the following examples:

- (51) HINDI  
*maī śāhar mē rahtī hū*  
 1SG city.OBL.SG in live.IPFV.SG.F be.1SG  
 'I live in the city.'
- (52) HINDI  
*maī śāhar jā rahī hū*  
 1SG city.OBL.SG go PROG.F be.1SG  
 'I am going to the city.'
- (53) HINDI  
*maī śāhar mē jā rahī hū*  
 1SG city.OBL.SG in go PROG.F be.1SG  
 'I am going into the city.'

These examples illustrate a default inessive interpretation of *mē* in (51) and the default construction encoding goals involving a bare oblique in (52). With a verb of movement, however, a nominal expression marked by *mē* can also be construed as a goal if the verb adds the necessary semantic component of directed movement as in (53). From a semantic point of view, then, *mē* continues the semantic range of a locative and not of an accusative.

There is corroborative quantitative evidence from varieties of Middle Indo-Aryan where forms deriving from *madhye* by far outnumber forms deriving from *madhyam*. A word count in the Pali *Tipiṭaka* (canonical texts only) yields a total of 1,836 instances of *majjhe* (from *madhye*) in contrast with only 126 instances of *majjham* (from *madhyam*). In the literature on grammaticalization, it is well known that it is typically the lexemes which are already highly frequent that tend to grammaticalize (see e.g. Bybee et al. 1994: 5).

Semantic and quantitative evidence thus points to an origin of *mē* in *madhye*, but this does not yet account for the nasalization that we find today. Note, however, that 'spontaneous nasalizations' (Grierson 1922) are a phenomenon well known in New Indo-Aryan languages and particularly in Hindi: that is, nasalizations of vowels where a nasal that could have triggered the nasalization is missing in the precursor form. A well-known example is nasalized Hindi *sāp* 'snake' deriving from Skt. *sarpa*. In some cases, the nasalization arises from adjacency with preceding nasal occlusives. Bloch (1965: 48–9) discusses such cases, among other examples adducing cognates of

*mē* in different New Indo-Aryan languages. Note also that in several other New Indo-Aryan languages, cognates have remained un-nasalized (see Masica 1991: 249), which suggests that ‘spontaneous’ nasalization is a matter of dialectal variation. Furthermore, in the Middle Indo-Aryan and Old Awadhi data studied here, both nasalized and un-nasalized variants occur side by side in apparent free variation. I conclude that nasalization does not stand in the way of a derivation of *mē* from *madhye*, the source of which I henceforth assume to be correct on the basis of the semantic as well as quantitative evidence outlined.

### 3.2 *Par* ‘on’

The second postposition whose origin has been established with much certainty is *par* ‘on’. *Par* is generally taken to derive from Skt. *upari* ‘above, on top’ (e.g. Beames 1970 [1875]: 298; Kellogg 1972 [1875]: 133; Oberlies 1998: 11), related for instance to Latin *super*, English *over*, and German *über* (Pokorny 1959). Only Hoernle (1880: 241) suggests an origin in Skt. *pare*, locative singular of *para-* ‘far, distant, beyond’. This form does not figure in Middle Indo-Aryan varieties as a postposition, however, and the semantics do not fit very well. Over the stages of Middle Indo-Aryan, different variants of descendants of *upari* are attested, such as *uppari* and *uvari* as well as *pari*. In contrast to the precursor lexeme of *mē*, *upari* is an invariant form in Old Indo-Aryan so that there is no question of an alternative inflectional form as a possible source.<sup>3</sup> In sum, *par* in all likelihood descends from Old Indic *upari*, a reconstruction which I will assume to be correct for the remainder of this study.<sup>4</sup>

### 3.3 Instrumental-ablative *se*

At an early point in time, *se* was analysed as deriving from a Sanskrit case suffix, locative plural *-su* (this suggestion is due to Weber 1877: 519). However, it is

<sup>3</sup> Note that not all Middle Indo-Aryan versions of *UPARI*, e.g. *uvari*, make for an obvious phonetic predecessor of *par*. However, these forms were used in free variation with forms that can be phonetic predecessors without a discernible syntactic difference, as will become clear in the course of this study. Therefore, I consider all variants of *UPARI* as syntactic—if not necessarily phonetic—precursor forms of Hindi *par*.

<sup>4</sup> Note that *UPARI* is assumed to be a derivation of the Old Indic local particle *upa* ‘near, to’. In Chapter 4, I will argue that the modern postpositions have no categorial connection to the old local particles, contrary to what has been claimed. On an etymological level, only *UPARI* is at least a candidate for an etymological connection. However, since formations in *-r* are also attested in other ancient varieties of Indo-European such as Latin and Ancient Greek, the derivation must date to a very early time. Also, Old Indic *upari* and *upa* behave differently both semantically (‘above’ vs. ‘near, to’) as well as syntactically, since only *upa* can function as a preverb. Synchronically from Old Indic onwards, there is no connection between the two elements, so that this does not posit a difficulty for my account in Chapter 4.

controversial whether degrammaticalizations of this kind, where an inflectional marker becomes an independent morpheme, occur in languages at all. If they should do so, they are very rare. What is more, the semantics of locative plural marking differ too much from the instrumental-ablative function of *se* for this suggestion to be seriously considered. *Se* was linked by others to the locative form *saṅge* from *saṅga-* ‘contact, relation,’ which can be construed as being at the origin of the instrumental and possibly also ablative usages of *se* (e.g. Kellogg 1972 [1875]: 132). Beames, by contrast, makes a case for the Old Indo-Aryan adverb *samaṃ* ‘together’, which also fits in at least with the instrumental semantics of *se*, a suggestion adopted by Oberlies (1998: 21). Besides quoting *saṅge*, Turner (1966) proposes an origin in Old Indo-Aryan *sahita-* ‘accompanied’, formed from *sam* ‘together’ and the perfective participle of  $\sqrt{dhā}$  ‘put’. What is not usually listed is a possible origin in Skt. *saha* which means ‘with’ in a comitative sense. Comitative meaning is not expressed by Hindi *se* (instead by the complex postposition *ke sāth*), but in Apabhramsha, where *saha* (as *sahu*) has also acquired instrumental meaning (cp. Bubeník 1998: 83), a cross-linguistically well-known path of semantic development, and therefore also a candidate.<sup>5</sup>

Strnad (2013: 355) states that most authors now assume an origin in *samaṃ*. However, it appears not to have been clarified how to derive from this form the usage of designating a source. It may be more promising to consider the merger of two forms in order to account for the origins of the various functions of Hindi *se*. In the various late Middle and early New Indo-Aryan stages attested, there is generally a host of forms found that phonologically resemble *se* and one or the other of the above-mentioned potential ultimate source forms in Old Indo-Aryan, but it is difficult to point to an uncontroversial etymology. Note that each of the proposals above refers to an element of a different word class. Whereas *saṅge* would be another locative noun with a spatial reading in parallel with *madhye* (and possibly *kākṣe*, see the next section), *samaṃ* and *saha* resemble *upari* in their adverbial character. *Sahita-*, in turn, as a perfective participle would potentially parallel the case of *kā/-e/-ī* (see section 3.6). In sum, then, it has not yet been possible to settle the origin of *se*. There is a small amount of syntactic evidence that pertains to potential forerunners of *se* in Apabhramsha, which will be touched on in section 8.3.1.

<sup>5</sup> There is a further candidate, namely early Middle Indo-Aryan *saddhiṃ*, according to Davids and Stede (2007 [1921–25]) from Old Indic *sadhrīm* ‘towards one goal’, which displays many of the functions of Hindi *se*. Its basic meaning ‘together with’, i.e. comitativeness, is not expressed by *se*, comitative function having been taken over by the secondary postposition *ke sāth* in Hindi. *Saddhiṃ*, however, shows several of the other usages that *se* has today, such as marking interlocutors, the opposite person in fights or in expressions of friendships, as well as standards of comparison or similarity (see Fahs 1989: 106–7). Furthermore, *saddhiṃ* follows a trajectory which closely parallels that of *upari*, the forerunner form of *par*, as will be outlined in Chapter 5.

### 3.4 Dative-accusative *ko*

The postposition *ko* is nowadays generally derived from Skt. *kakṣe* (armpit.LOC.SG.M) ‘in the armpit’. The initial suggestion of linking *ko* to Skt. *kṛta* (perfective participle of *√kr* ‘to make’) by Trumpp (1872: 115), accepted by the early Kellogg (cited in Kellogg 1972 [1875]: 130), was soon rejected. Rather, the latter is a good candidate for *kā/-e/-ī* (see section 3.6). The original suggestion of *kakṣe* is due to Rudolf Hoernle (1880: 224–7) and was taken up by John Beames (1970 [1875]: 252–62), who offers a detailed outline of the necessary phonetic changes, which has since been endorsed by many scholars (Kellogg 1972 [1875]: 130; Oberlies 1998: 11; Butt and Ahmed 2011: 564–5). However, ‘armpit’ is not among the perceptually salient body parts such as ‘head’, ‘back’, or ‘belly’, which are commonly observed to grammaticalize. An extended meaning of *kakṣa-* seems to have been the part below the armpit (Mayrhofer 1992), or ‘side (of the body), flank’, which may have been closer to the starting point for the grammaticalization.

These meanings of *kakṣa-* closely resemble the origin of the Hindi complex postposition *ke pās* ‘near’, with *pās* deriving from Sanskrit *pārśve* ‘at the ribs, flank of the body’. Recall that *madhya-* in a similar way has a secondary meaning of ‘waist’. Consequently, the grammaticalization paths of *mē* and *ko* (as well as *pās*) from body part terms would strongly resemble each other. In contrast with *madhye*, however, *ko* is used in a way approaching case marking at a much later point in time than *madhye*, which is considered a periphrastic locative construction at least since Pali (cp. Fahs 1989: 103).

The necessary transitory stages from *kakṣe* to *ko* are not well attested, but there is a small amount of evidence. Chatterji (1970 [1926]: 760–1) outlines that in Old Bengali, we find the form *kakhu*, as also *kāhū/kāhu* in Eastern Hindi (neighbouring Bengali), while Western Hindi (the main source of modern standard Hindi) has old forms such as *kahū*. In the twelfth and thirteenth centuries, the forms *kō*, *kū*, *ko* are attested in texts written in varieties of Old Urdu/Punjabi primarily marking indirect objects as well as direct objects when they are animate and/or definite, prefiguring the usage of modern *ko* (examples adapted from Butt and Ahmed 2011: 565):

(54) INDIRECT OBJECT

<i>farid</i>	<i>mē</i>	<i>janya</i>	<i>dukh</i>	<i>mōjh</i>	<i>ko</i>
Farid	1SG	know	grief/pain	OBL.1SG	DAT

‘Farid, I know I have grief (lit. grief is to/at me).’

(55) DIRECT OBJECT

<i>dhundēn</i>	<i>dīye</i>	<i>suhag</i>	<i>kū</i>
seek	give	husband	ACC

‘(You) are seeking a husband.’

These examples attest to the fact that the grammaticalization from spatial noun to abstract gram must have occurred prior to the emergence of the New Indo-Aryan varieties.<sup>6</sup> Now, how can we bridge the semantic gap between the original meaning of *kakṣe* and the grammaticalized usages of *ko*? Butt and Ahmed identify examples from approximately 1800 which are evidence for an intermediate stage of grammaticalization with *ko* encoding goals. A conceivable trajectory is a chronological shift from ‘in the armpit, at the side, flank’ via ‘to, towards’ encoding goals, and then on to recipients and patients. Compare the following examples by the author Dehalvi from around 1800, marking goals, both concrete and abstract (adapted from Butt and Ahmed 2011: 565–6):

- (56) *is manzil ko kab poācoge*  
 this destination.OBL.SG ACC when reach.FUT.2PL  
 ‘When will (you) reach this destination?’
- (57) *apne haq ko poāc kar*  
 self right ACC reach having  
 ‘having attained one’s right’

In sum, while the reconstruction for *ko* is not as well substantiated as for *mē* and *par*, it is probable that the correct etymology has been identified, as scholars agree on the phonetic derivation of *ko* from *kákṣe*, and because there is a certain amount of evidence which helps trace the semantic trajectory. The earliest attested usages in grammaticalized functions are from the twelfth and thirteenth centuries as illustrated by examples (54) and (55) for emerging Hindi/Urdu, with similar attestations from Old Bengali.

### 3.5 Ergative *ne*

The most contested etymology is that of the ergative marker *ne*. Trumpp (1872: 113) assumed that *ne* derives from the Skt. instrumental singular form *-ena*. However, starting with Beames (1970 [1875]: 266f), numerous scholars have pointed out that *-ena* was lost a long time before the first appearance of *ne* (‘over ten centuries’ in the view of Verbeke and de Cuyper 2009: 7, the exact time of the appearance of *ne*, however, remaining subject to debate: see the end of this section). The logically necessary sound changes do not conform to what is otherwise known about the history of Indo-Aryan either. As with the similar proposal to derive *se* from a suffix, the reinterpretation of a bound morpheme as a free grammatical morpheme would

<sup>6</sup> See also Old Awadhi data with these and other phonological variants and equivalent functions, Saksena (1971: 215–22).

need to be substantiated by strong evidence in order to be convincing (see Butt 2001: 113–16 or Verbeke and De Cuypere 2009: 6–8 for recent overviews).

While a reconstruction from *-ena* is hardly persuasive, *ne* was connected to the Skt. instrumental form not only on account of a vague phonological likeness, but for the good syntactic reason that the instrumental in former times encoded the agent in transitive constructions involving perfective ‘participles’, that is, the so-called *ta*-forms encoding resultative meaning. This construction, for example Sanskrit *Rāmeṇa kṛtam* (Rama.INS.SG.M do.PPP.NOM.SG.N) ‘it was done by Rama’ is commonly acknowledged as the source of today’s ergative construction. Thus, *ne* descending from an instrumental marker would present a continuation of instrumental-marking on the ergative argument. Below, I go into detail regarding the role that the old ergative construction seems indeed to have played. First, however, I want to briefly mention a few other etymologies that have been proposed.

Various lexical sources have been suggested, such as forms deriving from a Skt. root *√lag* ‘to stick’ (for details of proposed phonetic derivations, see Beames 1970 [1875]: 260ff, also Kellogg 1972 [1875]: 131–2), from Prakrit/Apabhramsha *taṇaya* ‘own’ (by Schwarzschild 1954: 93), or Skt. *janye* ‘for the sake of’ (by Aditi Lahiri as cited in Butt 2001: 116). So far, however, none of these proposals has been successfully substantiated with data for the necessary transitory stages.

The most promising suggestions involve borrowing.<sup>7</sup> *Ne* (and very similar forms) is employed as a case marker in several dialects and languages adjacent to or overlapping with the Hindi speaking area. Beames observed that *ne* was used in Persianized dialects in the environment of the British administration in the sixteenth and seventeenth centuries as a dative marker, which—and here the ergative function enters the picture—‘began gradually to be extended to the noun when used as the subject of a transitive verb in the past tense’ (1970 [1875]: 270, see also Hoernle 1880: 219; Grierson 1903: 484–5).<sup>8</sup> The ‘dative-ergative connection’ (Butt 2006) is exemplified by Old Rajasthani/Old Gujarati<sup>9</sup> and Haryani, all direct neighbours of Hindi, where *ne* (or similar forms) encodes both the dative and ergative (cp. Kellogg 1972 [1875]: 130–1; Butt and Ahmed 2011). Haryani, for instance, was used ‘in the environs of Delhi, which was one of the major residences of the Moghul court and which in turn was where the majority of Old Urdu was spoken and written in the Middle Ages’ (Butt and Ahmed 2011: 561). In addition to the usages mentioned above, a further function found in Haryani is that of locative meaning.

<sup>7</sup> While grammatical formatives are not easily borrowed, this is certainly not impossible if only the necessary sociolinguistic preconditions exist: see Thomason and Kaufman (1988).

<sup>8</sup> Beames (1970 [1875]: 272) furthermore adduces evidence from Marathi where inflectional forms were variously interpreted as either instrumental or dative, as well as from Punjabi where constructions with the closely resembling *nai* display a similar functional oscillation.

<sup>9</sup> Butt and Ahmed (2011: 561) refer to ‘Rajasthani’ and ‘Gujarati’ without specifying the precise historical period. The modern languages only employ *ne* for dative marking (e.g. Gusain 2004).



The ultimate source of the morpheme from which it spread to Hindi and other languages, finally, is identified as Rajasthani<sup>10</sup> in several studies (Butt and Ahmed 2011 and Montaut 2006, 2009 following Tessitori, as well as Verbeke and de Cuypere 2009). The most detailed version of this argument is Tessitori (1913) who depicts a typical phenomenon of deletion of initial syllables with *k-* and other consonants in Old Rajasthani. Tessitori connects *ne* with Apabhramsha *kaṇṇahī* from *karṇe*, the locative singular of Skt. *karṇa-* ‘ear’. In Old Rajasthani, variants of this form could mean ‘aside, near’. Furthermore, similarly to Haryani, it was used not only in locative function, but also as a dative and ablative marker, and a single example in ergative (‘agentive’) function is documented as well. Compare Tessitori’s (1913: 557–8) examples for the various functions with phonological variants retaining or dropping the initial *k-* including *kaṇi*, *kanhi*, and *naī* in the following examples:

- (58) OLD RAJASTHANI, LOCATIVE FUNCTION  
*na jāṇū kihā-kaṇi achai*  
 NEG know.1SG where-‘kaṇi’ be.3SG  
 ‘I do not know where (he) is.’
- (59) OLD RAJASTHANI, ABLATIVE<sup>11</sup> FUNCTION  
*Caturaka-kanhi pūchai vana-dhaṇi [...]*  
 Caturaka-‘kanhi’ ask.3SG forest-king.NOM.SG  
 ‘The king of the forest asks Caturaka [...].’
- (60) OLD RAJASTHANI, DATIVE FUNCTION<sup>12</sup>  
*āvai tihā-kaṇi*  
 go.3SG there-‘kaṇi’  
 ‘(He) goes there.’
- (61) OLD RAJASTHANI, ERGATIVE (‘AGENTIVE’) FUNCTION  
*Ādiśvara-naī dīkṣā līdhī jāṇī*  
 Ādiśvara-‘naī’ dīkṣā take know  
 ‘having learned that Ādiśvara had taken the dīkṣā’

The apparent functional gap between ergative and dative is reconciled by Tessitori who takes the ablative function as intermediate. He proposes that an original locative meaning gave rise to dative function on the one hand and to an ablative one on the other. The latter, in turn, can be construed as a possible source for the ergative function. Butt (2001: 131–29) argues that ergatives and datives have in common that

<sup>10</sup> ‘Rajasthani’ serves as an umbrella term for several varieties spoken in Rajasthan (e.g. Marwari or Jaipuri).

<sup>11</sup> The thematic role may be labelled ‘addressee’ instead of Tessitori’s ‘ablative’. However, the gram does not generally mark addressees, but only when something is asked, requested, or demanded from someone.

<sup>12</sup> This example adduced by Tessitori appears to involve a goal rather than a recipient or experiencer.

they serve as non-nominative subject markers which may explain their semantic connection without reference to an additional ablative usage.<sup>13</sup>

From a viewpoint of grammaticalization, Tessitori's etymology from ultimately *karṇe* (ear.LOC) is intriguing, with the semantics of 'aside, near' fitting in well with semantic changes that have been described, and the phonological transitory variants in Old Rajasthani furnishing the necessary intermediate forms between full lexeme and reduced gram.<sup>14</sup> However, the argument for Rajasthani as the donor language has the weakness that there is only the single example (61) identified by Tessitori with 'agentive' function, and that modern Rajasthani does not employ *ne* for ergative marking (e.g. Gusain 2004 for the Marwari variety of Rajasthani). Hence, actual ergative usage may only have developed after having been borrowed into other languages. Here, Haryani as a direct neighbour of Hindi seems a promising candidate.<sup>15</sup>

Regarding its entrance into the language, several authors have argued that *ne* somehow 'filled a gap' (Verbeke and De Cuypere 2009: 18) in the sense that the ergative argument was already in place, encoded by the bare oblique and continuing the functions of the old instrumental in constructions with perfective participles. *Ne*, it is argued, was inserted into this system as a secondary element 'reinforcing' the oblique when marking ergatives (e.g. Verbeke and De Cuypere 2009: 4). Compare the following example from the poet Kabir (1440–1518):

- (62) *masi kāgad chūyo nahī kalam gahī*  
 ink.NOM.SG paper.NOM.SG.M touch.PRF.SG.M NEG pen.NOM.SG.F take.PRF.SG.F  
*nahī hath ... kabir*  
 NEG hand Kabir.OBL.SG.M  
 'Kabir touched not ink nor paper, **he took no pen in hand.**'  
 (adapted from Butt and Ahmed 2011: 559)

In this example, while the agreement of *chūyo* does not give away the ergative pattern as both arguments are masculine singular, *gahī* agrees with *kalam* instead of with the agent *kabir*, showing ergative agreement. Many similar examples can be found somewhat later in the sixteenth century. Old Awadhi, for instance, one of the forerunner varieties of modern Hindi, did not know the postposition *ne* and employed the oblique for marking the ergative in constructions with perfective participles. Works by poets such as Jayasi and Tulsidas abound with examples such as *rānī sunā* (queen.OBL.SG.F hear.PPP.NOM.SG.M) 'the queen heard [sth.]' (Saksena

<sup>13</sup> Further support adduced by Butt and Ahmed (2011: 569) for a historical connection of dative markers evolving into ergative markers comes from Nepali where the postposition *le* at first only marks recipients but later develops ergative function.

<sup>14</sup> The *World Lexicon of Grammaticalization* (Heine and Kuteva 2002) also cites languages where words for 'ear' are sources for locatives, e.g. in Finnish.

<sup>15</sup> Hans Henrick Hock has pointed out to me that Braj varieties, too, should be taken into account.

1971: 241), where the participle agrees with a masculine direct object. This type of construction was commonly used until *ne* appears in addition to the oblique.

In sum, out of the numerous proposals, a borrowing sequence where *ne* was introduced from a neighbouring language or dialect, possibly Haryani with an ultimate origin in Rajasthani, is the scenario best supported by the evidence. For the remainder of this study, therefore, I will assume that *ne* was indeed borrowed, but I will not base any parts of my argument on the details of the borrowing scenario that has been outlined, as more research seems to be required to substantiate the proposal. As for its time of origin, *ne* appears very late in the history of the language. Butt and Ahmed (2011: 558) at one point mention 1400 as its time of appearance, but also give Beames' date of around 1600. Deo, as reported by Butt, identifies usages of *ne* in texts by the poet Amir Khusro who lived from 1253 to 1325, a finding which also requires further examination (see Butt 2001: 115, fn. 11). If Tessitori's account is correct, that *ne* arises from Old Rajasthani (ca. twelfth–fifteenth century), example (61) would constitute one of the earliest attestations of ergative usage (see also section 8.3.2.1 on *ne*).

### 3.6 Possessive *kā/-e/-ī*

This final section is devoted to the possessive marker *kā/-e/-ī*. A feature which sets it apart from the other postpositions is its partial inflection for case, number, and gender of the possessee. In the literature, this is commonly attributed to an origin in a participle agreeing with a noun that furnishes what today is the possessee. The suggestions all involve some participial formation of the root Skt.  $\sqrt{kr}$  'do'. Typically, authors either suggest the perfective participle *kr̥ta-* 'done' (e.g. Hoernle 1880: 234) or its enlarged variant *kr̥taka-*, *-ka-* being a frequently used derivational suffix for adjectives (accepted by Kellogg 1972 [1875]: 129 and Oberlies 1998: 11), or the future participle *kārya-* 'to be done' (e.g. Bubeník 1998: 76), also optionally with a suffix *-ka-*.

Whether *kr̥ta(ka-)* or *kārya(ka-)* is the source has not been settled. This is primarily because late Middle and early New Indo-Aryan varieties employ a host of phonological and spelling variants and it cannot be ascertained which forms exactly would have served as the predecessors of the modern forms. The semantics, however, strongly point to *kr̥ta-* and, for phonetic reasons, in all probability the extended variant *kr̥taka-*: it seems difficult to construct a path of change for a future participle 'to be done (in the future)' to evolve into a marker of possession. Something which has not happened does not metaphorically belong to someone yet, as the connection between the entity and the activity has not yet been established.<sup>16</sup> On the

<sup>16</sup> Grierson (1903: 485–6) attempts to bridge this semantic gap. His argument rests on the existence of a nominal usage of *kārya-* as 'matter, thing, occurrence' which, as he argues, involves the construal of a past event.

other hand, something which has been done or made by someone seems to be more easily construable as standing in a relation of possession (concrete or abstract) to someone.

In conclusion, while the historical reconstruction of *kā/-e/-ī* awaits further examination, scholars agree that it descends from a participial form of  $\sqrt{kṛ}$  'do'. Furthermore, an evaluation of its semantic structure points to a likely origin in *kṛta-*, an expansion of the perfective participle *kṛta-*. Regarding its time of appearance, it is from Apabhramsha onwards that forms going back to some participial form of  $\sqrt{kṛ}$  systematically start encoding relations of possession.

### 3.7 Summary

Drawing on the evidence that was presented in the preceding sections, it is clear that the Hindi simple postpositions by no means either descend from the same type of morphosyntactic source or grammaticalized at a similar point in time. Instead, the postpositions derive from very different word classes and historical periods. Regarding the morphosyntactic origins, *mē* (with certainty) and *ko* (in all likelihood) can be reconstructed as deriving from spatial or body part terms. *Par* derives from an invariant adverb *upari*. *Kā/-e/-ī* descends from a participial form of *kṛ* 'do', in all likelihood from *kṛta-* in the extended form *kṛta-*, the perfective participle form 'done, made'. The etymology of *se* remains uncertain with a nominal (*sañge*), adverbial (*samam*), or participial (*sahita-*) origin, besides other suggestions. Last but not least, *ne* seems most likely to have been borrowed from a neighbouring language at the dawn of Hindi/Urdu.

These etymologies are the starting point for the historical corpus study in this book. Since *mē* and *par* can be traced back to Old Indic with a high degree of certainty, I concentrate on them in studying the earliest developments of phrasal structures over the next chapters. With regard to the other elements, their developments at least since late Middle Indic or early New Indic can be reconstructed based on additional syntactic evidence that will be reviewed in Chapter 8. This will enable me to adopt a holistic view on the development of these elements of heterogeneous sources into a single formal and functional paradigm in the course of the last millennium.

## Local particles

### *The unique source of adpositions and configurationality in Indo-European?*

The topic of this book is the rise of low-level configurationality in Indo-Aryan, which hinges on the rise of the first phrasal structures in this branch of languages. The rise of phrase structure, in turn, I regard as a concomitant of the grammaticalization of certain elements. I identify these elements as the modern postpositions of Hindi and other New Indo-Aryan languages. In order to understand the origins of phrase structure, then, it is necessary to study the syntactic and semantic origins of these elements. However, should the modern postpositions merely be young members in an old category, one should instead look at this older category and its original members in search of the origins of phrase structure. In the literature, a scenario is assumed by many authors which does indeed link the postpositions back to such an old category. It is claimed that (i) simple adpositions of contemporary Indo-European languages (as well as related preverbs) etymologically go back to a group of old PIE spatial adverbs, the so-called local particles, and that, accordingly, (ii) it is the reanalysis of these local particles as adpositions which concomitantly gives rise to the first phrasal structure—adpositional phrases—in the various Indo-European branches. This proposal is often made in a general manner for the language family as a whole, but it has also been directed explicitly at Indo-Aryan. While the scenario correctly describes the developments in most branches of Indo-European, I argue in this chapter that it must be rejected for Indo-Aryan. Instead, the origins of phrase structure in Indo-Aryan lie elsewhere, namely in the precursor forms of the modern postpositions, which have no connection with the local particles. Before I begin with my argument, note that some authors have not drawn this link between local particles and New Indo-Aryan postpositions. For instance Bloch (1965) and Andersen (1979) refer to the origins of the postpositional phrases solely by taking recourse to their source constructions, and do not link them to the local particles.

This chapter is structured as follows. In section 4.1, I outline the scenario generally assumed, that modern Indo-European adpositions, including New Indo-Aryan

postpositions, continue the category of the old Indo-European local particles. In section 4.2, I evaluate the syntax and semantics of the local particles as attested in Old and early Middle Indic in order to explore whether they could in fact be the ancestors of the postpositions in this branch. First, I turn to evidence from the *Rigveda*, drawing on research by Heinrich Hettrich and colleagues (section 4.2.1). Then, I study the small handful of local particles which survive into post-Rigvedic times and on into Middle Indic as free morphemes with nominal orientation (i.e. in functions resembling adpositions). It is these residual elements which overlap in time with elements like *MADHYE* and *UPARI* entering the periphery of local case marking in early Middle Indo-Aryan, and which might therefore represent the link between the old and the new category (section 4.2.2). However, besides their very rare occurrence, I will unveil a preference for *prenominal* placement in contrast with the preference for the postnominal position shown by the precursor forms of the modern postpositions (see Chapter 6). Since there is hardly a temporal overlap and the syntactic properties do not match, I conclude that there is no connection between the two categories. Additional evidence that the modern postpositional constructions cannot be connected to the local particles comes from the constructional environments where they occur, which will be outlined in the next chapter. In view of the importance accorded to the local particles regarding the rise of phrase structure in Indo-European, I will address the question of why the local particles did not in fact survive in Indo-Aryan as adpositions, in contrast with the rest of Indo-European (section 4.3). I propose a scenario which takes into account the prosodic interplay of finite verb forms and local particles in dependency of position and clause type, showing that the interplay specific to Vedic Sanskrit was unfavourable to a reanalysis of the local particles as adpositions.

#### 4.1 The traditional scenario

I start with a brief introduction to the local particles—more detail will be given in section 4.2. These adverbial elements often specify the semantics of verbal or nominal expressions. Compare the following Homeric Greek examples showing the syntagmatic and semantic versatility of the local particle *épi* ‘on, to, near’ (from Hewson 2006: 4–7; see also Bortone 2010 on Ancient Greek and Wilhelm 2001 on Indo-European):

- (63) *éluth’ épi psukhè Agamémnonos*  
 come.AOR.3SG EPI soul.NOM Agamemnon.GEN  
 ‘The soul of Agamemnon approached.’ (Odyssey 24.20)
- (64) *epi gaían ap’ ouranóthen protrápētai*  
 EPI earth.ACC from heaven turn.3SG  
 ‘To earth from heaven (the sun) turns.’ (Odyssey 11.18)

- (65) *aletréuousi mülēys épi mēlopa karpón*  
 grind.3PL millstone.DAT EPI yellow.ACC grain.ACC  
 ‘They grind the yellow grain on the millstone.’ (Odyssey 7.104)
- (66) *kai epì knéphas hieròn élthēy*  
 and EPI darkness sacred come.3SG  
 ‘and the sacred darkness closes in’ (Iliad 11.209)
- (67) *kephalēy d’ epéthēke kalúptrēn*  
 head.DAT PRT EPI\_place.AOR.3SG veil.ACC  
 ‘and upon her head (she) put a veil’ (Odyssey 10.545)

This syntagmatic flexibility, which does not clearly map on usage, shows that we are still dealing with a single category. However, there are signs of a correlation of placement and meaning. Whereas *epi* appears to be used adverbially in example (63), its usage in examples (64) and (65) appears to foreshadow adpositional function, but shows variable placement with respect to the local case form, while it can be interpreted as modifying the verbal semantics of examples (66) and (67). Due to such different usages, the local particles in the ancient languages used to be referred to as adpositions or preverbs. However, since no clear split into two categories can be observed yet in Vedic Sanskrit (or Homeric Greek), I will speak of the local particles being used with either semantic *verbal* or *nominal orientation*—a distinction not always easy to make.<sup>1</sup> These local particles are part of the grammar of all old Indo-European varieties and in most of the daughter languages split at some point into adpositions on the one hand, and preverbs on the other, many of them still in existence today (see Vincent 1999 for a proposal as to how this split came about in Latin). For instance, English *of*, *at*, *by*, *in*, *on*, *to*, or *for*, or French *à*, *de*, *en*, or *par* originated as local particles. In many cases, one and the same local particle has given rise to both an adposition and a preverb, for example French *de la fille* ‘of the girl’ vs. *devenir* ‘to become’, etc. In accordance with these facts, it is frequently asserted that simple adpositions in modern Indo-European languages derive *in general* from the Proto-Indo-European local particles (for an early example see Meillet and Vendryes 1927: 497). A recent instance of this proposal can be found in Hewson (2006): ‘The development of systems of adpositions out of the ancient preverbs [i.e. local particles, UR] ultimately led to revolutionary results in the Indo-European languages’ (Hewson 2006: 9; see also Luraghi 2010).

Several authors, both non-specialists and specialists in Indo-Aryan, specifically mention this branch as also attesting to this generally assumed scenario that the

<sup>1</sup> The term *local particle* from German *Lokalpartikeln* is adopted from the research programme on these elements by Heinrich Hettrich and colleagues. While the term ‘particle’ has elsewhere been criticized for its vagueness and justly so (cp. Zwicky 1985), it is this vagueness that makes it a suitable label for the polyfunctional elements in question.

origin of the postpositions is to be connected with the local particles. For instance, Robert Coleman comments on the development of word order properties of ‘adpositional particles’/‘Ptc’ (i.e. the local particles) in different branches of Indo-European, paralleling the development in several branches including Indo-Aryan:

The direction of change in IE has not been from N Ptc to Ptc N but (i) towards elimination of the choice between the two orders; (ii) from optional to obligatory adjacency of the two items, as in Ptc N in English, N Ptc in Hindi, in contrast to the possibility of separating them from each other in Old English and Sanskrit; (iii) towards a restriction of each adpositional particle to one case, as in Classical Latin beside Hellenistic Greek. (Coleman 1991: 323)

The same idea that the modern postpositions are in some sense related to the local particles is hinted at in publications by Chatterji and Emeneau, two of the most influential writers on Indo-Aryan. Both authors contemplate the divergence of Indo-Aryan having postpositions instead of prepositions<sup>2</sup> and invoke a connection with the local particles (‘prepositions’ in Chatterji’s terminology, ‘adverbs in immediate constituency with nouns’ in Emeneau’s), based on the assumption that they, too, favoured the postnominal slot:

The most noteworthy thing [about NIA languages] is the gradual disuse of prepositions. All other IE. languages developed the prepositions as aids to the declinational system; and when the inflections died out, the prepositions took up their place, as in English and Persian, French and Bulgarian.... In Primitive IE., the preposition, in origin an adverb, came before or after the noun; but it is remarkable that the development of it in India, where it is not entirely suppressed, should be post-positional (as in Sanskrit); that and in late MIA. and NIA., a series of help-words of a different kind, the post-positions of nominal and verbal origin, should come in. (Chatterji 1970 [1926]: 172, emphasis added)

The absence of prepositions is striking to an Indo-Europeanist or a speaker of a Western Indo-European language; it should be remembered, however, that in Sanskrit itself (and it inherits this trait from Proto-Indo-European) there is no class of ‘prepositions’—the morphemes in question are rather ‘adverbs in immediate constituency with nouns’, the position being postpositional probably rather more often than prepositional. If these are replaced in Modern Indo-Aryan by noun forms invariably following the oblique form of the head noun, the construction is not too different from that of Sanskrit. Parallel constructions in Dravidian may possibly have helped toward the shift. (Emeneau 1956: 9, emphasis added)

In recent times, Bubeník (2006a) in particular has outlined in detail the rise of postpositions in Indo-Aryan as moulded into this scenario. Bubeník begins his argument for the connection of the modern postpositions with the ancient local particles by describing the situation in Vedic Sanskrit, pointing out that the local

<sup>2</sup> Others, of course, consider the opposite case more bewildering, namely that OV languages, e.g. Latin, have prepositions instead of postpositions, contrary to the expected word order harmony, only later changing to VO.



particles occur in syntactic apposition (and typically actual juxtaposition) with local case forms and often ‘reinforce’ the latter on a semantic level. For instance, the string *urvyām ādhi pṛthivyām* (broad.LOC.SG.F above earth.LOC.SG.F) literally ‘above, at/on/in/by the earth’ is rendered by Bubeník as ‘on the earth’. Bubeník continues with the assertion that the connection between local particle and case form is strengthened when the morphological affixes on the nominal form erode. This is a version of the push chain scenario that has often been assumed to explain the loss of morphological case and rise of adpositions in Indo-European, as outlined in section 1.1. The loss of morphological case is the crucial shift in the historical process, which leads to reanalysis of the string as a syntactic unit:

With the subsequent reduction of morphological case, it becomes practically impossible to ‘disrupt’ the noun phrase and the postpositional phrase became fully established. In terms of grammaticalization, the adverbial particle/adposition [i.e. local particle, UR] is revamped as postposition... (Bubeník 2006a: 110).

Bubeník couches his proposal also, in his own words, ‘in more formal syntactic terms’ (2006a: 110). By way of example, he juxtaposes two phrase structure trees (Figure 4.1), one for Vedic Sanskrit and one for Hindi, for the semantically roughly corresponding strings of *urvyām ādhi pṛthivyām* (Vedic Sanskrit) and *cauṛī prithvī (= pṛthvī) par* (Hindi), both translated by Bubeník as ‘on the broad earth’.

Thus, the modern construction is depicted as the direct successor of the former, the only difference lying in word order becoming fixed (presumably indicated by the dashed and solid lines respectively), accompanied by a change in labels from ‘AdvP’ to ‘PoP’. Bubeník, as others before him, does not address the fact that the Hindi and other New Indo-Aryan postpositions have etymologies that are entirely unrelated.<sup>3</sup>

The claim that the simple adpositions of modern Indo-European languages in general—including the Indo-Aryan postpositions—descend from the local particles is

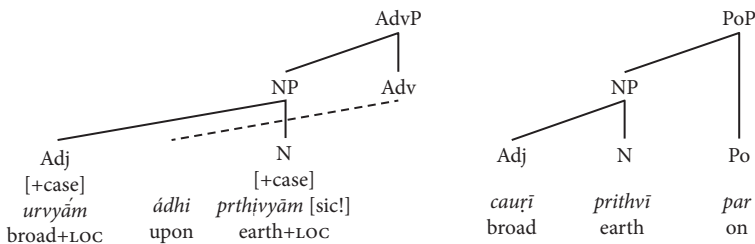


FIGURE 4.1 ‘on the broad earth’ in Vedic and Hindi

Source: Bubeník (2006a: 111). Reproduced with permission from John Benjamins.

<sup>3</sup> See Ch. 3, fn. 4 for Hindi *par*.

also the foundation for proposals that apply to the historical syntax of Indo-European more generally. The reinterpretation of local particles as adpositions is considered the initial step in the overhaul of the general syntactic types of these languages in that the newly created adpositions introduce phrasal organization which, depending on the language, also expands to higher clausal levels:

In the preceding sections I have shown how semantic constituency turned into syntactic constituency in the IE languages, starting from adpositional phrases, then spreading to NPs, and finally to the VP. (Luraghi 2010: 226)

The creation of the adpositional phrase as a new configuration was a model for similar configurational developments in the noun phrase and the verb phrase, and consequently the point of departure for the new typology. (Hewson 2006: 1)

Now, if the developments in a branch of Indo-European were to diverge from this scenario, this would also have implications for the general scenario of how Indo-European languages developed phrasal organization on higher syntactic levels.

Coming back to the data at hand, as there is no etymological connection, the claim of continuity between local particles and modern New Indo-Aryan postpositions is in need of support from syntax and semantics, as well as time. The argument for such a lineage would only be plausible if the Vedic local particles displayed (or acquired over time) such formal and functional properties as could prefigure those of the emerging postpositions. Also, coexistence of local particles and the emerging postpositions for a period of time would be a necessary condition.<sup>4</sup> In the following sections, I show that the evidence speaks against any such lineage, on the basis of data from the *Rigveda* as well as from post-Rigvedic stages of Old Indo-Aryan and from early Middle Indo-Aryan Pali.

#### 4.2 Syntax and semantics of the local particles in Old and Middle Indic

In order to ascertain whether the postpositions in some sense continue the category of local particles, one needs to consider the syntactic and semantic characteristics that the old elements would have had to display in order to be at all able to give rise to the modern elements and constructions. Some traits of the modern postpositional constructions simply could not yet have developed in Old Indo-Aryan, such as the oblique nominal morphology. Other properties of the modern constructions, however, would need to have been present in the earlier constructions. I claim that only elements which frequently occurred as (i) free morphemes in (ii) postnominal position with (iii) nominal orientation can be candidates as precursors of the modern

<sup>4</sup> I take it as uncontroversial that a category does not continue to exist when its members have all vanished from a language.

postpositions.<sup>5</sup> In the following two sections, I first turn to the syntax and semantics of the local particles in the *Rigveda* (section 4.2.1), the common reference point in the literature. I then move on in section 4.2.2 to the last survivors of the local particles, the only elements which still occur in the period of late Old and early Middle Indo-Aryan when MADHYE and UPARI begin to be used more frequently in periphrastic expressions of local cases.<sup>6</sup>

#### 4.2.1 Local particles in the Rigveda

It has been pointed out in section 4.1 that the local particles oscillate between functions approximating those of adverbs, preverbs, and adpositions. Some of this variety is illustrated in the following examples with *adhi*<sup>7</sup> ‘above’ (examples from Hettrich 1991):

(68) (= ex. (32))

<i>eṣá</i>	<i>súryeṇa</i>	<i>hāsate</i>	<i>pávamāno</i>	<i>ádhi</i>
DEM.NOM.SG.M	SUN.INS.SG.M	RACE.MID.3SG	PURIFY.PPM.NOM.SG.M	above
<i>dyávi</i>				
sky.LOC.SG.M				

‘This one, purifying himself, races with the sun, above, in the sky.’ (RV 9.27.5)

(69) *yádi stutásya maruto adhīthá*  
*yádi stutásya maruto adhi ithá*  
 if praise.GEN.SG.N Marut.VOC.PL.M above go.2PL  
 ‘if you, oh Maruts, are aware [lit. go above] of the praise’ (RV 7.56.15)

(70) *ádhi stotrásya sakhyásya gātana*  
 above praise.GEN.SG.N fellowship.GEN.SG.N go.IMP.2PL  
 ‘think of [lit. ‘come above’] the praise (and) the fellowship’ (RV 5.55.9)

(71) *ágne ní ṣatsi námasádhi barhísi*  
*ágne ní ṣatsi námasā ádhi barhísi*  
 Agni.VOC.SG.M down sit.IMP.2SG worship.INS.SG.N above barhis.LOC.SG.N  
 ‘Agni, sit down on the barhis [i.e. sacrificial grass] with worship.’  
 (RV 8.23.26)

<sup>5</sup> It is a logical possibility that the modern postpositions continue the old category of local particles and only came to ‘settle’ in the postpositional slot after the last local particles had died out. However, it will be shown in Chapter 6 that this word order preference is found from Old Indic times onwards.

<sup>6</sup> In this period, however, MADHYE and UPARI still clearly retain their concrete spatial meanings. Thus, they only occur in environments similar to local case forms when they actually denote concrete spatial meanings. The shift to synsemantic does not set in before late Middle Indic.

<sup>7</sup> I only note accentuation in quoted examples when present in the primary source, as I often make remarks that apply to local particles as attested not only in Vedic Sanskrit, or not only in those Vedic texts which note accentuation.

- (72) *síndhor*      *ūrmāv*      *ádhi*      *śrítāḥ*  
 river.GEN.SG.M    wave.LOC.SG.M    above    lean.PPP.NOM.SG.M  
 ‘(the wise one) leaning on the wave of the river’ (RV 9.14.1)

In example (68), *ádhi* does not seem to stand in a closer semantic relation with either the verb or a noun. Many scholars consider this adverbial function to be the original usage. In (69), by contrast, *ádhi* combines semantically with the verb  $\sqrt{i}$  ‘to go’, showing a lexicalized meaning ‘attend to, be aware’ instead of a literal interpretation of ‘go above’. This closer semantic connection is mirrored on the formal side, the local particle and the verb forming a prosodic unit, an effect of particular patterns of accentuation linked to clause type and position. Note, however, that prosodic fusion with the verb is not a necessity in cases of verbal orientation as illustrated in example (70) which contains an almost synonymous expression involving *ádhi* with *gā* ‘to come’. The next example (71) can be taken as displaying nominal orientation, the verb combining with the local particle *nī*, while *ádhi* combines with *barhiṣi*. In (72), finally, it is not certain whether we should interpret *ádhi* as construing with the noun or the verb as both interpretations, *síndhor ūrmāv (ádhi śrítāḥ)* and *(síndhor ūrmāv ádhi) śrítāḥ*, are possible options (cp. Hettrich et al. 2010: 20).

These examples give a brief impression of the breadth of usages of the local particles. They may occur in all kinds of positions in the clause with different types of semantic relations to other elements, verbal or nominal. Importantly, these examples clearly show that placement is not necessarily indicative of semantic orientation. What is more, some cases resist classification. Coming back to our original question, is it possible to discern a particular correlation involving (i) occurrences of the local particles as free morphemes in (ii) post-nominal position with (iii) nominal orientation, which could have given rise to the modern postpositions?

In a series of articles, Heinrich Hettrich, Antje Casaretto, and Carolin Schneider classify the attestations of the various local particles in the *Rigveda*. They register comparatively unambiguous mappings between placement and orientation and strictly classify only these:

Syntactic affiliation to the reference noun is indicated by pre- or postposed adjacent placement of the local particle and reference nouns ...; by contrast, affiliation with the verb is marked by directly preverbal position as the neutral option, and clause-initial position (tmesis)<sup>8</sup> as the marked option. (Hettrich et al. 2010: 20, translation UR)<sup>9</sup>

<sup>8</sup> *Tmesis* ‘separation’ refers to verbally oriented local particles removed from the finite verb form.

<sup>9</sup> ‘Syntaktische Zugehörigkeit zum BN [i.e. Bezugsnomen, UR] ist erkennbar an prä- oder postpositiver Kontaktstellung von LP [i.e. Lokalpartikel, UR] und BN ...; dagegen zeigt sich Zugehörigkeit zum Verb durch die unmittelbare präverbale Position als neutrale oder durch Satzanfangsstellung (Tmesis) als markierte Variante.’

Other cases, for example where a local particle stands between a case form and a verb, are excluded as ambiguous by default. Note that this is a syntagmatic strategy of classification, while I prefer referring to semantic verbal or nominal orientation as the functional breadth is still too large to speak of a syntactic split.

To my knowledge, this represents the first genuine attempt at categorizing the usages according to testable criteria which, however, presupposes the possibility of category assignment, the feasibility of which is not recognized by all authors. For example, Renou is very sceptical, saying that it 'is only an act of our mind faced with a linguistic structure which is fundamentally ambiguous' (1956: 119, translation UR<sup>10</sup>) which is why we try to classify the usages. However, the correlations between preverbal or clause-initial position with verbal orientation, on the one hand, and between juxtaposed position to a local case form and nominal orientation, on the other, are indeed typical. There may be some false positives, but the analyses in the publications by Hettrich and colleagues are convincing, especially considering that hundreds of attestations for each local particle are evaluated. Also note that theirs is a conservative approach. There are many examples where the position may allow for both options and which are therefore excluded as ambiguous, but where the semantic orientation is in fact reasonably clear.

When attempting a classification, Renou generally favours verbal orientation:<sup>11</sup> 'In principle, you should only admit a preposition [i.e. a local particle with nominal orientation] when the other interpretation [i.e. as verbally oriented] is excluded' (Renou 1956: 118, translation UR).<sup>12</sup> In comparison, the approach taken by Hettrich and associates is more favourable towards nominal orientation as they take adjacency with a local case form (if not standing in adjacency with the verb or in clause-initial position at the same time) as automatically pointing to nominal orientation. Keeping these assumptions particular to the approach by Hettrich and colleagues in mind, I now present a summary of the quantitative data given in the various publications by Hettrich and his colleagues.

The usages are grouped into five main categories (Table 4.1) of nominal orientation ('adnominal usage', category 1), verbal orientation ('adverbal', categories 2–4), and ambiguous cases (category 5). For present purposes, the most important categories are 1a and 1b, which distinguish between prenominal and postnominal position of nominally oriented local particles. Some of the examples in 5a–c may semantically be cases of 1a or 1b, but the syntagmatics are ambiguous. The figures are

<sup>10</sup> '[...] qu'elle est une surérogation de notre esprit devant une structure linguistique qui était fondamentalement ambiguë'

<sup>11</sup> At one point, Renou concedes that 'all else being equal, the distance between particle and verb raises the chances of a prepositional value' (1956: 119, translation UR).

<sup>12</sup> 'En principe il ne faut admettre la préposition que dans les cas où l'autre interprétation est exclue.'

---

**TABLE 4.1. Categories for usages of local particles in publications by Hettrich, Casaretto, and Schneider**


---

1a	adnominal usage: local particle directly preceding noun
1b	adnominal usage: local particle directly following noun
2a	adverbial: local particle precedes verb, unverbated
2b	adverbial: local particle follows verb directly
3	adverbial: clause-initial (tmesis)
4	adverbial: other position in the sentence
5a	ambiguous: noun–loc. part.–verb
5b	ambiguous: loc. part.–noun . . . verb
5c	ambiguous: verb–loc. part.–noun

---

presented in Table 4.2 (the percentages are rounded). I have added additional columns with percentages bearing on the questions at hand.

The figures in column A reveal that some of the local particles are never or only rarely attested with nominal orientation as categorized by Hettrich and his colleagues. This especially pertains to *apa*, *ava*, *ud*, *ni*, *niṣ*, *parā*, and *vi*, where less than 10 per cent of the attestations show nominal orientation. Some local particles, however, notably *acchā*, *anu*, and *tiras* are classified as displaying nominal orientation in half or more of the cases. Overall, the figures corroborate the assumption found frequently in the literature, that verbal orientation is the default.

Turning now to the question of syntagmatic position in usages with nominal orientation (column B):<sup>13</sup> here, the evidence is straightforward. Whether a local particle rarely (such as *niṣ* or *vi*) or frequently (such as *tiras* or *atī*) occurs with nominal orientation, with the exception of *acchā* and *ava*, we find postnominal placement only in up to a third of the cases and in part even less frequently. Some of the absolute numbers are of course small and do not allow for strong conclusions. However, the generally high percentages of prenominal position are stable across most of the local particles irrespective of absolute numbers. This stands in contrast to the widespread assumption that the local particles in Vedic occur in postnominal position much more often than in prenominal position. Note that also the accounts of Vedic prose (e.g. Cuny 1907) and Classical Sanskrit (e.g. Speijer 1968 [1886]) show that prepositional placement is found with many of the local particles and is not at all uncommon. In addition, the frequency of local particles with nominal orientation in postnominal position

<sup>13</sup> I exclude from consideration local particles which display nominal orientation in only around 1% of the cases. This involves *apa*, *ud*, and *parā*.

TABLE 4.2. Usages of Rigvedic local particles in publications by Hettrich, Casaretto, and Schneider

	sum	1a	1b	2a	2b	3	4	5a	5b	5c	A: How many with nom orient.? (1a/b)	B: How many postnom.? (1b out of 1a/b)	C: How many postnom. of total? (1b out of sum)
<i>acchā</i> 'to'	174	25	57	17	16	26	13	11	9	0	47% (82)	70%	33%
<i>ati</i> 'over'	199	50	17	49	2	7	15	32	8	19	34% (67)	25%	9%
<i>anu</i> 'after'	507	165	83	99	9	34	29	66	22	0	49% (248)	33%	16%
<i>apa</i> 'away'	235	2	1	107	2	80	35	3	4	1	1% (3)		
<i>abhi</i> 'to'	801	165	70	260	11	70	52	99	74	0	29% (235)	30%	9%
<i>ava</i> 'down'	243	8	10	133	4	47	12	22	7	0	7% (18)	56%	4%
<i>ud</i> 'up'	312	0	1	162	2	125	3	15	4	0	>1% (1)		
<i>upa</i> 'towards'	425	98	34	126	11	17	13	81	45	0	31% (132)	26%	8%
<i>tiras</i> 'across'	59	33	5	3	2	2	4	0	4	6	64% (38)	13%	8%
<i>ni</i> 'down'	665	23	3	468	11	100	9	31	20	0	4% (26)	12%	>1%
<i>nis</i> 'out'	128	6	3	62	3	31	6	13	4	0	7% (9)	33%	2%
<i>parā</i> 'away'	94	0	1?	62	0	23	1	7	0	0	>1% (1)		
<i>paras</i> 'before'	52	6	3	12	8	1	21	1	0	0	17% (9)	33%	6%
<i>purā</i> 'formerly'	59	12	1	0	0	0	46	0	0	0	22% (13)	8%	2%
<i>prati</i> 'against'	264	52	9	91	2	32	32	25	21	0	23% (61)	15%	3%
<i>vi</i> 'apart'	1049	15	5	624	41	179	94	73	13	5	2% (20)	25%	>1%

Note: The local particles are ordered as in Devanagari script according to the practice in the literature. The figures are derived from Casaretto (2010a, 2010b, 2010 [2011], 2011 [2012]a, 2011 [2012]b, 2011 [2012]c, 2012), Hettrich (1991, 1993, 2002), Hettrich et al. (2010), Schneider (2009 [2010], 2010a, 2010b, 2011, 2012). Not included are figures for *ā*, *api*, *paras*, *pra*, *sam*, which were not published at the point of writing, or on *adhi*, *antar*, and *pari*, which were analysed in Hettrich (1991, 1993, 2002), but not classified according to the five categories adopted in later publications. I furthermore exclude *sacā*, *saha*, *sākam*, *satrā*, and *s(u)mad*, which are not traditionally considered as belonging to the canonical local particles (see Delbrück 2009 [1888]: 470–1), as well as derived forms like *purastāt*.

measured against the respective total sum of occurrences (column C) is low, ranging below 10 per cent for all local particles but two, *acchā* and *anu*. Thus, the postnominal slot was certainly not well established in quantitative terms.

In sum, there are indeed some cases in Vedic Sanskrit where local particles occur as (i) free morphemes in (ii) post-nominal position with (iii) nominal orientation. However, the quantitative survey from the publications by Hettrich and associates reveals that it is in fact the *pre*nominal position which dominates with local particles in nominal orientation in the *Rigveda*. With regard to nominal orientation in general, the overall frequency is low and verbal orientation represents the default, even when categorized according to the grid applied by Hettrich and associates which is favourable

towards nominal orientation. All in all, there seems to be no foundation for assuming a special status of the postnominal position for local particles showing nominal orientation in the *Rigveda*.

#### 4.2.2 *Post-Rigvedic developments: The persistency of anu, ā, prati, and tiras*

I argued above that not only similarities in syntax and semantics, but also a temporal overlap between at least some of the local particles and the precursor forms of the postpositions in their earliest stages is necessary for a connection on a categorial level to be in any way a logical possibility. In this section, I examine the syntax and semantics of the small handful of local particles which survive beyond the Rigvedic stage in usages with nominal orientation.

While local particles frequently occurred as free morphemes in the *Rigveda*, most of them survive merely as bound preverbs by the stage of Classical Sanskrit, showing verbal orientation. Only a handful of local particles may on occasion still occur as free morphemes, which belong to those local particles that already have a high percentage of showing nominal orientation in the *Rigveda* (see Table 4.2). When they do, they now always display nominal orientation (or in some cases adverbial function) (see Whitney 1983 [1879]: 414). The local particles which are still attested as free morphemes in Classical Sanskrit are *anu* ‘after’, *antar* ‘inside’, *ā* ‘to, towards, from’, *purā* ‘before’, *prati* ‘against’, and *saha* ‘with’ (cp. Speijer 1968 [1886]: 115ff). Whitney (1983 [1879]: 414) narrows down this list, stating that only *anu*, *ā*, and *prati* occur with any frequency to speak of. Note that not only do these three elements survive into Classical Sanskrit as free morphemes, but they continue to be attested into Middle Indo-Aryan times. A fourth local particle is *tiras* ‘across’, only remarked on peripherally by Whitney and Speijer as one of the ‘inauthentic’ local particles which do not combine with verbs as freely (e.g. Whitney 1983 [1879]: 415]). *Tiras* is included here as a further local particle of Vedic Sanskrit which survives into Middle Indo-Aryan. These four particles are still used (if rarely) in Pali (as *anu*, *ā*, *tiro*, *paṭi*, cp. Fāhs 1989: 105–6) and there are stray attestations even at younger stages: for instance *prati* (as *pai*) figures in Jacobi’s collection of texts written in Mahārāṣṭrī, a Prakrit from the middle period of Middle Indo-Aryan (Jacobi 1967 [1886]). Now, do these residual elements show a correlation of form and function which prefigures the properties of the modern postpositions and might in this way represent the missing link between the elements in question?<sup>14</sup>

<sup>14</sup> A few surveys, notably Renou (1933) and Cuny (1907), deal with local particles in post-Rigvedic layers, but the focus is on usages in verbal orientation and the disappearance of tmesis. I am not aware of any study focusing on usages with nominal orientation, but there is a short section included in Cuny’s study of local particles in the *Śatapathabrāhmaṇa*. I will mainly draw on Cuny’s findings and add evidence from Pali.



When studying these elements, what emerges from the data is a particular syntactic correlate shared by them which directly bears on the question of a categorical link between local particles and postpositions. In the preceding section, I have surveyed the word order preferences of the local particles when nominally oriented. The figures for the long-surviving elements are 165/83 occurrences (pre- vs. postnominal placement) for *anu*, 52/9 pre- vs. postnominal placement for *prati*, and 33/5 for *tiras*. In other words, *anu*, *prati*, and *tiras* have a strong propensity for occurring in prenominal position when nominally oriented. While *ā* follows the noun in the majority of cases (there was at time of writing no article by Hettrich and colleagues, but a glance at Grassmann (1996 [1875]) gives an idea), there is one special usage where it occurs in prenominal position with ablatives (on which more below).<sup>15</sup> At first estimate, then, these four local particles stick out by showing high percentages of nominal orientation from the earliest attestations, and in addition display a special connection between nominal orientation and prenominal placement in the *Rigveda*.

For the *Śatapathabrāhmaṇa*, where we still find a certain number of local particles appearing as free morphemes with nominal orientation, Cuny (1907) notes that there is a division between those local particles occurring only in prenominal position, namely *ā* and *tiras*, and those occurring only in postnominal position, which are *adhi*, *ati*, *prati*, *antaḥ*, *abhi*, and *upa*. *Anu* is attested in both positions. The following examples show usages of *anu*, *ā*, and *tiras* in prenominal position and with nominal orientation in the *Śatapathabrāhmaṇa* (references from Cuny 1907):<sup>16</sup>

- (73) *tādetāsyaiṣvānu*                      *mātrām*  
*tād etāsya eva ānu*                      *mātrām*  
 PRT DEM.GEN.SG.M PRT after    measure.ACC.SG.F  
 ‘(and as this clarifying process takes place) in accordance with the measure of that (process of breathing)’ (ŚaB 1.1.3.2)
- (74) *sārvamātmānaṃ*                      *sāminddha*                      *ā*    *nakhebhyó*                      *’tho*  
*sārvam ātmānaṃ*                      *sāminddha*                      *ā*    *nakhebhyó*                      *’tho*  
 all.ACC.SG.M body.ACC.SG.M    burn.PPP.NOM.SG.M    to    nail.ABL.PL.M/N    and  
*lómabhyaḥ*  
*lómabhyaḥ*  
 hair.ABL.PL.N  
 ‘(He) burned the entire body from nails to the hair.’ (ŚaB 1.4.3.10)

<sup>15</sup> The article on *ā* has appeared in the meantime, but too late to be included here in detail. However, Antje Casaretto (pc) informs me that the percentage of *ā* with verbal orientation is in fact higher than suggested by Grassmann’s entry.

<sup>16</sup> Note that the system of marking accent in the *Śatapathabrāhmaṇa* differs from the system in the *Rigveda*.

- (75) *hyèṣátirastāmāṃsi* *darṣata* *iti*  
*hí eṣá tīras tāmāṃsi* *darṣata* *iti*  
 because DEM.NOM.SG.M through darkness.ACC.SG.N see.MID.3SG QUOT  
 ‘because he is visible **through darkness**’ (ŚāB 1.4.1.29)

In Pali, finally, the late survivors *anu*, *ā*, *prati*, and *tiras* have become very rare. It can be gathered from a search in the digital Pali texts that they largely retain their distribution as in Old Indo-Aryan: *anu* is hardly used but attested several times in prenominal position, *ā* now as a rule appears prenominally, *prati* (as *paṭi*) is attested in both positions but with a preference for postnominal placement (see also Fahs 1989: 105–6), while *tiras* (as *tiro*) seems only attested as a free morpheme in non-canonical literature with several cases of both pre- and postnominal position. This illustrates that the optional prenominal placement persists with all elements and is in fact the only pattern remaining for *ā*.

For *ā*, there is additional semantic evidence for a correlation of prenominal placement with nominal orientation already from early on. In the *Rigveda*, *ā* occurs both pre- and postnominally. In contrast with other local particles, the two positions are associated with distinct semantics. Starting with the more frequent postnominal placement, *ā* often does not serve as a local particle in the sense of making a semantic contribution as a spatial marker, but rather as a reinforcer, as it emphasizes the spatial case semantics. For instance, *ā* emphasizes the ablative in example (76). When preceding the noun, by contrast, it has the otherwise unattested effect of turning the ablative meaning into its exact opposite, so that *ā samudrāt* gives ‘to the ocean’ (ex. (77)).<sup>17</sup>

- (76) *ā* *no* *divó* *bṛhatāḥ* *párvatād* *ā*  
 LP ACC.1PL sky.ABL.SG.M great.ABL.SG.M mountain.ABL.SG.M Ā  
*sárasvatī* *yajatā* *gantu*  
 Sarasvati.NOM.SG.F worthy\_of\_sacrifice.NOM.SG.F go.IMP.3SG  
*yajñám*  
 sacrifice.ACC.SG.M  
 ‘From the great sky, from the mountain, Sarasvati, worthy of sacrificial acts,  
 should come to us for the sacrifice.’ (RV 5.43.11)
- (77) *yatí* *giribhya* *ā* *samudrāt*  
 go.PPA.NOM.SG.F mountain.ABL.PL.M Ā sea.ABL.SG.M  
 ‘going from the mountains to the sea’ (RV 7.95.2)

This reversal of semantics has long puzzled specialists (e.g. Brugmann 1911: 819). Delbrück at one point commits to regarding postnominal placement along with its

<sup>17</sup> In this example, *ā* must modify *samudrāt* and not *giribhya(s)* because otherwise, *samudrāt* would have an ablational reading.

semantics as the default and remarks that further examination will need to take into account the question of placement:

In explaining the construction of *ā* before the ablative, one will have especially to consider position. The original and, so to say, natural connection is present in *samudrād ā* ‘from the ocean’, the reversal in position and sense is *ā samudrāt* ‘up to the sea’. (Delbrück 2009 [1888]: 452, translation UR)<sup>18</sup>

I propose, by contrast, that *ā* shows its core spatial meaning when in prenominal position, while its meaning is bleached in postnominal position, a notion which has been entertained by various authors since the nineteenth century (see Osthoff 1879: 103, also Delbrück himself 2009 [1888]: 453). The same core meaning often shows in usages with verbal orientation, as illustrated by the pairs *gacchati* ‘he goes’ vs. *āgacchati* ‘he comes’ or *kṛṇoti* ‘he does/makes’ vs. *ākṛṇoti* ‘he brings, prepares’.<sup>19</sup>

In sum, those local particles which survive into early Middle Indo-Aryan with nominal orientation show a marked preference overall for prenominal position across Mantra Vedic, Vedic prose, and into Pali, in comparison with the other local particles which vanish earlier in this usage. This is especially clear from the totals of occurrence in the *Rigveda*, from the fact that they are the only elements at all attested in prenominal position in post-Vedic times, and from the development of *ā*, which in later stages only ever appears in prenominal position.

Thus, it is not only the case that there is no etymological connection between local particles and modern postpositions, and that the majority of local particles do not survive beyond Rigvedic Sanskrit with nominal orientation, but also that the few survivors in their rare occurrences do not fulfil the conditions to foreshadow the emerging postpositions, showing a preference for prenominal position. In fact, it appears that they may have survived so long because of their preference for prenominal position (see the next section). *MADHYE* and *UPARI* can be interpreted as occurring as periphrastic expressions of the morphological locative only from late Old and early Middle Indo-Aryan stages onwards. Since the last local particles surviving into this period as free morphemes do not show the necessary syntactic preconditions for being the ancestors of the modern postpositions, I conclude that there is no empirical basis for this claim.<sup>20</sup>

<sup>18</sup> ‘Bei der Erklärung der Construction von *ā* vor dem Ablativ wird man die Stellung besonders zu beachten haben. Die ursprüngliche und so zu sagen natürliche Verbindung liegt vor in *samudrād ā* [‘]vom Meere her[‘], die Umkehrung nach Stellung und Sinn ist *ā samudrāt* [‘]bis zum Meere[‘] hin.’

<sup>19</sup> Antje Casaretto (pc) suggests that we may not in fact be dealing with one and the same element in all these various usages of *ā*.

<sup>20</sup> In the next chapter, more data will be presented which also speak against a connection to the local particles, in that the modern postpositions go back to constructions involving dependent genitives (or dependent objects in the case of participles grammaticalizing), rather than local case forms that the local particles used to combine with.

Before I move on to enquire into the reasons for the loss of the local particles with nominal orientation in Indo-Aryan, note that the word order evidence presented aligns Indo-Aryan (and the same may be assumed for Iranian: see Bubeník 2006b) with other branches of the family with respect to the syntax of the local particles. It has been assumed that the local particles in western branches of the family from early on preferred the prenominal slot and thus grew into prepositions, while eastern branches show a preference for the postnominal slot (e.g. Hewson 2006: 13), which allegedly explains the rise of postpositions in Indo-Aryan. However, it is only when verbally oriented local particles are included that the numbers of local particles following local case forms are high. Nominally oriented ones prefer the prenominal slot. Thus, there may not have been such a marked split between east and west in Indo-European after all, as regards the syntax and semantics of the local particles. The marked split, by contrast, lies in the loss of the local particles in Indo-Aryan and the missing connection to the modern category of postpositions.

### 4.3 Why were the local particles not reanalysed as adpositions?

Because the local particles have been made protagonists in the story of the rise of phrase structure in the Indo-European family, I address the question of why nominally oriented local particles fell out of use in Indo-Aryan. In doing so, I aim at a better understanding of the factors that led to Indo-Aryan taking a different route not only to adpositions, but also to phrasality as a new syntactic principle of organizing nominal expressions. I propose that language-specific differences in the interplay of position and clause type bearing on accentuation created unfavourable conditions for a reanalysis as adpositions in Indo-Aryan in contrast with other branches of Indo-European, including its closely related sister branch of Iranian.

A prosodic difference in Vedic distinguishes finite verbs occurring in main and subordinate clauses.<sup>21</sup> In the former, they lose their accent and encliticize to whichever element stands to their left, while they retain their accent in subordinate clauses (and when initial in main clauses). When enclitic in main clauses, the verb often cliticizes to a verbally oriented local particle as in example (78)—this is the pattern that becomes the model for the unverbated forms that survive through the ages: see

<sup>21</sup> Some of the oldest Vedic texts note accentuation (e.g. *Rigveda*, *Atharvaveda*, *Śatapathabrāhmaṇa*), on the basis of which the Vedic system of word and sentence accentuation is fairly well understood (e.g. Whitney 1855–56, 1869–70). The Vedic pitch accent system consists of a high tone (*udātta*), a low tone (*anudātta*), and a falling tone (*svarita*). The main accent is contributed by the high tone and in some cases the falling tone, the latter being a secondary creation as a fusion of a high and a low tone of formerly two separate syllables. Thus, if a word carries a high tone or an ‘independent’ falling tone (dependent falling tones resulting from the fall *after* a high tone syllable, not marking an accented syllable themselves), it is considered accented (i.e. bearing a main accent), otherwise it is unaccented. In the Roman transliterations of the Vedic texts which note accentuation, an acute is commonly used for a high tone accent, and a grave sign for the independent falling tone accent. Low tones and dependent falling tones are not marked.

Kuryłowicz (1958) or Turner (1916, 1936) (prosodic unity can be gleaned from the lack of an acute sign marking accentuation on the inflected verb form). However, the finite verb may also encliticize to other kinds of element, for instance to a noun as in the first verse of the *Rigveda* in example (79) or to nominally oriented local particles as in examples (80)–(84) (prosodic units are marked in bold):<sup>22</sup>

- (78) *índrasya nú vīryàṇi prá vocaṃ*  
 Indra.GEN.SG.M now power.ACC.PL.N forward proclaim.AOR.INJ.1SG  
 ‘I now proclaim Indra’s strengths.’ (RV 1.32.1)
- (79) *agním īle puróhitam yajñásya*<sup>23</sup> *devám*  
 Agni.ACC.SG.M praise.1SG principal.ACC.SG.M sacrifice.GEN.SG.M god.ACC.SG.M  
*ṛtvijam hótāram*  
 sacrificing\_at\_right\_time.ACC.SG.M sacrificer.ACC.SG.M  
*ratnadhátamam.*  
 distributing\_great\_riches.ACC.SG.M  
 ‘I praise Agni, the one in charge of the sacrifice, the god who sacrifices at the right time, the sacrificer who bestows great riches.’ (RV 1.1)
- (80) *ṛtásya pánthām ánv emi sādhuṃyá*  
 right.GEN.SG.N path.ACC.SG.M after go.1SG right\_way  
 ‘I go along the path of truth in the right way.’ (RV 10.66.13)
- (81) *parāyatīm mātāram ánv acaṣṭa*  
 leaving.ACC.SG.F mother.ACC.SG.F after look.IMPF.3SG  
 ‘He looked after [i.e. his gaze followed] the leaving mother.’ (RV 4.18.3)
- (82) *sá devó devān práti paprathe*  
 DEM.NOM.SG.M god.NOM.SG.M god.ACC.PL.M against expand.PRF.3SG  
 ‘The god expanded against the gods.’ (RV 2.24.11)

<sup>22</sup> Apart from discourse particles and clitic pronouns, a Vedic word usually carries one main accent. Whereas this is a system primarily of word accent, sentence level effects exist as in the deaccentuation of non-initial vocatives, and—centrally to the present topic—regarding a systematic variation in the accentuation of finite verbs (and, linked to it, of local particles). In search of an explanation for the difference of accentuation between finite verbs in main and subordinate clauses, Kuryłowicz (1958: 97–8) proposes that it is the interplay of verbs with local particles which accounts for the deaccentuation. Accented local particles deprived finite verb forms of their accent, a pattern that was later extended to all instances in main clauses (apart from instances in initial position). This scenario does not seem convincing for the reason that finite verb forms do not bear an accent in main clauses irrespective of position (apart from the clause-initial slot), and no matter what kind of word they lean to. Moreover, finite verbs ‘win out against’ local particles on the level of accentuation in subordinate clauses. Hettrich (1988: 779), following earlier authors, proposes another argument, based on the common pattern of subordinate clauses preceding main clauses. When the sentence pitch is falling over the length of the utterance of the complex sentence, the very final element, typically the finite verb form, would have the lowest pitch. At a later point in time, this pattern would have extended to other positions within the main clause.

<sup>23</sup> The slash indicates a line break.

- (83) *bhadrām*                      *te*                      *agne*                      *sahasinn*                      *ānikam*  
 pleasing.NOM.SG.N      GEN.2SG      Agni.VOC.SG.M      mighty.VOC.SG.M      face.NOM.SG.N  
*upākā*                      *ā*      *rocate*                      *sūryasya*  
 proximity.LOC.SG      to      shine.3SG      sun.GEN.SG.M  
 ‘Your pleasing face, oh mighty Agni, shines beside the sun.’      (RV 4.11.1)
- (84) *mātā*                      *pitāram*                      *rtā*                      *ā*      *babhāja*  
 mother.NOM.SG.F      father.ACC.SG.M      order.LOC.SG.N      to      distribute.PRF.3SG  
 ‘The mother distributed to the father according to rule.’      (RV 1.164.8)

That example (80) is nominally oriented can be gathered from the fact that the accusative does not by itself occur in the function of the ‘accusative of extension’ but only with accompanying local particles (see Casaretto 2010a: 15–16). Example (81) is another instance of an accusative of extension. In (82), *prāti* specifies a goal in combination with the accusative. In (83), *ā* emphasizes the locative meaning of the preceding nominal element, a common usage of this particular local particle (see the preceding section). Finally in (84), *ā* appears to combine with a nominal element even more removed, namely *pitāram*.<sup>24</sup>

What we see in examples (80)–(84) is a mismatch between prosody and semantics. The local particle forms a prosodic unit with the enclitic verb, but is nominally oriented.<sup>25</sup> In other ancient varieties of Indo-European (possibly apart from Iranian—see the next paragraph), by contrast, there is no such general encliticization of finite verbs, and thus no comparable prosodic unification with local particles, no matter what their semantic orientation (see Kurylowicz 1958). I suggest that the resulting mismatch between prosody and semantics of local particles in Vedic Sanskrit was a factor in the loss of local particles with nominal orientation in this particular branch. The constellation leads to a skewing of the data towards prosodically unified strings, which blocks reanalysis of nominally oriented local particles as adpositions.<sup>26</sup>

The sister branch of Iranian furnishes evidence that this analysis is on the right track. It is assumed that Old Iranian languages show the same prosodic interaction between finite verbs and local particles in terms of dependency of clause type and position as Vedic Sanskrit on a general level (see Delbrück 1900: 64–5; Reichelt 1978

<sup>24</sup> I thank Antje Casaretto for pointing out to me the particular structure of this example.

<sup>25</sup> Note that these examples involve some of the four long surviving local particles, which have a particular propensity towards nominal orientation. Even here, however, we find instances where they merge prosodically with the finite verb irrespective of orientation.

<sup>26</sup> In addition, prosodic unification is already frequent even if we only consider local particles with verbal orientation, covering univentions in subordinate clauses as well as enclitic verbs leaning to local particles in main clauses. Furthermore, while the merger in main clauses is at first merely prosodic, recall that it is this pattern which lexicalizes—evidence for the emergence of a tight unity of the merged forms in main clauses. These aspects further support the scenario that prosodic unification in general, and the main clause pattern in particular, may indeed have stood in the way of a reanalysis of local particles as adpositions, which requires unbound morphemes.

[1909]: 266; West 2011: 117). However, Bubeník (2006b: 136) points out a stronger preference for prenominal rather than postnominal placement of nominally oriented local particles in the Old Iranian varieties in comparison with Vedic Sanskrit. As a consequence, prosodic mergers of nominally oriented local particles with finite verbs would have been rarer than in Vedic Sanskrit. It is in accordance with this difference between the two sister branches that local particles survive in Iranian as prepositions, for example in modern Persian (see Bubeník 2006b: 142ff). This scenario of the development in Iranian, however, requires more detailed examination in the future.

#### 4.4 Conclusion

In this chapter, I have examined how the rise of adpositions and adpositional phrases in Indo-Aryan is commonly depicted in the literature. It has been claimed that Indo-Aryan postpositions descend from local particles, just like adpositions in other Indo-European branches. That the Indo-Aryan postpositions are not connected to the local particles etymologically had not been raised as a problem for this claim. Testing this proposal, I reviewed the syntax and semantics of the local particles in Old and Middle Indic as to their usages with nominal orientation. It turns out that, in the language of the *Rigveda*, nominally oriented local particles display a tendency for prenominal rather than postnominal position. Moreover, the only local particles which still rarely occur beyond Rigvedic times as free morphemes with nominal orientation show a certain correlation with prenominal placement in comparison with other local particles, or even only occur prenominally. Thus, there is no evidence for a special postnominal slot for nominally oriented local particles either in Old or early Middle Indic where the oldest of the modern postpositions slowly begin to be used as periphrastic expressions of case, showing a preference for the postnominal position since Old Indic times. Therefore, contrary to what has been assumed, there appears to be no foundation for assuming a connection between the local particles and the postpositions.

The reason for the disappearance of the local particles with nominal orientation must be sought in the period of the oldest Vedic layers, since it is in the transition to younger stages of Old Indic that these usages greatly decrease. I pointed out that the Vedic patterns of accentuation can lead to mismatches of prosody (fusion with the verb) and semantics (nominal orientation) in main clauses with regard to local particles hosting enclitic verbs. I proposed that this constellation was a factor in the loss of local particles as free morphemes with nominal orientation. The local particles having been ruled out as potential predecessors of the modern New Indo-Aryan postpositions, I will look for their origins elsewhere in the next chapter, namely in the earliest attestations in Old and Middle Indic of the oldest Hindi postpositions *mē* and *par*. It will become clear that we are not only dealing with younger elements grammaticalizing, but—crucially—with a syntactically alternative road to phrase structure.





## The components of the source constructions

In the preceding chapter, I showed that the assumption that Indo-Aryan postpositions continue the category of PIE local particles cannot be confirmed. In this chapter, I turn to the original components of the modern postpositional phrases themselves. There are different ways of approaching the question of which original constructions the modern phrases arose from. I have pointed out that the Hindi postpositions generally govern the oblique case. Thus, reconstructing the origin of the oblique suffixes would seem a straightforward strategy to find out what cases *MADHYE*, *UPARI*, etc. used to combine with. The forms in need of reconstruction are the oblique singular *-e* of masculine *ā*-stem nouns (other classes do not have a distinct oblique singular form), as well as the oblique plural in *-ō* which exists in all declensions. While there is general unanimity in the literature that oblique plural *-ō* ultimately derives from (versions of) genitive plural suffixes (Beames 1970 [1875]: 218f; Kellogg 1972 [1875]: 128; Oberlies 1998: 1–2), there are different proposals for the origin of *-e*. While Oberlies (1998: 2) favours the Old Indo-Aryan locative singular ending in *-asmin* (a pronominal form which later also attaches to nouns) and Kellogg a descent from Apabhramsha genitive singular *-he* (1972 [1875]: 126), most scholars assume a hybridization of several late Middle Indo-Aryan forms including the genitive (genitive-dative), ablative, locative, and even instrumental (e.g. Beames 1970 [1875]: 211; Oberlies 1998: 6, fn. 9; Verbeke and de Cuypere 2009: 7–8).

Thus, morphological reconstruction does not yield a straightforward answer to the question of the components of the source constructions. What is more, note that even a successful identification of a morphological forerunner could not solve the question of the original components, as this is a matter not only of form but also of function. For instance, the genitive at an early time came to be used in dative function and, by the stage of Pali, the dative had almost completely vanished. So, while the oblique plural most likely descends from a morphological genitive, it functionally continues a genitive-dative with extensions into other case functions. Similarly, the reconstruction of oblique singular *-e* is not only a challenge on the morphological

level, but the polyfunctionality of probable forerunner forms is just as difficult to capture. To give only one example, in Old Awadhi, the oblique is used for instrumental, locative, and ergative functions, and is the inflectional form governed by postpositions, a result of strong syncretism (cp. Saksena 1971: 122–3).

As a reconstruction of the history of the oblique forms will only take us so far, I examine the contexts in which the earliest forerunner forms of the modern postpositions occurred during the Old and Middle Indic period, attempting to discern the original components of today's phrases. Since we are dealing with non-configurational language varieties, I am not searching for morphosyntactic constructions in the sense of phrases with fixed word order and elements obligatorily filling specific slots at this point. The selection of a specific syntagm does not happen before grammaticalization sets in in the transition from Middle to New Indic. Here, I only concentrate on the question of which cases these forerunner forms combined with, irrespective of order.<sup>1</sup> On this foundation, I develop a scenario that can account for the genesis of the postpositional phrases with regard to their components. The origins of the postpositional word order will be treated in the next chapter. This chapter is structured as follows.

In section 5.1, I outline the syntactic properties of *MADHYE* and *UPARI* in Vedic Sanskrit, which at first glance suggest a scenario similar to that of the local particles in other Indo-European branches, namely the reanalysis of groups with local case forms. While this would support assertions of a single path to configurationality in Indo-European (e.g. Hewson and Bubeník 2006), I show in section 5.2 that a look at data from Pali blurs the picture. Here, and in part already in late Old Indic stages, not only relational nouns like *MADHYE*, but also formerly zero-valent elements like *UPARI* start construing with genitives. This opens up the possibility of an alternative scenario, namely the origin of the postpositional phrases in relational noun constructions (i.e. constructions consisting of a spatial noun, body part term, or other relational noun together with a possessor encoded as a genitive).<sup>2</sup> But how was it possible for *UPARI* and several other elements to acquire a slot for a genitive? In sections 5.3.1 and 5.3.2, I outline a proliferation of relational noun constructions in post-Vedic times, in the wake of which spatial adverbs acquired argument slots, a development I refer to as the *post-Vedic genitive shift*. In subsequent sections, I show how this reanalysis of zero-valent elements like *UPARI* as nominalized heads in

<sup>1</sup> In this chapter, I distinguish between neutral reference to 'combinations' of elements (i.e. occurring in any position) vs. 'groups,' i.e. cases where the elements in question stand in juxtaposition. This distinction is important as only syntagmatically connected elements are affected by grammaticalization processes.

<sup>2</sup> Note that I use the term 'possessor' in this study for any type of ground used for spatial relations encoded by *MADHYE*, *UPARI*, etc. This not only includes possessors in a strict sense, i.e. sentient beings, but indicates any entity with regard to which a certain spatial, or in later historical stages also non-spatial, relation is encoded. In the types of constructions that are the focus of this study, possessors in nominal expressions are overwhelmingly expressed as genitives and later obliques. However, note that there are also alternative possibilities, which do not play a major role in this study.

relational noun expressions became possible. I argue that combinations of case syncretism (5.3.3.1) and semantic ambiguity (5.3.3.2) allowed for reanalyses of local case forms as genitives. While we witness a considerable expansion of relational nouns, with other adverbs, adjectives, and participles ‘aligning’ to their syntax of taking dependent genitives, combinations with local case forms also remain attested until Apabhramsha. Eventually, it is on the basis of semantic data presented in section 5.4 that relational noun expressions can be identified as the input constructions for the grammaticalization processes at hand. Based on this evidence, I conclude in section 5.5 that phrases can arise in non-configurational languages not only from symmetrical groups of syntactically independent, co-ranking elements, but also from asymmetrical groups that involve a nominal or nominalized head and a genitive dependent (or, alternatively, a participial head and a dependent nominal expression). I link these findings to grammaticalization theory, arguing that we are here seeing two subtypes of grammaticalization processes showing differences on several levels, both formal and functional.

### 5.1 From symmetrical group to phrase?

It was pointed out in section 2.6 that syntactically fixed constructions are absent from non-configurational languages such as the Old and Middle Indic varieties. Instead, groups in the sense of Lehmann (1991) or Himmelmann (1997) of syntactically independent elements in juxtaposition form the basis for grammaticalization processes. I quote Himmelmann’s (1997) definition of a ‘group’:

[A] group is a syntagm, in which syntactically independent and co-ranking elements are juxtaposed. The elements may form a sequential and prosodic unit, but they do not have to. That is, the relations between the elements of a group may be of a purely semantic and possibly also of a morphological nature. Nominal syntagms in Latin or Nunggubuyu are typical examples. (Himmelmann 1997: 137, translation UR)<sup>3</sup>

Adpositional phrases in other Indo-European branches arise from just such groups of syntactically independent and co-ranking elements, local particles and local case forms, as shown in the previous chapter (see also sections 1.1 and 2.5). I argued at length that there is no connection between the ancient local particles and the modern postpositions in Indo-Aryan. However, this does not preclude the possibility of a structurally parallel reanalysis of a group of co-ranking adverbial elements being reanalysed as an adpositional string in Indo-Aryan. Indeed, this development seems

<sup>3</sup> ‘[E]ine Gruppe ist ein Syntagma, in dem syntaktisch selbständige und gleichrangige Elemente juxtaponiert sind, Die Elemente können eine sequentielle und prosodische Einheit bilden, müssen es aber nicht. Das heißt, die Beziehungen der Elemente einer Gruppe untereinander können rein semantischer und ggf. auch morphologischer Natur sein. Nominale Syntagmen im Lateinischen oder Nunggubuyu sind typische Beispiele dafür.’

at first glance to be supported by evidence from Vedic, as I will show below. If this scenario turned out to be correct, this would align Indo-Aryan at least on this structural level to other branches of the family after all, irrespective of a lack of etymological connection and irrespective of a different word order pattern. In the sections thereafter, however, I will show how the inclusion of evidence from Middle Indic disproves this scenario and that instead, an origin in relational noun expressions must be assumed. But first, let us turn to the situation in Vedic Sanskrit.

Compare the following examples showing the distribution of the local particle *ádhi* as an adverb on its own and combining with a local case form, and the same two options for MADHYE and UPARI:

- (85) (= ex. (32)) *ádhi*

<i>eṣá</i>	<i>súryeṇa</i>	<i>hāsate</i>	<i>pávamāno</i>
DEM.NOM.SG.M	sun.INS.SG.M	race.MID.3SG	purify.PPM.NOM.SG.M
<i>ádhi</i>	<i>dyávi</i>		
above	sky.LOC.SG.M		

‘This one, purifying himself, races with the sun, **above**, in the sky.’ (RV 9.27.5)

- (86) (= ex. (71)) CASE<sub>loc</sub> + *ádhi*<sup>4</sup>

<i>ágne</i>	<i>ní</i>	<i>ṣatsi</i>	<i>námasā́dhi</i>	<i>barhíṣi</i>
ágne	ní	ṣatsi	námasā́ <b>ádhi</b>	<b>barhíṣi</b>
Agni.VOC.SG.M	down	sit.IMP.2SG	worship.INS.SG.N	above
				barhis.LOC.SG.N

‘Agni, sit down **on** the barhis with worship.’ (RV 8.23.26)

- (87) MADHYE

<i>náinam</i>	<i>úrdhvám</i>	<i>ná</i>	<i>tiryáñcam</i>	<i>ná</i>	<i>mádhye</i>	<i>pári</i>
ná enam	úrdhvám	ná	tiryáñcam	ná	mádhye	pári
NEG DEM.ACC.SG.M	above	NEG	below	NEG	MADHYE	LP

*jagrabhat*

comprehend.PRF.3SG

‘(No one) comprehended him above, below, or **in the middle**.’ (YV 32.2)

- (88) CASE<sub>loc</sub> + MADHYE

<i>śúṣaṃ</i>	<i>ná</i>	<i>mádhye</i>	<i>nábhyām</i>	<i>índrāya</i>
energy.ACC.SG.M	as	MADHYE	navel.LOC.SG.F	Indra.DAT.SG.M
<i>dadhur</i>	<i>indriyām</i>			
put.PRF.3PL	power.ACC.SG.N			

‘Like energy, (three goddesses) have laid power into the navel, in the middle, for Indra.’ (YV 21.54)

<sup>4</sup> I use CASE<sub>loc</sub> for case forms denoting spatial and other ‘semantic’ usages. In the remainder of this study, I will use the ‘+’-sign to indicate expressions where word order is irrelevant—the elements only need to occur in the same clause. If I refer to strings without the plus sign, I imply the respective order, i.e. GEN MADHYE differs from MADHYE GEN.

- (89) UPARI  
*nīcīnā* *sthur* *upāri* *budhnāḥ*  
 directed\_downward.NOM.PL stand.3PL UPARI foundation.NOM.SG.M  
 ‘(They) [= trees in the universe that grow upside down, UR] stick out down-  
 wards, the root is **above**.’ (RV 1.24.7)

- (90) CASE<sub>loc</sub> + UPARI  
*tisráḥ* *prthivīr* *upāri* *pravā* *divó*  
 three.ACC.PL.F earth.ACC.PL.F UPARI hovering.NOM.DU.M sky.GEN.SG.M  
*nākaṁ* *rakṣethe*  
 canopy.ACC.SG.M protect.2DU  
 ‘Hovering **above**, to the three earths, (you) protect the sky’s canopy.’  
 (RV 1.34.8)

These data show that MADHYE and UPARI possess similar distributional options to nominally oriented local particles (see section 4.2.1), as they may either occur on their own as in examples (87) and (89), or in combination with local case forms as in (88) or (90).<sup>5</sup> The examples quoted were selected because here MADHYE/UPARI and the local case form stand in immediate adjacency. These juxtaposed elements or ‘groups’ would have been the input strings to a reanalysis that parallels that of the local particles in other branches of Indo-European.<sup>6</sup> This hypothetical reanalysis is formalized in (91) and exemplified for local particles (represented by Latin *ad*) as also for MADHYE and UPARI:

- (91) REANALYSIS OF AN ADVERBIAL GROUP AS AN ADPOSITIONAL PHRASE  
 (ADV) (CASE<sub>loc</sub>) > {(ADV) (CASE<sub>loc</sub>)} > (ADP CASE<sub>loc</sub>)<sup>7</sup>  
 a. \*(*ad*) (*Romam*) > {(*ad*) (*Romam*)} > (*à Rome*) ‘at/to Rome’  
 b. (*nābhyām*) (*mādhye*)<sup>8</sup> > {(*nābhyām*) (*mādhye*)} > (*nābhi mē*) ‘in the navel’  
 c. (*prthivīr*) (*upari*) > {(*prthivīr*) (*upari*)} > (*prthivyō par*) ‘on(to) the earths’

<sup>5</sup> Note that an adjectival interpretation of MADHYE is a theoretical possibility in cases where a locative it combines with is masculine or neuter singular, because here MADHYE can be construed as agreeing with the latter. Such a modifying usage of MADHYE, however, is not attested in my corpus. This is important to note as the majority of singular locative case forms which stand in apposition with MADHYE in my corpus are masculine or neuter singular. In other cases, this interpretation can be ruled out on morphological grounds, for instance when the locative is feminine as in ex. (88).

<sup>6</sup> On a semantic level, there is certainly a difference between local particles and relational nouns like MADHYE. MADHYE still shows the literal meaning of ‘in the middle’ in Vedic Sanskrit. However, in certain cases it does not so much point to an actual centre but expresses a location within a space away from its boundaries, e.g. *madhye samudre* can be rendered as ‘far out at sea’ or German ‘mitten im Meer’ rather than ‘in the (more or less literal) middle of the sea’. In such cases, MADHYE shows a function not unlike that of the local particles. See Chapter 7 for detailed discussion of its semantic development.

<sup>7</sup> The brackets only encode syntactic unity, not dependency relations.

<sup>8</sup> In the quoted passage (ex. (88)), the order of the two elements is reversed. In fact, in the Vedic and Pali data of my corpus, there is no case of MADHYE following a local case form in direct adjacency, the pattern which would theoretically prefigure the modern phrase. In view of the general syntactic freedom, this gap is in all probability due to chance.

This formalization depicts an initial stage where syntactically independent and co-ranking adverbial elements on the one hand, be it a local particle or adverbially used relational noun, and a local case form on the other hand, co-occur in a clause. Since both express spatial relations which may combine (but do not interdepend) semantically, they often occur in juxtaposition. A transitory stage, in which the elements in question occur in a group more and more often, but are still syntactically independent, is marked by the additional curly brackets in the second step of the development. At some point during this stage, the adverbial form about to grammaticalize may be reanalysed as monomorphemic, should it have borne inflectional or derivational marks (as happened with *MADHYE*). In a third stage, the group is reanalysed as a syntactic unit, indicated by the single set of round brackets, and the case form is reinterpreted as being governed by the former adverbial form, now an adposition.<sup>9</sup> After reanalysis as a syntactic unit, the case marking on the noun may erode over time, as when Latin *Romam* (ACC) becomes French *Rome*.

In sum, the reanalysis of an adverbial group is the one obvious way in which Vedic *MADHYE* and *UPARI* may logically acquire a slot for an argument. In the case of *UPARI*, this is in fact the only way, judging by the Vedic data, as it only occurs with local case forms and does not, for example, take genitive dependents. The scenario sketched not only corresponds to Himmelmann's proposal that the grammaticalization of function words initiates with groups of co-ranking elements in non-configurational languages, but also, it corroborates the idea of a single path to configurationality in Indo-European—the only difference being that, while other branches rely on local particles, Indo-Aryan makes use of a different set of elements and prefers the reversed word order.

## 5.2 Relational noun expressions in Pali

We saw in the preceding section that the syntax of Vedic *MADHYE* and *UPARI* perfectly matches the scenario of a reanalysis of a group of co-ranking elements in juxtaposition. Particularly, the fact that *UPARI* is zero-valent does indeed seem to force this scenario, as *UPARI* must somehow acquire an argument slot for a dependent. What I am going to outline in this section, however, is an alternative trajectory that only emerges if data from Middle Indic are taken into account, which I will argue to be the correct one.

Vedic *MADHYE*, besides occurring with local case forms, could also take genitives (exx. (92) and (93)) and the same is attested in Pali (exx. (94) and (95)).

<sup>9</sup> An adverb may be reanalysed as an adposition in combination with different case forms. Thus, for instance German *in* continues to construe with datives and with accusatives, e.g. *Ich gehe im Raum (hin und her)* 'I am walking (back and forth) in the room' vs. *Ich gehe in den Raum* 'I walk into the room'.

(92) VEDIC SANSKRIT

*tasya madhye nividam dadhāti*

DEM.GEN.SG MADHYE Nivid.ACC.SG.F put.3SG

‘In the middle of it [= a recital, UR], (he) places a Nivid [= an invocation, UR].’ (Śā 15.3a)

(93) VEDIC SANSKRIT

*nimankṣye haṃ salilasya madhye*

plunge.FUT.MID.1SG 1SG water.GEN.SG.N MADHYE

‘I will plunge into the middle of the water.’ (AiB 8.21)

(94) PALI

*so haññati ñāṭisakhāna majjhe*

he kill.PASS.3SG relation\_and\_companion.GEN.PL.M MADHYE

‘He gets killed in the middle of his relatives and friends.’ (Jā II, p. 228)

(95) PALI

*pāyāsi rājā bahu sobhamāno*

go\_forth.AOR.3SG king.NOM.SG.M much shine.PPM.NOM.SG.M

*purakkhato nāriganassa majjhe*

honoured.NOM.SG.M group\_of\_women.GEN.SG.M MADHYE

‘The very splendid and honourable king comes forth in the middle of a group of women.’ (Jā IV, p. 464)

In contrast with its zero-valent syntax in Vedic Sanskrit, UPARI now also (and already in Epic and Classical Sanskrit, see Speijer 1968 [1886]: 124–5) takes genitive dependents at times:

(96) PALI

*seyyathāpi [...] payasotattassa nibbāyamānassa upari santānakam*

just\_as boiled\_milk.GEN.SG cool.PPM.GEN.SG UPARI scum.NOM.SG

*hoti, evam evaṃ pātur ahosi*

become.3SG just so manifest be.AOR.3SG

‘Even as scum forms on top of boiled milk that is cooling, so did (the earth) appear.’ (Dīgh III, p. 85)

(97) PALI

*atha kho Bhagavā tassa brahmuno upari vehāsaṃ*

now PRT Lord.NOM.SG.M DEM.GEN.SG.M Brahmā.GEN.SG.M UPARI in\_the\_air

*pallāṅkena nisīdi tejodhātum samāpajjitvā*

cross\_legged sit\_down.AOR.3SG element\_of\_flame.ACC.SG attain.CONV

‘And the Blessed One sat down cross-legged in the air above Brahmā, having transformed into flames.’ (Saṃ I, p. 144)

This data for UPARI open up the possibility of another origin for the modern New Indo-Aryan postpositional phrases, namely in relational noun expressions consisting of a nominal (MADHYE) or nominalized<sup>10</sup> head (UPARI) taking a genitive dependent, instead of an origin in adverbial groups, as hypothetically depicted in the previous section. Sometime in post-Vedic times, UPARI begins to appear with genitive dependents despite its earlier zero-valent syntax. In view of the relatively low absolute number of examples for UPARI in my corpus of Vedic Sanskrit (which, however, covers a large part of the attested Vedic literature), it is a theoretical possibility that it is only by chance that UPARI is not attested with genitives at this early stage. However, Delbrück (2009 [1888]: 163) suspected that UPARI does not in fact occur in this combination even though this had been previously claimed.<sup>11</sup> Evidence from cognates which also do not take genitives suggests that this conclusion is correct. For instance, Latin *super* only combines with accusatives and ablatives, and Old English *ofer* only with datives and accusatives. Therefore, I assume that UPARI indeed did not take genitives in Vedic Sanskrit.<sup>12</sup>

Compare now the figures for MADHYE (Table 5.1) and UPARI (Table 5.2) in different syntactic combinations in Vedic Sanskrit, Pali, and Apabhramsha which may help solve the puzzle of which original source construction to assume for OBL *mē/par*.<sup>13</sup>

These figures show that MADHYE in Vedic Sanskrit most frequently occurs in isolation or takes a genitive. However, it also occurs more frequently in combination

TABLE 5.1. Historical attestations of MADHYE in various syntactic environments

MADHYE	alone	+ LOC	+ ACC	+ GEN	+ GEN/LOC
Vedic Sanskrit (137 out of 168) <sup>a</sup>	63	6	1	66 <sup>b</sup>	1
Pali (118 out of 355)	41	13	0	43	21
Apabhramsha (35 out of 39)	0	1	0	34	0

<sup>a</sup> The first figure is the number of examples actually included in the analysis (i.e. without usages in compounds and without repetitions—see section 1.3 and the appendix). The second figure provides the number of all attestations.

<sup>b</sup> This excludes three instances where *madhye* is emphasized by a following *ā*.

<sup>10</sup> I use ‘nominalized’ here to refer to the acquisition of a slot for a genitive dependent, which partially aligns elements like UPARI to the syntax of relational nouns.

<sup>11</sup> ‘Für die Verbindung eines Genitivs mit echten Praepositionen, welche für *āti ānu upāri sácā* [sic!] angenommen worden ist, habe ich ein sicheres Beispiel nicht gefunden.’

<sup>12</sup> Note that it is an unrelated matter whether UPARI in origin was a locative (e.g. Dunkel 1992: 159). The important point here is that it is not attested with genitives in Vedic Sanskrit.

<sup>13</sup> I only provide the overall figures here. References to passages in the primary sources will be given in Chapters 6 and 7, where I go into the particulars of the different constructions.



TABLE 5.2. Historical attestations of UPARI in various syntactic environments

UPARI	alone	+ LOC	+ ACC	+ GEN	+ GEN/LOC
Vedic Sanskrit (33 out of 62)	29	1	3	0	0
Pali (36 out of 99)	11	13	1	9	2
Apabhramsha (43 out of 55)	5	1	0	35	2

with local case forms in Pali and it becomes increasingly difficult to discern whether a form is a genitive or a locative due to syncretism.<sup>14</sup> While UPARI only occurs on its own or in combination with local case forms in Vedic Sanskrit, it construes with genitives in almost a third of the cases in Pali. The Apabhramsha data for both MADHYE and UPARI suggest that the relational noun expressions win out. However, these figures are too low to allow for strong conclusions. Based on these figures, it cannot be declared with certainty which constructions to posit as the sources of today's postpositional phrases.

While relational noun expressions are well-known sources of adpositional constructions (e.g. Kahr 1975, 1976; Starosta 1985; Lehmann 2002 [1982]: 67ff), among which are the secondary adpositions of many Indo-European branches, they have not been identified as possible sources of phrases in strongly non-configurational languages. However, because of the data on post-Vedic usages of UPARI, one should consider the possibility of such an origin. Below, I show what the trajectory from such asymmetrical groups would have looked like on a structural level in contrast to (91), paralleling English *in front of*:

- (98) REANALYSIS OF RELATIONAL NOUN EXPRESSIONS AS ADPOSITIONAL PHRASES<sup>15</sup>
- |  |   |   |
|--|---|---|
| $N_{\text{loc}} (N_{\text{gen}})$                  | $> \{N_{\text{loc}} (N_{\text{gen}})\}$       | $> (\text{ADP } N_{\text{gen}})$                |
| a. <i>in front (of X)</i>                          | $> \{\text{in front (of X)}\}$                | $> (\text{in front of X})$ <sup>16</sup>        |
| b. <i>(samudrasya)</i> <sup>17</sup> <i>madhye</i> | $> \{(\text{samudassa}) \text{ madhye}\}$     | $> (\text{samudra mē})$ 'in the ocean'          |
| c. <i>(payasotattassa) upari</i>                   | $> \{(\text{payasotattassa}) \text{ upari}\}$ | $> (\text{uble hue dūdh par})$ 'on boiled milk' |

<sup>14</sup> The low number of ambiguous forms in Apabhramsha is due to the fact that there are few cases involving nouns of feminine declensions or *i*- and *u*-stems where the local cases have merged with the genitive.

<sup>15</sup> See Lehmann (2002 [1982]: 69) for a similar formalization.

<sup>16</sup> In English, many relational nouns also develop an adverbial usage, e.g. 'The man is standing in front', 'What will you do instead?' etc. In German, by contrast, relational nouns do not typically develop this usage, e.g. *Er steht im Rücken des Hauses* 'He stands at the back of the house' vs. \**Er steht im Rücken*.

<sup>17</sup> I exemplify this change with *samudra*- 'ocean', an etymon attested across all historical stages (borrowed into Hindi from Sanskrit).

Here, the initial situation is that of a relational noun (typically a spatial or a body part term) optionally taking a genitive that encodes the spatial reference frame. The brackets around only the possessor indicate that the possessor presupposes the head noun, but the head noun does not require the possessor to be present. As in (91), frequent occurrence in juxtaposition (indicated by an outer set of curly brackets) at some point leads to the reanalysis of the string as a syntactic unit where both elements are obligatory as in the Hindi version on the right, marked by a single set of brackets. As before, examples (92)–(97) were selected because they illustrate just such constructions that hypothetically developed into the modern phrasal pattern. In contrast with the scenario depicted in (91), there is a semantic and morphological dependency relation between the elements prior to reanalysis as phrasal units.

### 5.3 The post-Vedic genitive shift

That UPARI at some point acquires a slot for a genitive is in need of explanation. In the following sections, I will outline what I refer to as the *post-Vedic genitive shift*. This shift is driven by the proliferation of relational noun expressions, in the course of which not only UPARI, but also a host of other adjectival and adverbial elements were reanalysed as taking genitives. The case of UPARI is only the tip of the iceberg of a systematic expansion of periphrastic expressions of case consisting of nominal and nominalized heads and genitive dependents in post-Vedic times. In the search for the original components of the modern postpositional phrases, an examination of exclusively Vedic data would lead to the wrong conclusion that the Indo-Aryan postpositional phrases, in parallel with prepositional phrases in other Indo-European branches, originate in adverbial groups. Only the exploration of the post-Vedic genitive shift will point us to the correct source of the oldest of the Hindi postpositional phrases headed by *mē* and *par*, an origin in relational noun expressions.

What I refer to as the post-Vedic genitive shift has previously been briefly hinted at in a short passage in Bloch's *L'indo-aryen du Veda au temps modernes* (1934) and there is another brief hint in Fahs (1989: 103, see below). Bloch's words are here quoted only in a footnote, because they are written for specialists in Indo-Aryan.<sup>18</sup> In essence, Bloch points out that elements like MADHYE increasingly occur in

<sup>18</sup> '[I]n RV III 8, 2 *sámidddhasya śráyamāṇaḥ purástād* installed in front of the lighted (fire), *purástād* no longer accompanies an ablative or an accusative like *puráh* (in the Brāhmaṇas *upāriṣtād* accompanies the accusative like *upāri*), it is a noun which governs a noun. Besides *mádhye samudré*, cp. Pa. *majjhe samudde* in the ocean in the middle, we read for example *mádhye árṇasaḥ* in the middle of the surge. We find later ŚBr *ātmana upari* upon oneself. The Iranian and Vedic construction of *upari* is with the accusative and instrumental, the new construction is that of a normal locative. This process is widely extended in Sanskrit and Middle Indian. Thus for example *antike*, *samīpe*, *prṣṭhe*, *arthe*, *arthāya* (Pa. *atthāya*, *atthaṃ*, *hetoh* (Pa. *hetu*), *nimittam*, *nimittena*, *vaśād*, *vaśena* etc., are formed. It is this extension of noun groups which explains the absence of the prepositional system' (Bloch 1965: 159).

constructions with genitives rather than local case forms and that, in parallel, former zero-valent adverbs like UPARI begin taking genitives. These constructions consisting of a nominal or nominalized head and a dependent genitive become the new productive expressions in the domain of local cases, while the ‘prepositional system’—that is the old local particles—is not used any more.<sup>19</sup>

Over the following sections, I will outline the post-Vedic genitive shift in detail. I start with a general description of the expansion of relational noun expressions in Pali (section 5.3.1). In a second step, I focus on the reanalysis of adverbs such as UPARI (section 5.3.2). I then explore how such reanalyses could come about, identifying passages involving ambiguous case morphology and ambiguous semantics (section 5.3.3).

### 5.3.1 Proliferation of relational noun expressions

In contrast with Vedic Sanskrit, Pali heavily relies on various types of periphrastic expression of case, including the last surviving local particles, old adverbs, relational nouns, converbs (‘absolutives’ in Indo-Aryan terminology), and participles. I concentrate on relational noun expressions.<sup>20</sup> Consider the following near-minimal pair (from Fahs 1989: 104):

(99) PALI

- |   |                                 |                             |
|---|---------------------------------|-----------------------------|
| a. <i>antevalaṇṇjakādisu</i>              | <i>aguṇavādim</i>               | <i>adisvā</i> <sup>21</sup> |
| indoor_people.LOC.PL                      | saying_sth._unvirtuous.ACC.SG.M | not_see.CONV                |
| b. <i>antovalāṇṇjakānam</i> <sup>22</sup> | <i>antare</i>                   | <i>kañci</i>                |
| indoor_people.GEN.PL                      | inside.LOC.SG                   | anyone                      |
| <i>aguṇavādim</i>                         | <i>adisvā</i>                   |                             |
| saying_sth._unvirtuous.ACC.SG.M           | not_see.CONV                    |                             |

Both: ‘Among the people living in the house, (he) did not see (anyone who) uttered something infelicitous.’

In these two examples, a locative *antevalaṇṇjakādisu* (indoor\_people.LOC.PL) in a virtually parallel context is matched by the relational noun expression *antovalāṇṇjakānam*

<sup>19</sup> I disagree with Bloch only in one point, namely when he establishes a causal link between the proliferation of relational noun expressions and the disappearance of local particles in Indo-Aryan: ‘It is this extension of noun groups [i.e. of relational noun expressions, UR] which explains the absence of the prepositional system [i.e. of local particles, UR].’ It was shown in the previous chapter that the local particles do not survive as free morphemes with nominal orientation for independent reasons. In general, the loss and rise of syntactic classes resists explanation by simple causal relations.

<sup>20</sup> Participles play a role in the context of the origins of *kā/-e/-ī*, which will be taken up in Chapter 8.

<sup>21</sup> For this form, compare Oberlies (2001: 269).

<sup>22</sup> *Antevalaṇṇjakādisu* and *antovalāṇṇjakānam* are formed from shortened versions of the stem *antovalāṇṇanaka-* (Davids and Stede 2007 [1921–25]: ‘anto’).

*antare* (indoor\_people.GEN.PL inside.LOC.SG).<sup>23</sup> Overall, Fahs (1989) lists forty-nine such relational nouns as *MADHYE* and *antare*, each inflected in some local case form, which take genitive dependents and play an increasingly important role in the realm of local case. Other frequent forms are for instance GEN + *atthāya* (purpose.DAT) for purposes and reasons, GEN + *santike* (proximity.LOC) 'near to', GEN + *samaye* (time.LOC) 'at the time of', or GEN + *hetu* (reason.ABL) 'because of' (Fahs 1989: 107–10). Fahs (1989: 102) notes that relational noun expressions are the most frequent type of periphrastic expression. These are followed in decreasing order by adjectives and perfective participles, converbs, relative and correlative pronouns, and finally local particles.<sup>24</sup> In a selected section of the comparatively young texts of the *Jātakas* (the stories of Buddha's former births), where the propensity for periphrastic constructions is especially pronounced, Fahs (1989: 101–2) gives a percentage of 48 per cent relational noun expressions, followed by 32 per cent converbs, out of the entire number of periphrastic expressions he counts.

Let us take a closer look at the various relational nouns identified by Fahs. Of Fahs' list of forty-nine elements, only forty-two in fact derive from Old Indo-Aryan nouns. The other elements are, first, three nominal formations including a perfective participle (GEN + *kate* (from Skt. *kṛte* do.PPP.LOC.SG) 'for the sake of') and two adjectives (GEN + *dūre* (far.LOC.SG) 'far from' and its opposite GEN + *avidūre* (not\_far.LOC.SG) 'not far from').<sup>25</sup> I add to this group *param* 'beyond' (listed by Fahs in a different category of 'prepositions' that includes the remnant local particles besides other elements) as it also has an adjectival basis. The remaining four elements are *pacchato* 'behind', *purato* 'before', *puratthato* 'before', and *heṭṭhā* 'below'.<sup>26</sup> These elements have an adverbial base, extended by a suffix indicating a local case relation, either ablative *-tas* (Pali *-to*) or the locative suffix *-tāt* (the latter survives in *heṭṭhā* from *adhastāt*).<sup>27</sup> Etymologically, *pacchato* descends from an adjectival stem *paśca-* 'late', from which an adverb *paścā* 'later' is formed. *Purato* and *puratthato* derive from *puras* 'in front'

<sup>23</sup> It is difficult to assess precisely the degree to which the synthetic and analytic forms can be claimed to be synonymous. Regarding the shift towards periphrastic expressions in Pali, I follow the assessments of such widely read scholars as Bloch (1965: 158–9) or Fahs (1989: 102–5) who diagnose a strong increase in the use of such periphrastic constructions for local case meanings in late Old Indo-Aryan and Pali.

<sup>24</sup> This may appear like a rather variegated and in part surprising set of elements, as for instance pronouns do not commonly figure as periphrastic markers of case. However, these Old Indo-Aryan elements indeed have assumed functions which may be interpreted as expressing meanings at the periphery of case marking (cp. Fahs 1989: 105–7 for details).

<sup>25</sup> I pointed out in section 2.4.1.3 that the noun–adjective distinction in Vedic Sanskrit is weak. Here, too, the classification rests mainly on usage as modifiers and not on general morphological or syntactic properties.

<sup>26</sup> These elements differ from other nominal elements in that they cannot occur as final members in compounds.

<sup>27</sup> These suffixes function as case markers and are frequently employed in the formation of adverbs. However, while they correspond to ablative and locative-directional function respectively and may be added to different nominal and pronominal stems, they do not take part in the regular formation of case forms in the way the normal case suffixes do (cp. Whitney 1983 [1879]: 403–5).

(*puratthato* is a tertiary formation from *puras* > *purastāt* > \*(*purastāt* + *tas*) > *puratthato*). *Heṭṭhā*, finally, descends from an adverb *adhas* ‘below’ (< *adhas* + *tāt*). In the *Rigveda*, *paścā*, *puras*, and *adhas* occur as zero-valent adverbs and the extended forms are not yet attested, with the exception of *adhastāt*.

In contrast to UPARI, which I will discuss besides other elements in the next section, all eight elements just listed possess nominal morphology and thus, while not themselves nouns, allow for nominalized usage in the syntactic pattern of relational noun expressions. What is interesting about these eight elements is that, apart from the participle *kate*, their forerunner forms in the *Rigveda* are not attested with genitives (in the case of the four extended adverbs, this relates to their adverbial base forms: see below). In Pali and in part already in Classical Sanskrit, by contrast, they typically take genitives in the wake of the general proliferation of relational noun expressions.<sup>28</sup> For instance, the three adjectival forms, *dūre*, *avidūre*, and *param* occur in Vedic Sanskrit either in isolation or in combination with ablatives, for example *param vijñānāt* ‘beyond human knowledge’ (cp. the entries in Grassmann 1996 [1875], Böhtlingk and Roth 1855–75). In Pali, however, *avidūre* and *param* appear with genitives (Fahs 1989: 106, 108).<sup>29</sup>

The situation is similar in the case of the four extended adverbs. Their old adverbial base forms die out over time (only *adhas* as *adho* still occurs infrequently in Pali).<sup>30</sup> The extended forms begin to occur from later Old Indic layers (*puratas*) or from Pali (*pacchato* and *puratthato*) onwards and, as soon as they are attested, they take genitives. The added case suffixes lend these elements a nominal form and a slot for genitives along with it. Compare the following contrastive pairs of zero-valent base forms in Vedic Sanskrit and the extended forms with genitives in Pali, that is *paścā* vs. *pacchato*, *purás* vs. *purato* and *puratthato*, and *adhás* vs. *heṭṭhā*.

(100)	VEDIC <i>paścā</i>					
	<i>tvám</i>	<i>tyám</i>	<i>indra</i>	<i>súryam</i>	<i>paścā</i>	<i>sántam</i>
	2SG	DEM.ACC.SG.M	INDRA.VOC.SG.M	SUN.ACC.SG.M	behind	be.PPA.ACC.SG.M

<sup>28</sup> For the reasons outlined in section 1.2.1, I put a focus on Pali rather than Classical Sanskrit. However, the elements show the reanalysed usage in part already in Classical Sanskrit—hence the general label of the post-Vedic genitive shift.

<sup>29</sup> *Dūre* however is not attested with genitives, only with ablatives, according to Fahs (1989: 198).

<sup>30</sup> In one *Rigvedic* verse, *adhás* occurs with a form *padós* from *pád* ‘foot,’ which is either genitive or locative dual. Since there is no indication that *adhás* had already acquired a slot for a genitive, I follow Grassmann (1996 [1875] ‘*pád*’) who classifies *padós* as locative. Semantically, both cases would seem compatible. Compare the example:

<i>adháh</i>	<i>sapátnā</i>	<i>me</i>	<i>padór</i>	<i>imé</i>	<i>sárve</i>
below	enemy.NOM.PL.M	GEN.1SG	foot.DU.GEN/LOC.M	DEM.NOM.PL.M	all.NOM.PL.M
<i>abhiṣṭhitāḥ.</i>					
kick_below.PPP.NOM.PL.M					

‘All (my) enemies are kicked below (my) feet.’

(RV 10.166.2)

*purás kṛdhi / devānām cit tiró vāsam*  
 before do.IMP.2sg god.GEN.PL.M PRT against will.ACC.SG.M  
 ‘You, Indra, bring that sun which is **behind** to the front, even against the will  
 of the gods!’ (RV 10.171.4)

- (101) PALI *pacchato*  
*sobhenti āpagākūlaṃ, mahāleṇassa pacchato*  
 adorn.3PL river\_bank.ACC.SG great\_hermitage.GEN.SG behind  
 ‘(The rose-apple trees) adorn the river bank, **behind the main hermitage.**’  
 (Thera, p. 36)

- (102) VEDIC *purás*  
*áyám gáuḥ pṣ́nir akramīd áśadan*  
 á ayám gáuḥ pṣ́nir akramīd áśadan  
 to DEM.NOM.SG.M ox.NOM.SG.M spotted.NOM.SG.M step.IMP.3SG sit.IMP.3SG  
*mātáraṃ puráh / pitáraṃ ca*  
 mother.ACC.SG.F before father.ACC.SG.M and  
 ‘This spotted ox walked here and sat down **in front** of father and mother.’  
 (RV 10.189.1)

- (103) PALI *purato*  
*yā devatā candimasūriyānaṃ purato*  
 REL.NOM.PL.F divine\_being.NOM.PL.F moon\_and\_sun.GEN.PL.M before  
*dhāvanti*  
 run.3PL  
 ‘the divine beings which run **ahead of sun and moon**’ (Saṃ II, p. 266)

- (104) PALI *puratthato*  
*esa bhikkhu mahārāja paṇḍavassa*  
 DEM.NOM.SG.M monk.NOM.SG.M great\_king.VOC.SG.M Paṇḍava.GEN.SG.M  
*puratthato nissino vyagghusabho va, sīho va*  
 in\_front sit\_down.PPP.SG.M tiger.NOM.SG like lion.NOM.SG like  
*girigabbhare*  
 mountain\_cave.LOC.SG  
 ‘This monk, oh great king, sits **before the Pandava** like a tiger, like a lion in a  
 mountain cave.’ (Sutta, p. 73)

- (105) VEDIC *adhás*  
*tisráḥ pṛthivír adhó astu víśvāḥ*  
 three.ACC.PL.F earth.ACC.PL.F below be.IMP.3SG all.ACC.PL.F  
 ‘May he be **below all three earths.**’ (RV 7.104.11)

(106) Pālī *heṭṭhā*

*pupphassa*      *heṭṭhā*    *tiṭṭhāmi*  
 flower.GEN.SG.N    below    stand.1SG

‘I stand **below** the flower.’

(Apa II, p. 161, translation Black)

In sum, relational noun expressions proliferate in Pali and often occur in contexts which are otherwise occupied by synthetic morphological case forms. Moreover, not only relational nouns proper, but also elements other than nouns such as adjectives and adverbs start taking genitives at some point in post-(R̥ig)Vedic times. In the case of adjectives, we are dealing with nominalized usages where the adjective occupies the place of a nominal head. In the adverbial domain, old zero-valent adverbs fall out of use and instead extended forms with added case suffixes increase in frequency. Note that the elements surveyed in this section all possess nominal morphology, if only in the form of secondary case suffixes. Thus, while it is intriguing that they shift to new usages with dependent genitives, they show the morphological properties that allowed them to be employed in nominalized usages. In the next section, by contrast, I will turn to elements lacking any kind of nominal morphology, but which nevertheless also acquire a slot for a genitive dependent in post-Vedic times, one element of which is UPARI.

### 5.3.2 Pseudo-nouns

For the adverbs I consider now, which complete the list of Fahs’ periphrastic expressions of case in Pali, no origin in nouns or other elements with nominal morphology (e.g. adjectives, participles, or extended adverbs) can be ascertained. However, I show that they also are affected by the post-Vedic genitive shift.

In R̥igvedic Sanskrit, the elements in question function as simple zero-valent adverbs. But from some post-Vedic stage onwards, in part only from Pali onwards, they are suddenly attested with genitives in the same way as the elements discussed in the preceding section. In contrast with elements such as *pacchato*, *purato*, *puratthato*, and *heṭṭhā*, they do not carry a nominal case suffix, which could explain the nominalized usage. It is conspicuous that some of these elements have a phonological shape which ‘looks’ (or rather sounds) deceptively nominal. There are a limited number of possible nominal stem endings and a limited number of inflections which could be interpreted as local case forms. The elements in question, UPARI (pseudo-LOC), *antarā* ‘between’ (pseudo-INS), and also *saddhim* ‘with’ (pseudo-ACC), conform to such phonologically possible local case forms and may have been interpreted as such by a speaker of Pali (that is, of a variety that Pali is based on).<sup>31</sup> A further

<sup>31</sup> I continue to write UPARI and MADHYE in small capitals when I refer to them as used across several historical layers. When referring to a specific period, and in the case of the other elements discussed here, I use the phonological form of the respective historical period and write it in italics.

element that acquires a slot for a genitive is *anto* ‘inside,’ the sound shape of which does not conform to any local case form. I will first discuss *UPARI*, *antarā*, and *saddhim*, which are reasonably frequent, and later turn to *anto*.<sup>32</sup>

It was pointed out at the start of section 5.3 that *UPARI* in Pali (and already in Classical Sanskrit) occurs with a genitive, which is not attested for Rigvedic Sanskrit. What explains the appearance of the expression with genitives and its progressive expansion? Synchronically, Rigvedic *UPARI* is best described as an adverb similar in function to German zero-valent *oben*. In terms of sound shape, *-i* is the distinctive marker of locative singulars of consonantal stems. From the synchronic point of view of a speaker of early Middle Indo-Aryan, then, *UPARI* may very well have looked like a locative formation of a stem *upar-*, as analysed by Bloch for Middle Indic (Prakrit) and New Indic varieties (cp. fn. 12 above):

Hin[di] *ūpar* [figuring in the complex postposition *ke ūpar* ‘above,’ equally stemming from *upari*] and *par* [are] not a question of Skt. *upari*, but of a word constructed as a locative, Pkt. [Prakrit] *uppari*, Panj[abi] *uppar*; locatives in form are Rom[ani] *opre*, cf. *oprāl* abl., Mar[athi] *vaŕ*) (Bloch 1965: 179, emphasis added)

But why was *UPARI* not reanalysed as a locative before? I propose that the analogical influence from proliferating relational noun expressions in late Old and early Middle Indo-Aryan triggered the reanalysis. By contrast, relational noun expressions are much less dominant in the spatial domain in earlier stages of Old Indic, so that there would not have been a comparable analogical push to reanalyse *UPARI*.

Like *UPARI*, *antarā* ‘between’ also has the phonological shape of a regular local case form, in this case an instrumental of a hypothetical stem *antar-* (see the discussion in the following paragraphs for the invariant form *antar*). In parallel with *UPARI*, *antarā* functions as an adverb in the *Rigveda*, but in younger layers also occurs with genitives. Compare the following examples:

- (107) VEDIC *antarā*  

<i>yád</i>	<i>antarā</i>	<i>parāvátam</i>	<i>arvāvátam</i>	<i>ca</i>	<i>hūyáse</i> /
if	between	distance.ACC.SG.F	proximity.ACC.SG.F	and	call.PASS.2SG
<i>índrehá</i>		<i>táta</i>	<i>á</i>	<i>gahi</i>	
índra ihá		táta	á	gahi	
Indra.VOC.SG.M	here	there	to	go.IMP.2SG	

‘If you are called **near and far**, then come here from there, oh Indra.’  
(R̥V 3.40.9)

<sup>32</sup> A further element which figures in Fahs’ list is *anti* ‘opposite’. It is reconstructed to be an old locative and synchronically retains a locative shape, so that combinations with genitives do not come as a surprise. As it is very rare in Pali, I will not discuss this element further.



(108) PALI *antarā*

*sīho bhikkave [...] pāde pādaṃ accādhāya*  
 lion.NOM.SG.M brother.VOC.PL foot.LOC.SG foot.ACC.SG put\_on\_top.CONV  
*antarā satthināṃ/ naṅgutthaṃ anupakkhipitvā*  
 between thigh.GEN.PL tail.ACC.SG throw.CONV  
 ‘The lion, oh brothers, ... puts one foot on the other and throws his tail  
 between his thighs.’ (Aṅg II, p. 245)

*Antarā* probably derives from a coalescence of *antar* ‘inside’ and *ā* ‘at, to, towards’ which often occurred in juxtaposition (see Grassmann 1996 [1875]: ‘*antarā*’).<sup>33</sup> So, in the same way as with UPARI, *antarā* is not in fact based on some nominal *r*-stem, but only has the appearance of it. In the context of an increasing reliance on relational noun expressions, *antarā* like UPARI may very well be interpreted as an instrumental.<sup>34</sup>

The next element, *saddhiṃ*, is the only one on the list that is actually frequent in Pali.<sup>35</sup> As with the other elements, it is only after the *Rigveda* that we find it in combination with genitives. Compare the following examples of adverbial usage in Vedic Sanskrit and ‘nominalized’ usage in Pali with a genitive:

(109) VEDIC *sadhrīm*

*sadhrīm ā yanti pári bibhratiḥ páyo*  
 fixed\_on\_a\_goal to go.3PL around carry.INT.PPA.NOM.PL.F milk.ACC.SG.N  
 ‘(They) go fixed on a common goal, carrying milk.’ (RV 2.13.2)

(110) PALI *saddhiṃ*

*mamañ ñeva saddhiṃ paṭisammodati*  
 GEN.1SG PRT with return\_friendly\_greeting.3SG  
 ‘(He) greets me back friendly.’ (Vina II, p. 154)

The reconstruction of *saddhiṃ* from *sadhrīm* (< *sa* ‘with’, possibly with the root *dhar* ‘carry’, see Mayrhofer 1996) is now generally accepted.<sup>36</sup> In parallel with UPARI and *antarā*, *saddhiṃ* can synchronically be interpreted as a local case form, in this case as an accusative of an *i*-stem.

<sup>33</sup> In Pali, there are other local case forms based on an *a*-stem form *antara-*, which is assumed to derive from *antar* (cp. Mayrhofer 1992: ‘*ántara*’ for details). Examples are *antaram* ‘in, between’ (direction) and *antarena* ‘in, between’ (direction and location); see Fahs (1989: 108). Among the local case forms is a form *antarā* ‘from inside’ from an earlier ablative *antarāt*, which is not to be confused with the *antarā* at issue here. While the formation of *antarā* ‘between’ differs from these *a*-stem formations, the homophony might have been an additional factor in its nominal interpretation.

<sup>34</sup> This analysis is also implicit in the analysis given by the lexicographers Davids and Stede (2007 [1921–25]: ‘*antarā*’).

<sup>35</sup> This frequency of *saddhiṃ* supports the argument that it may be one of the origins of Hindi *se*, as proposed in Ch. 3, fn. 5. Elements that grammaticalize are known to be frequent before grammaticalization sets in.

<sup>36</sup> Some authors note that, semantically, *saddhiṃ* does not seem a straightforward continuation of *sadhrīm* (cp. Davids and Stede 2007 [1921–25]: *saddhiṃ*). Monier-Williams (1899) lists a usage of *sadhrī* with a short *-i* as ‘with’ for Classical Sanskrit, which may have had an influence.

Besides pseudo-nominal *UPARI*, *antarā*, and *saddhiṃ*, another element, *anto*, also only starts taking genitives in post-Rigvedic times, but its phonological shape does not evoke some local case form. Compare examples from Vedic and Pali (the latter not unambiguous; see clearer Skt. examples in Böhtlingk and Roth ‘antar’):

- (111) VEDIC *ántár*  
*ṛtásya* *gopá* *ná* *dádbhāya* *sukrátus* *trí*  
 truth.GEN.SG.N cowherd.NOM.SG.M NEG deceive.INF.DAT wise.NOM.SG.M three  
*śá* *pavitrā* *hṛdy* *àntár* *á* *dadhe*  
 DEM.NOM.SG.M strainer.ACC.PL.N heart.LOC.SG.N inside to put.PRF.MID.3SG  
 ‘The wise cowherd of the truth cannot be deceived, he carries three strainers  
 in (his) heart.’ (RV 9.73.8)

- (112) PALI *anto*  
*phalaṃ* *dadeyya,* *evam* *etissā* *santatiyā*  
 fruit.ACC.SG.N give.PRS.OPT.3SG SO DEM.GEN.SG.F continuation.GEN.SG.F  
*atthi anto*  
 be.3SG inside  
 ‘(and then, after the shoot has grown, in due time having reached maturity . . . ,  
 (he)) would give the fruit (to the ground), by way of continuation (of it)’  
 (Mili, p. 51, translation Davids)

One may speculate whether the fact that reanalysed *antarā* is based on *anto* (Vedic *antar*) was a factor in the nominal interpretation of *anto*.<sup>37</sup> Note also that *anto* is very rare in Pali, so that its syntax as a zero-valent adverb may not any longer have been as entrenched. This, in turn, may have allowed for a reanalysis as a relational form with a dependent genitive despite the lack of any nominal appearance. In any case, *anto* can be regarded as the most extreme case of reanalysis in the context of the genitive shift, as it undergoes the shift without having either nominal morphology or a pseudo-nominal sound shape.

In sum, we have seen in this section and in the preceding one that relational nouns proliferate in post-Vedic times, and adjectives, participles, and adverbs with and even without nominal morphology follow suit and acquire slots for genitives. Besides the structural level of reanalysis, this shift is a shift in usage. Thus, while relational nouns such as *MADHYE* had always been able to take genitives, over time they occur in this constellation more and more often, and develop into frequently used periphrastic expressions of case. Similarly, old zero-valent adverbs fall out of use, while nominalized forms based on them occur more and more often. For instance in Pali, while

<sup>37</sup> A further morphological connection with a pseudo-nominal is drawn by Grassmann (1996 [1875]: ‘ántár’) who posits an original form *antari* on the basis of compounds such as *antarikṣa* ‘atmosphere’ in parallel with the sound shape of *upari*.

*puro* has vanished, *purato* and *puratthato* are regularly employed. While *adho* (Vedic *adhas*) as a simple adverb still occurs, but rarely so, *hetṭhā* with genitives is more frequently attested. While *anto* is rare, *antarā* with genitives is common. While zero-valent UPARI slowly vanishes, reanalysed UPARI with genitives becomes a frequent construction in the course of Middle Indic, and so on. To put it in Fahs' words: 'The aversion against prepositions in Pali shows in the transformation of a preposition [i.e. referring to *anto*] into an auxiliary noun [*'Hilfsnomen'*].' (1989: 103).<sup>38</sup>

### 5.3.3 Reanalysis of local case forms

In the following two sections, I will show contexts that are ambiguous on both a syntactic and a semantic level, which allowed for the zero-valent adverbs lacking any nominal morphology, introduced in the preceding section, to acquire a slot for genitives. I will first discuss patterns of syncretism between local case forms and genitive forms that yield syntactic ambiguity between adverbial groups and relational noun expressions (section 5.3.3.1). Then I will turn to various semantic constellations of the combinations in question. Only some of these allow for a reinterpretation of the syntagms as containing a relational element with a semantic slot for a spatial ground (section 5.3.3.2).

**5.3.3.1 Syncretism** In Harris and Campbell's words, reanalysis is 'a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any immediate or intrinsic modification of its surface manifestation' (1995: 61). I will propose the assignment of such a new analysis to adverbial groups which are reinterpreted as relational noun expressions in the context of the post-Vedic genitive shift. I show that one precondition for reanalysis was syncretism, which blurred the distinction between local case forms and genitive forms.

The Old Indo-Aryan seven cases (plus the vocative) are morphologically distinct with a separate form per case only in singular *a*-stem masculine stems. In other declensions, two or three cases may have the same form. Syncretisms accrue in the transition to Middle Indo-Aryan. In Pali, the dual has virtually vanished, and the genitive and dative are typically expressed by a single form. Importantly for the present discussion, the genitive-dative shares a form with local cases in several declensional classes. This pertains especially to all of the major feminine singular declensions, which have just one form not only for the genitive and dative, but also for the instrumental, ablative, and locative (with some cases having one or more

<sup>38</sup> 'Die Abneigung gegenüber Präpositionen im Pāli wird daran deutlich, dass eine Präposition zum Hilfsnomen umgeformt wird.'

TABLE 5.3. The case paradigm of feminine *ā*-stems in Classical Sanskrit and Pali

	Classical Sanskrit	Pali
NOM	<i>kanyā</i> ‘girl’	<i>kaññā</i>
ACC	<i>kanyām</i>	<i>kaññam</i>
INS	<i>kanyayā</i>	<i>kaññāya</i> ( <i>kaññā</i> )
DAT	<i>kanyāyai</i>	<i>kaññāya</i>
ABL	<i>kanyāyāḥ</i>	<i>kaññāya</i> , <i>kaññato</i> , <i>kaññāto</i>
GEN	<i>kanyāyāḥ</i>	<i>kaññāya</i>
LOC	<i>kanyāyām</i>	<i>kaññāya</i> , <i>kaññāyaṃ</i> ( <i>kaññāye</i> )

separate secondary forms).<sup>39</sup> For illustration, see the singular paradigm of feminine *ā*-stems in Sanskrit and Pali, the most frequent feminine stem class, in Table 5.3. While in Old Indo-Aryan, only ablative and genitive share a form, now all local cases may be expressed by the single form *kaññāya*, the paradigm being effectively reduced to three distinct forms (cp. Fahs 1989: 51–2 and Oberlies 2001: 141; forms in brackets are rare).

Because of such extensive case syncretism, it sometimes cannot be determined, even in Old Indo-Aryan, whether for instance MADHYE is combining with a local case form or with a genitive. In accordance with the spread of syncretic forms, such ambiguous cases proliferate in the transition to Middle Indo-Aryan. Compare the following ambiguous examples with MADHYE from Vedic Sanskrit and Pali:

## (113) VEDIC SANSKRIT

*mádhye vasiṣva tuvinṛmṇorvór*

**mádhye** vasiṣva tuvinṛmṇa ūrvór

MADHYE wear.IMP.2SG very\_valient.VOC.SG.M thigh.GEN/LOC.DU.M

*ní dāsāṃ śíśnatho háthaiḥ*

down demon.ACC.SG.M strike.2SG blow.INS.PL.M

‘Cover yourself **between the thighs**, oh very valiant one, strike down the demon with blows.’ (RV 8.70.10)

## (114) PALI

*saṃviggaraṭṭho*

*parisāya*

*majjhe*

in\_state\_of\_fear.NOM.SG.M group.GEN/LOC.SG.F MADHYE

‘(He was) frightened **in the middle of the assembly**.’ (Jā III, p. 348)

The same type of ambiguity can be found for those adverbs which started off as zero-valent and at some point acquired a slot for a genitive (*avidūre*, *param*, UPARI,

<sup>39</sup> There are a few exceptions: for instance *dhitar* ‘daughter’ has a separate locative form (see Fahs 1989).

*antarā*, *saddhiṃ*, and *anto*).<sup>40</sup> I propose that it was such ambiguous examples which made the reanalysis possible, aligning the adverbs to the syntax of the proliferating relational noun expressions by acquiring a slot for a genitive dependent. Compare the following examples from Pali which involve syncretic forms:

(115) GEN/ABL + *avidūre*

<i>tena kho pana</i>	<i>samayena</i>	<i>aññatarā</i>	<i>bhikkhuni</i>
DEM.INS.SG even but	time.INS.SG	certain.NOM.SG.F	Buddhist_nun.NOM.SG.F
<i>jarādubbala</i>	<i>caraṇagilānā</i>	<i>Sundarīnandāya</i>	
decay_weak.NOM.SG.F	walking_sick.NOM.SG.F	Sundarīnandā.GEN/ABL.SG.F	
<i>bhikkhuniyā</i>	<i>avidūre</i>	<i>nipannā</i>	<i>hoti</i>
Buddhist_nun.GEN/ABL.SG.F	not_far.LOC.SG	sit_down.NOM.SG.F	be.3SG

‘So, at some point, a certain nun who was weakened by decay and sick from walking sat down **not far from the nun S.**’ (Vina IV, p. 212)

(116) GEN/LOC + UPARI

<i>mahantassa</i>	<i>phalakaṃ</i>	<i>soṇḍāya</i>	<i>upanibandhati,</i>
great.GEN.SG	shield.ACC.SG	trunk.LOC.SG.F	tie.3SG,
<i>tomarahattho</i>	<i>ca</i>	<i>puriso</i>	<i>upari</i> <i>gīvāya</i>
holding_a_lance.NOM.SG.M	and	man.NOM.SG.M	UPARI neck.GEN/LOC.SG.F
<i>nisinno</i>	<i>hoti</i>		
sit_down.PPP.NOM.SG.M	be.3SG		

‘He ties a shield to the great beast’s trunk, a man holding a lance is sitting on top of his [an elephant’s, UR] neck.’ (Majjh, p. III 133)

Evidence that ambiguous examples of this kind made the reanalysis possible also comes from apparent free variation in attested versions of a single text. For instance, example (108) above contains *antarā* in combination with *satthināṃ* (GEN.PL), which was selected for the text edition. However, other versions of the text have *satthī* (ACC.PL) and *sattimhi* (LOC.SG) (cp. Woodward’s comment in his translation). Such variability is indicative of a diminishing functional distinction between the different cases, making a reanalysis easily possible.

Regarding relational nouns, the distribution was already skewed towards a propensity for taking genitives (see Table 5.1 for MADHYE); so it is likely that syncretic forms accompanying relational forms would be interpreted as genitives by default. Now, note that many of the reanalysed nominal and adverbial forms presented in sections 5.3.1 and 5.3.2 are still comparatively infrequent in Pali in comparison with the relational nouns proper. For instance, while MADHYE occurs 118 times (without

<sup>40</sup> The situation is different for the extended forms *pacchato*, *purato*, etc. Their adverbial base forms vanished, and instead these new forms emerged. In accordance with their extension with a nominal case suffix, they construe with genitives from the beginning.

compounds, repetitions, etc.), UPARI is attested only thirty-six times in the *Tipiṭaka*. Being a low-frequency form by comparison, its syntax would be less entrenched. Therefore, when co-occurring with a syncretic case form, it would easily be reanalysed as a nominal head with a dependent genitive on the model of the much more frequent relational noun expressions.

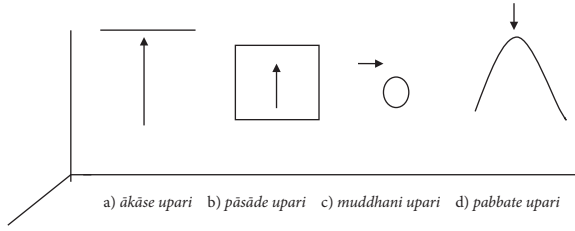
5.3.3.2 *Gaining ground: The semantic changes of zero-valent adverbs undergoing the genitive shift* In this section, I focus on the semantic preconditions for reanalyses of adverbial groups as relational noun constructions. Compare at first the following examples of UPARI in combination with local case forms in Pali:

- (117) *upari bhikkhave ākāse verambā nāma*  
 UPARI monk.VOC.PL sky.LOC.SG Veramba.NOM.PL named  
*vātā vāyanti*  
 wind.NOM.PL blow.3PL  
 ‘Oh brothers, in the sky above winds blow which are called Veramba.’  
 (Saṃ II, p. 231)

- (118) *caṅkī brāmaṇo upari pāsāde ...*  
 Caṅkī.NOM.SG.M Brahmin.NOM.SG.M UPARI palace.LOC.SG  
*upagato hoti*  
 lie\_down.PPP.NOM.SG.M be.3SG  
 ‘The brahman Caṅkī was lying down upstairs, in the palace.’  
 (Majjh II, p. 164)

- (119) *atha kho Mucalindo nāgarājā ... bhagavato*  
 now PRT Mucalinda.NOM.SG snake\_rāja.NOM.SG Exalted\_one.GEN.SG  
*kāyaṃ sattakkhattuṃ bhogehi parikkhipitvā*  
 body.ACC.SG seven\_times coil\_of\_a\_snake.INS.PL encircle.CONV  
*upari muddhani mahantaṃ phaṇaṃ vihaṇṇa*  
 UPARI head.LOC great.ACC.SG hood.ACC.SG strike.CONV  
 ‘So Mucalinda, the king of snakes ... having encircled the body of the Exalted One seven times with his coils and having struck his great hood above, over the Exalted One’s head ...’  
 (Ud, p. 10)

- (120) *seyyathāpi bhikkave upari pabbate thullaphusitakesu*  
 just\_as brother.VOC.PL UPARI mountain.LOC.SG rain-god.LOC.PL  
*devesu vassante*  
 god.LOC.PL rain.MID.3PL  
 ‘Just as, brothers, when above, on the mountain, the rain-gods rain.’  
 (Saṃ V, p. 396)

FIGURE 5.1 Image schemas of CASE<sub>loc</sub> + UPARI<sup>41</sup>

The locative case, which UPARI is construed with here,<sup>42</sup> is semantically what has been called a ‘general locative’ (e.g. Levinson, Meira, and The Language & Cognition Group 2003). It expresses an underspecified localization roughly at some place which in English would have to be rendered by one of a choice of prepositions, *in*, *at*, *on*, etc. This semantic breadth can be gleaned from the examples above where locatives co-occur with UPARI which in all the examples encodes a space ABOVE.<sup>43</sup> While in (117), *upari* and *ākāse* independently refer to the upper region of the three-dimensional world, in (118) *upari* refers to the upper area (i.e. the highest floor) *within* a space encoded by the locative (a palace).<sup>44</sup> In (119) and (120), *upari* refers to the space ABOVE in relation to something else expressed as a locative, an object (a snake) in the first example, and a surface (a mountain) in the second example, that is without or with contact. These different image schemas can roughly be depicted as in Figure 5.1.

Now consider the following examples (121) and (122) with genitives in Pali:

- (121) *Saccakassa Nigaṇṭhaputtassa upari vehāsaṃ*  
 Saccaka.GEN.SG son\_of\_Jains.GEN.SG UPARI air.ACC.SG  
*ṭhito hoti*  
 stand.PPP.NOM.SG.M be.3SG  
 ‘The Yakṣa came to stand in the air above Saccaka, the son of Jains.’  
 (Majjh I, p. 231)

<sup>41</sup> For readability, all examples in Fig. 5.1 are given in the order: CASE<sub>loc</sub> UPARI.

<sup>42</sup> This is the most frequent local case that UPARI construes with in Pali: see Table 5.2.

<sup>43</sup> I use the shorthand ABOVE for the period when UPARI still identifies a region (or complex spatial relation) rather than a simple spatial relation ON or OVER. Here, in Pali, UPARI is always used with reference to a region that is located at a high place along a vertical axis, as will be shown. Before Hindi it does not mark simple relations, e.g. usages such as ‘on the floor’ where relation to a point on the ground is made without elevation along a vertical axis. This latter shift will be outlined in section 7.4.2.

<sup>44</sup> One could also construe the first example as UPARI referring to the upper area *within* the sky if it is interpreted as a three-dimensional space itself. However, the context makes clear that UPARI does not specify the upper region within the sky, but rather simply means *above* in a general construal of a three-dimensional space, which in this case coincides with reference to the sky.

(122) (= ex. (96))

*seyyathāpi* [...] *payasotattassa nibbāyamānassa upari santānakaṃ*  
 just\_as boiled\_milk.GEN.SG cool.PPM.GEN.SG UPARI scum.NOM.SG  
*hoti evam evaṃ pātur ahosi*  
 become.3SG just so manifest be.AOR.3SG  
 ‘Even as scum forms **on top of boiled milk that is cooling**, so did (the earth)  
 appear.’ (Dīgh III, p. 85)

In terms of image schemas, these two examples roughly parallel the previous ones in (119) and (120). As the snake above the head in example (119), in (121) a *yakṣa* (type of semi-divine being) is located above Saccaka. Similarly, just as the surface of a mountain is evoked in (120), scum forms on the surface of boiled milk in (122). Now, *only* these two types of conceptualizations, of an object located ABOVE in relation to some other entity, without or with contact, occur with relational noun expressions in my corpus. By contrast, conceptualizations as in examples (117) and (118), where UPARI simply refers to a region high up in a three-dimensional space without reference to another object below, are not found in usages of UPARI when taking genitives.

My proposal is that it is not so much the case that UPARI itself has a different meaning in the two combinations. Rather, the meaning of GEN + UPARI differs from CASE<sub>loc</sub> + UPARI because the genitive only allows for a restricted set of interpretations.<sup>45</sup> While UPARI as a zero-valent element simply makes reference to an upper area, the reanalysis of the local case form as a genitive is mirrored on the semantic side in the form of an argument slot for a ground. The semantic reanalysis is presented as a transition from the image schema on the left to the image schema on the right in Figure 5.2.

On the left, we see the figure (F) located in the upper area in relation to whichever spatial reference frame is set by the context, whether an object on the ground, a

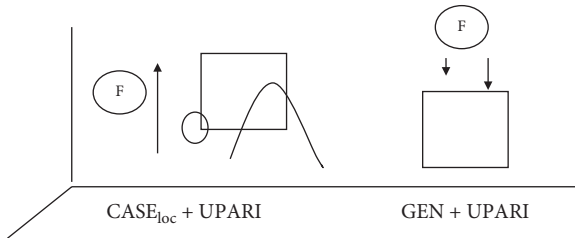


FIGURE 5.2 Image schemas of CASE<sub>loc</sub> + UPARI and GEN + UPARI

<sup>45</sup> Note that this does not have anything to do with grammaticalization, which would always involve a shift in the meaning of the grammaticalizing element (i.e. UPARI). At this point, we are dealing with a reanalysis of UPARI and other adverbs as relational elements. Grammaticalization sets in later, affecting reanalysed GEN UPARI, the individual sub-processes of which will be explored in Chapter 7.



surface, or a three-dimensional frame. As a result of the combination, a pragmatic reading of ABOVE in relation to something else below, or ABOVE in relation to some spatial frame emerges. In the schema on the right, by contrast, UPARI has a semantic slot for a ground in relation to which it localizes something as being in some way ABOVE. The two arrows in this schema mark the fact that localization without contact or with contact in relation to some entity or surface is possible.

Note the important difference in construal between the combination with local case forms and the relational noun expression. While both *pabbate* (LOC) *upari* ('above, in/on/at the mountain') and *payasotattassa* (GEN) *upari* ('on top of boiled milk') yield roughly the same image schema of a localization on some surface, in the combination with local cases it is pragmatic inference which has us construe this image schema. In the relational noun expression, by contrast, it is in the semantics of the expression itself that UPARI locates something as being ABOVE in relation to some ground, without or with contact.<sup>46</sup>

Now, I propose that the transition from adverbial groups to relational noun expressions was possible precisely in such cases where not only was the case form ambiguous, but the combination of UPARI and local case form would pragmatically allow for the inference of a location of something as 'above' or 'on top of' something else, as in examples (119) and (120). By contrast, even when there is syntactic ambiguity, image schemas such as in (117) or (118) would not lead to a reinterpretation, lacking semantic ambiguity ('above, in the sky' vs. 'above, on top of the sky'). On the other hand, solely semantic ambiguity without accompanying formal ambiguity does not allow for reanalysis either. The following passage attests to the existence of examples that display both formal and semantic ambiguity, the combination of which allows for reanalysis. I claim that this type of example marks the threshold between the two alternative syntactic analyses as either an adverbial group or a relational noun expression. Here, 'above, at/on/by the neck (LOC) (of an elephant)' by inference can be construed as simply 'on top of the neck':

- (123) GEN/LOC + UPARI (= ex. (116))
- |                          |                 |                |                         |
|--------------------------|-----------------|----------------|-------------------------|
| <i>mahantassa</i>        | <i>phalakam</i> | <i>soṇḍāya</i> | <i>upanibandhati,</i>   |
| great.GEN.SG             | shield.ACC.SG   | trunk.LOC.SG.F | tie.3SG                 |
| <i>tomarahattho</i>      | <i>ca</i>       | <i>puriso</i>  | <i>upari gīvāya</i>     |
| holding_a_lance.NOM.SG.M | and             | man.NOM.SG.M   | UPARI neck.GEN/LOC.SG.F |

<sup>46</sup> While a great deal has been published on metaphorical, metonymic, and other types of semantic shift involved in grammaticalization processes, semantic shifts involved in a reanalysis yielding a change in syntactic class as depicted here, which is not connected to grammaticalization, seem to have been little explored in the literature. Therefore, I rely here solely on the semantic evidence apparent in the data. The rich grammaticalization literature on spatial terms will be an important foundation for my examination of the relation between semantic and formal processes in the grammaticalization of the postpositional phrases in Chapter 7.

*nisinno* *hoti*  
 sit\_down.PPP.NOM.SG.M be.3SG

‘He ties a shield to the great beast’s trunk, a man holding a lance is sitting **on top of his** [an elephant’s, UR] **neck**.’ (Majjh III, p. 133)

The semantic shift to a relational structure undergone by UPARI is paralleled by the other elements which acquire a slot for a genitive. For instance, the extended form *heṭṭhā* mirrors UPARI in that it shifts from meaning ‘below’ as Vedic zero-valent *adhas* to ‘under’ as *adhastāt* (Pali *heṭṭhā*), where the genitive adds a semantic slot for a ground. While *puras* as a simple adverb expresses ‘in front’, the extended forms *purato* and *puratthato* mean ‘before’ with a spatial or temporal ground. While *paścā* meant ‘later’, *pacchato* in Pali expresses ‘after (a certain point in time)’. The pseudo-nominal *antarā* when combining with an accusative encodes a relation of ‘between two points’ (Fahs 1989: 106), the meaning of endpoints reached being inherent in the accusative semantics.<sup>47</sup> However, with genitives it expresses a general relation of ‘between’ without the specific semantic contribution of the accusative, but instead in relation to an intrinsic ground. Note that not only zero-valent adverbs show such a shift in meaning. Relational nouns differ in the same way in that, with local case forms, they may combine with all sorts of semantically independent spatial relations, whereas a genitive always simply specifies the ground. Compare the following example with MADHYE, where *barhis* (*barhir*) does not encode the ground for MADHYE (the ground instead remains unexpressed) but a goal, one of the accusative’s functions:

- (124) *ní* *ṣatsi* *mádhyā* *ā* *barhír*  
 down sit.IMP.2SG MADHYE LP sacrificial\_grass.ACC.SG.N  
*ūtāye* *yajatra*  
 support.DAT.SG.F worthy\_of\_worship.VOC.SG.M  
 ‘Sit down in the middle, on the barhis, for solace, oh worthy one of worship!’ (R̥V 3.14.2)

## 5.4 Semantic evidence

Note that, notwithstanding the scenario I have just outlined, which accounts for the developments in Indo-Aryan, the reanalysis of an adverbial group as an adpositional phrase (without an intermediate step ‘via’ relational noun expressions) is an attested possibility. It is this precise development which turned the local particles into prepositions in western branches of Indo-European, as shown in the previous chapter. Also, we know that cognates of UPARI, such as English *over*, German *über*, or French *sur* (in the same way as the core PIE local particles) were reanalysed in

<sup>47</sup> With verbs of movement, accusatives may generally encode goals.

groups with local case forms, and like Hindi *par* express a localization of something above something else, for example *the picture over the door* or *put the jacket over the chair*. Examples (119) and (120) demonstrate that such pragmatic interpretations are also possible in the case of UPARI when combining with local case forms. So, in view of the fact that both scenarios are logical possibilities and both types of combinations with local case forms and with genitives are attested until Apabhramsha (see Table 5.2), does this return us to the beginning of this chapter, wondering about the true source constructions of the New Indo-Aryan postpositional phrases? In this section, I point out fine semantic properties of OBL *mē/par* as a final piece of evidence that the source constructions must indeed be the relational noun expressions.

First of all, while UPARI expresses ABOVE with and without contact, its cognate forms English *over* and German *über* mean ABOVE in the first sense (left arrow in the image schema on the right in Figure 5.2):<sup>48</sup> the general relation that they also expressed with local case forms at earlier times (English has *on* and German has *auf* to express vertical localization with contact).<sup>49</sup> I outlined that UPARI in combination with local case forms can be construed as ABOVE without or with contact. However, I argued that this is a product of pragmatic inference. By itself, UPARI denotes simply a region that is ABOVE in a three-dimensional space. This matches the semantic structure of *over* and *über*, but it does not match the semantics of *par*. *Par* encodes a relation of contact with some object or surface in contemporary Hindi, a difference that can be traced to the source construction involving a reference frame encoded as a genitive, above which something is situated, whether with or without contact (see section 5.3.3.2).

Another way in which English *over* and German *über* continue usages in combination with local case forms in contrast with Hindi *par* is the expression of a traversal sense (combining PATH and elevation along a vertical axis) as in *walk over a bridge*, *jump over a fence*, or *über eine Mauer klettern* ‘climb over a wall’. This usage is found in ancient varieties of Indo-European when in combination with accusatives (see Hofmann 1965: 281 for Latin). Compare the following examples from Old Indo-Aryan, Old English, and Latin:

(125) Epic Sanskrit

*upary upari gacchantah śailarājam*

UPARI UPARI go.PPA.NOM.PL.M king\_of\_mountains.ACC.SG.M

‘They are going across the great mountain.’ (MBh 1.111.5e)<sup>50</sup>

<sup>48</sup> While French *sur* in the same way as Hindi *par* means ‘above’ or ‘on’, *sur* only partially represents a cognate of *par*, as it descends not only from Latin *super*, but from a merger, formally and functionally, of *super* with *supra* and *sūsum* (see De Mulder and Vanderheyden 2002).

<sup>49</sup> See Brugman (1983) and Lundskær-Nielsen (1993) for English *over*, and Bouillon (1978) and Müller (2013) for German *über*.

<sup>50</sup> Cited after the critical edition by Sukthankar provided online at <http://titus.uni-frankfurt.de/texte/etcs/ind/aind/mbh/mbh.htm> (last accessed 13 October 2013).

## (126) OLD ENGLISH

*ðylæs he ofer ðone ðerscold his*  
 lest 3SG over DEM.ACC.SG.M threshold GEN.3SG  
*endebyrdnesse stæppe*  
 authority.GEN.SG.F step.SUBJ.3SG

'lest he should step **over the threshold** of his authority'

(Old English version of Gregory's *Pastoral Care*, edition  
 Sweet 1871: 76; cited in the *OED*)

## (127) LATIN

*volat super impetus undas*  
 fly.3SG above momentum.NOM.SG.M wave.ACC.PL.F  
 'The momentum flies **over the waves**.'

(Ennius ann. 386, cited after Hofmann 1965: 281)

To this day, *over* and *über* conserve these usages. UPARI, by contrast, while it also figured in such constructions with accusatives as seen in example (125), does not encode such a traversal sense in constructions with genitives in a single case in my corpus. In Hindi, this sense is expressed by secondary postpositional phrases *ke ūpar* (another younger, nominalized usage of UPARI) and *ke pār* (from Skt. *pāraḥ* 'the other side/riverbank on the other side' from *para-* 'far, beyond,' cp. *param* in section 5.3.1).

Vice versa, there are usages of GEN + UPARI that are only found in this construction and not in groups with local case forms. One such case, which shows an expansion into an abstract domain, is the expression of a feeling for someone or something, a usage already attested in Classical Sanskrit and Pali (cp. also Böhtlingk and Roth 1855–75: 'upari'). Compare the following examples from Sanskrit, Apabhramsha, and Hindi:

## (128) CLASSICAL SANSKRIT

*putrasyopari kruddhaḥ*  
*putrasya upari kruddhaḥ*  
 son.GEN.SG.M UPARI get\_angry.PPP.NOM.SG.M  
 'angry with the son'

(Hitopadeśa 66.17, cited after Böhtlingk and Roth 1855–75 'upari')

## (129) APABHRAMSHA

*mā rūsaḥi ammaho uvari sāmī*  
 NEG be\_angry.IMP.2SG GEN.1SG UPARI lord.VOC.SG.M  
 'Do not be angry with me, my lord!'

(KA 3.12.10)

## (130) HINDI

*vah us par gussā hai*  
 DEM.DIR.SG DEM.OBL.SG on angry.DIR.SG.M be.3SG  
 'He is angry **with him**.'

Thus, whereas cognates of UPARI such as English *over* and German *über* only continue functions attested at earlier times in constructions with local case forms, Hindi *par* continues usages that are only attested in constructions with genitives. I take this as a final piece of evidence that relational noun constructions are indeed at the origin of postpositional phrases in Hindi, contrary to what is suggested by evidence of UPARI in Vedic or by the developments in other branches of Indo-European, and contrary to what has been claimed in the literature.

## 5.5 Symmetrical and asymmetrical groups

In section 2.5 and in Chapter 4, I outlined a scenario that depicts the development of phrases in non-configurational languages. This scenario is based on the idea that co-ranking and syntactically independent but semantically co-referential elements are at some stage analysed as a single syntactic unit, that is as a phrase. Examples are adverbial groups of demonstratives and nouns that are reanalysed as nominal phrases with the demonstratives turning into articles, or adverbial groups of local particles and local case forms being reanalysed as adpositional phrases, as happened in many branches of Indo-European.

I have argued in this chapter that Indo-Aryan deviates from other branches of Indo-European and from this scenario because adpositional phrases do not develop from such symmetrical groups. Instead, I propose that Indo-Aryan postpositional phrases descend from relational noun expressions, with zero-valent adverbs like UPARI aligning to the syntax of relational nouns by acquiring a slot for a genitive dependent. These source constructions do not qualify as symmetrical groups because they show formal and semantic asymmetries, rather than consisting of syntactically independent, co-ranking, and semantically co-referential elements. Formally, the relational noun can occur on its own, while the adnominal genitive cannot. Semantically, the relational noun has an argument slot for a possessor.

Note that relational noun expressions are cross-linguistically well-known sources of adpositional phrases (e.g. Kahr 1975; Starosta 1985). However, so far they have been identified as sources in cases where nominal expressions are already organized phrasally, such as in the more recent grammaticalization phenomena in English or German, for example *on top of*, *by virtue of*, *im Rücken von* ‘at the back of’, etc. I therefore propose an extension of Himmelmann’s model, namely the distinction between the rise of phrases in strongly non-configurational languages from symmetrical groups on the one hand (corresponding to Himmelmann’s groups, consisting, for example, of demonstratives and nouns, or adverbs and local case forms), and asymmetrical groups on the other hand (e.g. relational noun expressions). Symmetrical groups consist of co-ranking, syntactically independent, and semantically co-referential elements. Asymmetric groups, by contrast, are constructions where

one element or construction depends on another one semantically and possibly morphologically.

Besides relational noun expressions, there are also other cases where a phrase arises from an asymmetrical group. Among the Hindi postpositional phrases, the possessive marker *kā/-e/-ī*, for instance, descends from a participle with a dependent object (see section 3.6), also an asymmetrical group. Another example is the combination of an inflected verb which grammaticalizes into an auxiliary in construction with an infinitive form. For instance, the French future descending from a combination *cantare habeo* represents an asymmetrical group with the infinitive depending on the finite verb form (see too the quotation from Hopper in section 2.6).<sup>51</sup>

Thus, we can identify not one structural pathway to phrase structure in strongly non-configurational languages, but two. This has implications for the study of the historical syntax of Indo-European. Recall the widespread assumption that adpositional phrases in Indo-European generally originate in adverbial groups (see section 4.1). While this is correct for the majority of branches, Indo-Aryan limits the scope of this scenario. In this branch, there is no etymological connection between the local particles and the modern postpositions, nor a connection on a categorical level, as shown in the previous chapter. Instead, the genesis of adpositional phrases in the respective branches progresses along an unrelated pathway, taking off in asymmetric groups. As pointed out in section 1.1, this calls into question push-chain accounts for the change from synthetic to analytic strategies of case marking, since the local case forms never become functionally redundant inside newly formed adpositional phrases (see also Ch. 5, fn. 19).

I would like to round off this section with a remark on the types of phenomenon subsumed under ‘grammaticalization’. In section 2.1, I stressed that it is not individual lexemes that are being grammaticalized, but lexemes in constructions. The development of both symmetrical groups and asymmetrical groups into adpositional phrases offers insights into the types of construction we are dealing with. Once more, the reanalysis of symmetrical groups is a reanalysis of syntactically unconnected elements into a syntactic unit. Before reanalysis, only semantic co-reference and possibly parallel morphological marking indicated that the elements in question belong together. Note that these are relations between the elements and the referent, but not among the elements themselves in the sense of belonging together as a formal and functional unit. Thus, the reanalysis of an asymmetrical group as a phrase is concomitant to the birth of a construction.<sup>52</sup> In asymmetric groups, by contrast, the

<sup>51</sup> While Latin already had phrasally organized structures at this point, this pertains to nominal expressions but does not extend to verbal expressions. Thus, *cantare habeo* is another example of a construction united by semantic and morphological dependency relations developing into a syntactic phrase.

<sup>52</sup> As a consequence, while it may be legitimate to claim that Hindi *mē* descends from Vedic *MADHYE*, it is strictly speaking incorrect to derive Hindi *par* from Vedic *upari*. The reason is that Vedic *upari* did not

elements are already necessarily connected by semantic and possibly morphological dependency relations. They may not yet form a phrasal unit, but they are already a construction through these dependency relations. The reanalysis of such a construction as a phrasal unit can be seen as a syntactization of already existing semantic (and morphological) dependency relations.

I hypothesize that the difference between a reanalysis of symmetrical groups and of asymmetrical groups also shows differences on further levels beyond the question of the structural relations between the elements involved. I contrast below the reanalysis of adverbial groups in western branches of Indo-European with the reanalysis of MADHYE and UPARI as in Indo-Aryan, but the differences should also apply to other symmetric and asymmetric groups.

It is generally assumed that frequency plays an important role in the reanalysis of symmetrical groups as phrasal units (e.g. Luraghi 2010: 217). Only as a result of frequent recurrence does the syntagm receive a new analysis as a syntactic unit (see section 2.6 on the aspect of ritualization and entrenchment in grammaticalization phenomena). The reanalysis entailed by such frequent recurrence, however, is not (or not necessarily) accompanied by a semantic change pertaining to the gram, but only involves the local particle acquiring an argument slot. The difference can be captured in many cases simply as the addition of an argument slot for a ground, rather than as an intrinsic semantic shift of the local particle. Compare once more different usages of Homeric Greek *epi* ‘on, to, near’ (see section 4.1):

(131) (= ex. (63))

<i>éluth'</i>	<i>épi</i>	<i>psukhè</i>	<i>Agamémnonos</i>	
come.AOR.3SG	EPI	soul.NOM	Agamemnon.GEN	
‘The soul of Agamemnon approached.’				(Odyssey 24.20)

(132) (= ex. (64))

<i>epì</i>	<i>gaîan</i>	<i>ap'</i>	<i>ouranóthen</i>	<i>protrápētai</i>	
EPI	earth.ACC	from	heaven	turn.3SG	
‘To earth from heaven (the sun) turns.’					(Odyssey 11.18)

An adverbial meaning ‘towards, near, on’ when combined with an accusative yields an interpretation of a prepositional ‘to’. There appears to be no major semantic change apart from the fact that *epi* as a zero-valent adverb encodes a spatial location or direction, whereas it encodes a location or direction with regard to some other spatial entity when combining with a local case form.<sup>53</sup>

have a slot for genitive dependents. Instead, it is reanalysed UPARI which represents the precursor form of Hindi *par* (see Ch. 8, fn. 16).

<sup>53</sup> Certainly, lexicalized usages may evolve (see the examples in section 4.1), but this is not a necessary change, and it may set in before or after reanalysis as an adpositional phrase.

The reanalysis of asymmetric groups as adpositional phrases, by contrast, crucially involves semantic shifts that affect the gram's meaning. I will show in Chapter 7 how semantic shifts—namely metaphorical and metonymic extensions of GEN + MADHYE/UPARI—entail a syntactic fixation of the asymmetric groups as phrases. In particular, the shift of MADHYE from referring to a spatial area to expressing a purely spatial relation will be shown to entail the development of phrasality. While the details of how these changes come about have to wait until Chapter 7, the contrast between what is barely a semantic shift, in the case of the local particles, and the significant metaphorical and metonymic extensions in the case of the constructions in Indo-Aryan, is manifest.

Thus, in terms of grammaticalization theory, when it comes to the reanalysis of symmetrical groups, we are dealing with the birth of constructions at the same time. Only then do grammaticalization processes take off, applying to the constructional unit and affecting the meaning of the grammaticalizing elements. For example, through reanalysis, a local particle turns into a preposition, that is, it has become a grammatical marker. From this point onwards, grammaticalization proceeds; for instance the preposition may develop from a concrete spatial meaning to an abstract one, as in the case of Latin *ad* 'to, at' developing into a marker of recipients and benefactives. In the case of asymmetric groups, there is a construction from the start. What asymmetric groups share with symmetric ones, however, is that only in groups—that is adjacent syntagms—may grammaticalization processes set in.

## 5.6 Conclusion

In this chapter, I went back as far as historical attestations allow and examined the oldest attestations of the modern Hindi postpositions *mē* and *par* in order to identify the main components of the source strings of the modern postpositional constructions. Comparative evidence from other branches of Indo-European, as well as from Vedic Sanskrit itself, suggests that the modern constructions descend from groups of adverbial elements. I showed however, based on Middle Indic evidence, that relational noun expressions are the true sources of the modern postpositional constructions OBL *mē/par* and in this way represent the oldest roots of the modern postpositional constructions of Hindi. Elements such as UPARI, along with a host of other elements that could not be used as heads in relational noun expressions in Vedic, were reanalysed as nominal, allowing for genitive dependents, in the course of what I refer to as the post-Vedic genitive shift. I identified contexts of formal and functional ambiguity that made possible the reanalysis and thus led to an alignment to the syntax of relational noun expressions. In Chapter 8, I will argue that at least *ko* in all probability emerged along the same trajectory. On a structural level, this also applies to OBL *kā/-e/-ī*, which similarly takes off from an asymmetrical group.



As a consequence of these findings, it is necessary to correct the view that configurationality, in Indo-European, arises exclusively from adverbial groups. In order to accommodate the alternative pathway to adpositional constructions and hence configurationality in Indo-Aryan, I introduced a new distinction between origins of phrases in either *symmetrical* or *asymmetrical groups*. Himmelmann's notion of 'group' corresponds to symmetrical groups: combinations of syntactically independent, co-ranking elements in juxtaposition, connected through semantic co-reference. Asymmetrical groups comprise elements in juxtaposition that are linked by semantic and possibly morphological dependency. Examples are relational nouns with dependent genitives as described in this chapter, participles with objects, or finite verb forms with dependent infinitives. From the perspective of grammaticalization theory, these two types of groups differ in that the reanalysis of symmetrical groups as syntactic units is concomitant with the birth of constructions, while asymmetrical groups are already constructions in the sense of functional and formal units. To conclude, the insight gained here that not only symmetrical but also asymmetrical groups can stand at the origin of the development of configurational syntax widens not only our understanding specifically of the historical syntax of Indo-European, but also our general understanding of configurationality as a historical, emergent phenomenon.



## The origin of the postpositional syntagm

This chapter, together with the next, explores the question of how the components identified in Chapter 5 ended up as grammaticalizing constructions. I here focus on the emergence of the postpositional syntagm. In the literature, there is an ongoing debate about whether language-internal or language-external factors are responsible for New Indo-Aryan having postpositions instead of prepositions like other branches of Indo-European (section 6.1). I will first outline the different proposals (section 6.1.1) and subsequently turn to an evaluation of the claims. The different proposals rest on the assumption—despite basic free word order—of an OV syntax from early on, resulting in the expected development of postpositions. I will, however, show that the data in my corpus paint a somewhat different picture (section 6.1.2). Because the traditional approaches to the topic only take us so far, I take a step back and concentrate on the question of juxtaposition in section 6.2. Here, I will explore whether the loss of the Wackernagel position can be argued to be implicated in the emergence of obligatory syntagmatic continuity. However, no direct causal link can be detected. In section 6.3, I focus on the subgroup of relational noun expressions which involve pronominal possessors. In contrast to nominal possessors, these show the pattern GEN N (i.e. GEN MADHYE/UPARI) in a great majority of examples already from Vedic Sanskrit onwards, and may thus be a source of the postpositional syntagm (section 6.3.1).<sup>1</sup> In search of explanations for this difference between the syntax of pronominal and the syntax of nominal possessors, I take a closer look at the demonstrative forms which serve as 3rd person pronominal possessors. The stem-class they belong to shows a general tendency to appear at the left edge of the clause in anaphoric function. I will argue that the leftward tendency of pronominal possessors together with the disappearance of certain types of discontinuity are a crucial factor in the consolidation of GEN<sub>pro</sub> N strings

<sup>1</sup> Henceforth, I subsume UPARI under the syntactic category of nominal heads, it having acquired a slot for genitives in the course of the post-Vedic genitive shift.

(section 6.3.2). The question of how this pattern spreads to constructions with nominal possessors is one of the topics discussed in the next chapter.

## 6.1 Why does Indo-Aryan have postpositions?

Both assumptions, that postpositions are due to language-internal or to language-external factors, can be found in the literature in an equal measure. I first outline the claims (section 6.1.1) and then test them against my data (section 6.1.2).

### 6.1.1 *Language-internal and language-external explanations*

It has been assumed in the literature that New Indo-Aryan has postpositions for the simple reason that Indo-Aryan has and always has had a tendency at least towards OV syntax with the expected harmonic correlates (e.g. Masica 1976). As Vedic Sanskrit is a free word order language, this assumption must be treated with caution, however. While the *Rigveda* is notorious for showing great syntagmatic variation, Vedic prose appears to largely adhere to this order, but in principle also allows otherwise. Certain younger Classical Sanskrit texts again show a much freer order (see Hock 1997 for an overview). The issue is complicated by the fact that early periods coincide with the most artificial language of Mantra Vedic, while prose—itsself not without artifice—is attested in younger layers, which makes it difficult to disentangle ‘chronology and genre’ (Hock 1997). Also, a comprehensive study of order which takes into account information structure is wanting.

In any case, the order of verb and object cannot of course directly ‘explain’ the placement of adpositions. Whatever Greenbergian correlations may predict, explanations for word order patterns must be sought in the history of the constructions themselves. Harmony with OV order, then, often simply stems from the fact that adpositional constructions derive from OV constructions, for example from participial expressions. This line of argument has been taken up for Vedic Sanskrit by Andersen (1979) who argues that it is OV constructions consisting of participles or converbs and their objects which lead to deverbal postpositions (see also Harris and Campbell 1995: 211), and GEN N with relational nouns and adnominal genitives which lead to denominal postpositions. With regard to the latter (I turn to deverbal *kā/-e/-ī* in Chapter 8), Andersen’s proposal agrees with what is generally assumed for Old and Middle Indic, namely that GEN N is the rule with exceptional cases of N GEN (e.g. Delbrück 1878 for Vedic prose, Thommen 1905 for post-Vedic Sanskrit and early Middle Indo-Aryan). The data being skewed towards GEN N, this pattern would at some point have simply become obligatory. The first hypothesis, accordingly, is that the syntagmatic pattern of the modern adpositional constructions arises from an age-old propensity for GEN N.

Others have argued that contact influence is responsible for New Indo-Aryan having postpositions instead of prepositions. Language contact especially between Indo-Aryan and Dravidian, sometimes including Munda languages, has been discussed since the nineteenth century.<sup>2</sup> Today, India (or even the entirety of South Asia) is considered to be one of the empirically best established linguistic areas (e.g. Ebert 2001: 1539). While Dravidian languages borrowed a great deal of Indo-Aryan (especially Sanskrit) vocabulary, they in turn are argued to have impacted on Indo-Aryan at deeper levels of structure, for instance in the domain of phonology where Indo-Aryan retroflexes have been attributed to Dravidian influence.<sup>3</sup> With regard to syntax, Chatterji claimed that Dravidian and Indo-Aryan syntax 'are one', the convergence starting in early Middle Indo-Aryan:<sup>4</sup>

**It is in syntax that Indian Dravidianism and Aryandom are one.** A sentence in a Dravidian language like Tamil or Kannada becomes ordinarily good Bengali or Hindi by substituting Bengali or Hindi equivalents for the Dravidian words and forms, without modifying the word-order, but the same thing is not possible in rendering a Persian or English sentence into a NIA. language. The most fundamental agreements are thus found between NIA. and Dravidian, and all **this began from early MIA.**, as is seen from a comparison of the syntax of Pali and the Prakrits with that of the modern vernaculars. (Chatterji 1970 [1926]: 176–7, emphasis added)

Again, it is in the context of a general OV syntax that postpositions enter the picture (early references are Bloch 1965: 327; Chatterji 1970 [1926]: 172). Especially since Emeneau mentioned postpositions in his seminal paper on 'India as a Linguistic Area' (1956), postpositions are commonly put forward in argument for this Sprachbund. In an influential study, Masica (1976) draws 'isoglosses' of word order and a handful of other traits not only within South Asia, but across the entire Eurasian continent.<sup>5</sup> With respect to word order, he claims that this domain alone suffices to demarcate an Indian linguistic area as a core left-branching area (i.e. postpositional placement and other sub-clausal ordering options abiding by a general OV syntax), fading out at its fringes (think of Kashmiri V2 order, for instance). While Masica abstains from a discussion on the direction of influence,<sup>6</sup> others clearly identify

<sup>2</sup> An early reference is Caldwell (1875 [1856]); influential contributions are e.g. Emeneau (1956, 1980), Kuiper (1967), Hock (1975, 1996a), Gumperz and Wilson (1971), Southworth and Apte (1974), and Masica (1976).

<sup>3</sup> Some authors have discussed the possibility of internal origin (Hamp 1996; Hall 1997). Hock has pointed out that the evidence adduced for the contact scenario is not coercive (Hock 1975, 1984, 1996a).

<sup>4</sup> Other features typically listed as defining an Indian linguistic area are echo-word formations, the direct-oblique alternation of nominal forms, lack of prefixes, converbs, compound verbs, morphological causatives, dative subjects, quotative particles, as well as a handful of other suggestions.

<sup>5</sup> Masica only registers one data point per language. He concedes that this represents an overgeneralization, but assumes its validity in Greenbergian tradition, arguing that the most frequent pattern in a language deservedly may be taken as the relevant datum (1976: 18–19).

<sup>6</sup> 'The clearer areal definition of zones of convergence, if it is possible, would have a bearing on some of the historical questions regarding such zones.... Meanwhile, such purely synchronic study defers for the

Dravidian as having endowed Indo-Aryan with OV syntax, including postpositions. For instance, Thomason and Kaufman (1988: 141ff) invest several pages in dismissing ‘anti-Dravidian-influence reasoning’ (1988: 142) of more sceptical authors:

From what is known now, in any case, **Dravidian is by far the best candidate for the source of the features in Indic**. We cannot, at present, look elsewhere for a different source, since supporting historical evidence is lacking. More important, we do not need to, because the Dravidian hypothesis accounts for the facts very well. (Thomason and Kaufman 1988: 144)

Such statements of Dravidian contact influence on Indo-Aryan do not always make explicit the time of influence. However, at least a handful of studies assume substrate effects for Vedic or pre-Vedic times with regard to other proposed contact-induced elements and structures, especially the retroflexes (e.g. Kuiper 1967; Hock 1996a), while Chatterji in the quotation above points to early Middle Indo-Aryan, arguing that it is in this period that Dravidian and Indo-Aryan start converging. As the Hindi postpositions, for instance, were not borrowed from Dravidian, I take it that claims of areal influence must be interpreted as claims that Dravidian induced stable GEN N (and possibly general OV) order in Old Indo-Aryan or Middle Indo-Aryan, a pattern which later in the transition to New Indo-Aryan gave rise to the earliest postpositional phrases.<sup>7</sup> As a result, the language-internal and language-external claims converge in a single hypothesis, namely that a reasonably stable pattern GEN N (and OV), either simply present in the language from early on or induced by contact, gave rise to the modern postpositional constructions. As a first step, then, I evaluate in the next section whether this claim holds true for GEN + MADHYE/UPARI. If this should be confirmed, one can in a second step address the question of whether the language-internal or the language-external scenario (or a combination of both) is most plausible.

#### 6.1.2 *Word order patterns of GEN + MADHYE/UPARI in Old and Middle Indo-Aryan*

In all three historical varieties of Vedic Sanskrit, Pali, and Apabhramsha, GEN N strings are frequent with MADHYE and from Pali onwards also with UPARI in accordance with what has generally been said about nominal word order in Old and Middle Indo-Aryan.<sup>8</sup> However, the reversed order of N GEN is not merely a rare exception, but

moment the controversies regarding direction of borrowing and the precise nature of the evolution of various features that have occupied so much attention heretofore’ (Masica 1976: 6).

<sup>7</sup> Recall that some authors also make reference to the local particles in this context. Emeneau, for instance, suggests a combined influence of local particles allegedly placed in postnominal position with ‘[p]arallel constructions in Dravidian [which] may possibly have helped toward the shift’ (see section 4.1 for the full quotation). However, the local particles were shown to have a tendency for the prenominal slot, if only nominally oriented ones are considered, as shown in section 4.2.

<sup>8</sup> For accounts of word order see e.g. Delbrück (1878, 2009 [1888]), Schäufele (1991b), and Hock (ed., 1991) for Vedic, and Thommen (1905) and Canedo (1937) for Pali. To my knowledge, there is no comparable study on word order in Apabhramsha.

indeed occurs numerous times across all stages (figures will be given below). Not only in Vedic Sanskrit and Pali, but even in Apabhramsha we can still find examples such as the following:

- (133) APABHRAMSHA  
*dummahila*                      *ji*    *vagghi*                      *majjhe*    *gharaho*  
wicked\_woman.NOM.SG.F    PRT    tigress.NOM.SG.F    MADHYE    house.GEN.SG  
‘A wicked woman is a tigress **in the house**.’                      (PA 15.13.8)

- (134) APABHRAMSHA  
*uppari*    *dumaho*    *paḍāya*                      *nivandhivi*  
UPARI    tree.GEN.SG    flag.ACC.SG.F    take\_off.CONV  
‘(He) took the flag from **on top of the tree**.’                      (BH 110.1)

What is more, instances of discontinuity are also attested until late Middle Indo-Aryan. Recall from section 2.3.2 that discontinuity comes in many forms and the insertion of a clitic, for instance, does not have to be taken as evidence against phrasal structure, while other types of discontinuous pattern indicate a lack of phrasal cohesion. In the following, I will outline the different types of discontinuity found in some detail. I group the different kinds into four types<sup>9</sup> ranging from what I refer to as ‘weak’ to ‘strong’ discontinuity, with Type 1 discontinuity being the ‘weakest’ type and Type 4 discontinuity the ‘strongest’.<sup>10</sup>

In Type 1 discontinuity, the cohesion of a group is not severely disrupted, as only a clitic is inserted into the string, excluding the special case of P2 clitics.<sup>11</sup> P2, in turn, which has a special status with an elaborate internal structure and is filled by elements of very different functions in the clause, is involved in Type 2 as well as Type 3 discontinuity. Both types were introduced in Chapter 2 as ‘P2 effects’. First, in Type 2 discontinuity, which involves clitics as in Type 1 discontinuity, a possessor of the relational noun expression, if pronominally realized, is placed in the Wackernagel position. In Type 3 discontinuity, semantically connected elements are separated because one element is placed in initial position preceding the Wackernagel slot, this slot being restricted to one lexeme. In Type 4 discontinuity, discontinuity is not

<sup>9</sup> I do not elaborate on aspects of information structure, but restrict myself to outlining the syntagmatic patterns found.

<sup>10</sup> My categorization is certainly not without alternatives. I strive to categorize the different types from the perspective of the factors that induce discontinuity, and largely ignore other considerations. To give an example, coordinating conjunctions may figure both in my categories of Type 1 and Type 2 discontinuity. For a different approach, see Schäufele (1991) in the two frameworks of GB and LFG.

<sup>11</sup> As already mentioned in section 2.4, this type of clitic insertion is also well known from configurational structures and is attested, for instance, for spoken Hindi where an emphatic clitic may be inserted at various points inside the postpositional construction (e.g. Sharma 1999). Note that, while I use the term ‘insertion’ here, I am agnostic regarding the mechanics (and formalizations) of this phenomenon.

caused by inserted clitics, nor does it have a relation to P<sub>2</sub>; the words in question stand apart for independent reasons.

Before I turn to discussing examples, recall that the constructions examined here cannot be cases of apparent discontinuity. We are not dealing with a symmetrical group of two syntactically co-ranking and equally morphologically marked elements, but with a nominal head and its adnominal genitive dependent. These cannot be unconnected syntactic units, since they are an asymmetrical group connected by semantic as well as morphological dependency relations.

One might interject here that possessors may also be moved out of their nominal domain in languages which have a well-established noun phrase (cp. fn. 33 in Chapter 2), so that a diachronic examination of their syntactic behaviour is not necessarily of interest. However, a diachronic change can clearly be identified regarding expressions involving elements such as *MADHYE* and *UPARI* that grammaticalize into postpositions. Old Indo-Aryan and Middle Indo-Aryan varieties do not possess a nominal phrase out of which the possessor could be moved. Also, there is no apparent constraint on the 'landing site' of the possessor, so that the concept of movement rules fails in the absence of a way of constraining it. With time, however, the adnominal genitive develops the obligatory constraint of appearing in the slot directly preceding *MADHYE*, *UPARI*, etc. While it is true that possessors may be drawn out of the nominal phrase in modern Hindi, this is a recent development that is tied to the secondary development of a new adnominal possessive construction, itself headed by the postposition *kā/-e/-ī* (see exx. (41) and (42)).<sup>12</sup> That said, I now turn to the data.

Starting with Type 1 discontinuity, enclitic discourse particles or conjunctions may disrupt a group in a position other than the Wackernagel slot.<sup>13</sup> Examples can be seen in example (135) where an emphatic particle interrupts the string, and in example (136) which involves the conjunction *ca* 'and'.<sup>14</sup> I also consider vocatives placed in between a head noun and possessor to be a variant of Type 1 discontinuity.

<sup>12</sup> In Hindi, then, one may indeed speak of movement out of the nominal domain because the movement applies to a defined embedded phrasal structure (the possessive PostP) and is restricted to specific slots within the clause, i.e. may not be inserted into other postpositional phrases and is typically drawn to the edges of the clause.

<sup>13</sup> I am lumping together conjunctions and emphatic particles in this category as I focus on patterns of discontinuity, not on the functional relation between clitic and prosodic host.

<sup>14</sup> If at least a part of the genitive expression (if complex) is adjacent to *MADHYE/UPARI*, I categorize this, for the purposes of this chapter, as continuous, e.g.: *niya-māhappu samaggaha* (complete.GEN) *vi / jayaha* (world.GEN) *majhi payadantu* 'having made evident his magnanimity in the entire world' (SA 713.7). Similarly, I also classify structures as continuous, if *MADHYE/UPARI* is one of two or several nominal elements that together take a genitive possessor, and where a different one of the nominal heads is in adjacency with it.



(135) PALI

*atthaccaye*                      *mā ahu*                      *sammūlho /*  
 perishing\_of\_riches.LOC.SG    NEG    be.AOR.2SG    bewildered.NOM.SG  
*bhinnaplavo*                      *sāgarasseva*                      *majjhe*  
 bhinnaplavo                      sāgarassa eva                      majjhe  
 ship\_wrecked.NOM.SG    ocean.GEN.SG PRT    MADHYE

‘In view of perishing riches be not distressed (like someone) shipwrecked in the middle of the ocean.’  
 (Jā III, p. 158)

(136) PALI

*te*                      *tattha jūte*                      *ubhaye samāgate,*  
 DEM.NOM.PL    there    fight.LOC.SG    both    come\_together.PPP.NOM.PL  
*raññaṃ*                      *sakāse*                      *sakhīnañ*                      *ca majjhe*  
 king.GEN.PL.M    presence.LOC.SG    fellow.GEN.PL    and    MADHYE

‘The two met there in contest in the presence of the kings and among the friends.’  
 (Jā VI, p. 282)

Type 2 and Type 3 discontinuity consist of P2 effects, already encountered in section 2.3.1. The Wackernagel position is an internally complex position in Old and Middle Indo-Aryan and can be filled by different accented and unaccented elements such as conjunctions, different types of pronoun, and discourse particles. Hock (1982) is the first modern day elaboration of Jacob Wackernagel’s findings for (Mantra) Vedic, and he proposes the following maximal template (1982: 11, the additional line specifying the P1 and P2 position is my own):

(137) WACKERNAGEL POSITION IN MANTRA VEDIC (adapted from Hock 1982)

Ḍ / Ḍ́	(Prt <sup>15</sup> )	(Ḍ́rt)	(Ḍ)	(Ḍ́)	Ḍ́
P1	P2				

In initial position preceding P2 (henceforth P1), we often find an accented pronominal element (Ḍ́) or alternatively some other lexeme (Ḍ́).<sup>16</sup> Inside the Wackernagel position, the first slot hosts unaccented sentence particles (Prt) such as *u* ‘and, but, now’ or *ha* with emphatic function, followed by accented sentence particles (Ḍ́rt) such as *khālu* ‘indeed’ or *tú* ‘but’. The third slot is occupied by unaccented

<sup>15</sup> Hock uses ‘P’ for discourse particles in P2.

<sup>16</sup> Sometimes, elements appear in a position even before P1. These are discourse linkers such as *atha* ‘then’ which may either occur in P1 or precede another element that occupies this position (see Schäufele 1991b: 40–2 on such ‘nexus’ elements).

demonstrative forms or clitic forms of personal pronouns (D), while the final position is occupied by accented demonstratives (Ā) (see Hock 1982: 11ff for illustrations).<sup>17</sup> In general, each of the positions inside P2 can be filled with more than one element and none has to be filled. Also, there are many cases where a clause does not have a P2 nor, by implication, a P1. The following two examples illustrate typical instantiations of the Vedic clause-initial string (from Hock 1982: 12; see also examples in section 2.4.1.2):

- (138) *ādanti ha sma vaí etásya purāñnam*  
 eat.3PL PRT PRT PRT DEM.GEN.3SG early\_food.ACC.SG  
 Ȳ P<sub>rt</sub> P<sub>rt</sub> P<sub>rt</sub> Ā Ā  
 ‘They eat his earlier food.’ (Kāthaka Saṃhitā 32.2)

- (139) *tām tú naḥ āgatām pratiprábrūtād*  
 DEM.ACC.SG.F but 1PL come.PPP.ACC.SG.F announce.IMP.2SG  
 Ā P<sub>rt</sub> D Ȳ  
 ‘Announce her to us as having come.’ (ŚaB 3.2.1.22)

The Wackernagel position, then, may be implicated in discontinuity of GEN + N in two different ways. As Meillet and Vendryes have it, ‘the second position often has the effect of separating words united by their meaning’ (1927: 526, translation UR).<sup>18</sup> First of all, genitives, if pronominal, may enter P2 (Type 2 discontinuity). This refers to genitive forms of 1st and 2nd personal pronouns and genitive forms of demonstratives for 3rd person pronouns, which function as pronominal possessors. Vedic has a few other possessive pronouns morphologically derived from the personal pronouns, but these are hardly used. In the following, I will use the cover term of ‘pronominal possessors’ for the above-mentioned elements.

Second, either the genitive (whether nominal or accented pronominal) or the head noun may be placed in initial position demarcated by Wackernagel elements, with the respective other element further back in the sentence (Type 3 discontinuity). This overlaps partly with what others have identified as P2 ‘splitting’ sentence-initial ‘constituents’ (see references in Taylor 1990: 72f). As pointed out in section 2.4.1.2, however, the elements in question do not necessarily have to form an immediate frame around P2 in Vedic, but the second element may also come later in the clause as illustrated by example (25).

An example for Type 2 discontinuity can be seen in example (140) from Vedic Sanskrit. Even though *no* and *madhye* are syntagmatically continuous this example is classified as Type 2 discontinuity, because *no* is a clitic in P2 hosted by *katham*, while

<sup>17</sup> Since Hock (1982), several publications on the details of P2 have appeared, especially regarding the relative order of elements, syntactic and/or prosodic constraints, as well as differences between historical layers and genres of Vedic literature, e.g. Hock (1992, 1996b), Hale (1995, 1996), and Schäufele (1996).

<sup>18</sup> ‘[E]lle [= la seconde place, UR] a souvent pour effet de séparer des mots unis par le sens.’

*madhye* stands in the slot directly following P2. Thus, the elements are placed in their respective positions for independent reasons having to do with clitic placement, not because they form a constructional unit. Example (141) from Pali is a combination of Type 2 and Type 3 discontinuity. Type 3 discontinuity is further illustrated in (142) and (143) from Vedic Sanskrit.

- (140) VEDIC SANSKRIT, TYPE 2

<i>dāsyāḥ</i>	<i>putraḥ</i>	<i>kitavo</i>	<i>’brāhmaṇaḥ</i>
slave.GEN.SG.F	son.NOM.SG.M	cheat.NOM.SG.M	non_Brahmin.NOM.SG.M
<i>kathaṃ</i>	<i>no</i>	<i>madhye</i>	<i>’dīkṣiṣṭeti</i>
kathaṃ	no	madhye	adikṣiṣṭa
how	GEN.1PL	MADHYE	consecrate.AOR.MID.3SG
‘The child of a slave woman, a cheat, no Brahman; how has he been consecrated in our middle?’			

(AiB 2.19.1, cp. Keith’s translation)

- (141) PALI, COMBINATION OF TYPE 2 & 3

<i>majjhe</i>	<i>va</i>	<i>no</i>	<i>bhāsa</i>	<i>samantacakkhu</i>
MADHYE	PRT	GEN.1PL	speak.IMP.2SG	all_seeing.VOC.SG.M
<i>Sakko</i>	<i>va</i>	<i>devānam</i>	<i>sahassanetto</i>	
Sakka.NOM.SG.M	like	god.GEN.PL.M	thousand-eyed.NOM.SG.M	
‘Speak in our very middle, oh all-seeing one, like Sakka the thousand-eyed one (in the middle) of the gods.’				

(Sutta 60)

- (142) VEDIC SANSKRIT, TYPE 3 (= ex. (24))

<i>madhye</i>	<i>ha</i>	<i>saṃvatsarasya</i>	<i>svargo</i>	<i>lokaḥ</i>
MADHYE	PRT	year.GEN.SG.M	sky.GEN.SG.M	world.NOM.SG.M
‘(For one half year, he worships the fire with the Vātsapra,) for the world of heaven is in the middle of the year.’				

(ŚaB 9.3.1.10)

- (143) VEDIC SANSKRIT, TYPE 3

<i>madhye</i>	<i>vā</i>	<i>idamātmano</i>	<i>’nnaṃ</i>	<i>dhiyate</i>
madhye	vā	idam ātmano	’nnaṃ	dhiyate
MADHYE	or	DEM.NOM.SG.N	body.GEN.SG.M	food.NOM.SG.N
‘In the middle of the body, the food is put.’				

(ŚānB 15.3)

‘Strong’ or Type 4 discontinuity involves cases in which discontinuity does not arise in connection to (non-P2 related) clitic insertion or P2 effects, but where the elements in question stand apart for independent reasons, as in examples (144) and (145) from Vedic Sanskrit and Apabhramsha.

- (144) VEDIC SANSKRIT (= ex. (4))

<i>devasenānām</i>	<i>abhibhañjatīnām</i>	<i>jáyantīnām</i>
army_of_gods.GEN.PL.F	smash.PPA.GEN.PL.F	conquer.PPA.GEN.PL.F

*marúto*                      *yantu*                      *mádhye*

Marut.NOM.PL.M      go.IMP.3PL      MADHYE

‘In the middle of the smashing conquering armies of the gods, the Maruts should go.’ (AV 19.13.9)

(145) APABHRAMSHA

*majjhi*    *gantu*                      *kayalīharaha*

MADHYE    go.CONV      banana\_house.GEN.SG

‘He went into the banana house.’ (SA 566.5)

UPARI shows a similar breadth of syntactic variation to MADHYE, not only in Pali but also in Apabhramsha, where besides GEN N order, also N GEN, as well as Type 1 (with a lexicalised pronoun-particle combination) and Type 4 discontinuity, are attested:

(146) N GEN (= ex. (134))

*uppari*    *dumaho*                      *paḍāya*                      *nivandhivi*

UPARI      tree.GEN.SG      flag.ACC.SG.F      take\_off.CONV

‘(He) took the flag from on top of the tree.’ (BH 110.1)

(147) TYPE 1 DISCONTINUITY

*ahaha*    *ari*                      *jiya*                      *karisi*                      *mā*    *rōsu* /

EXCL      EXCL      soul.VOC.SG      do.IMP.2SG      NEG      anger.ACC.SG

*iyarassu*                      *kassu*                      *vi*                      *uvari!*

other.GEN.SG.M      Q.GEN.SG.M      PRT      UPARI!

‘Oh soul, don’t be angry with some other person!’ (SA 693.2)

(148) TYPE 4 DISCONTINUITY

*uppari*    *caḍiu*                      *taho*                      *aṭṭhāvayaho*

UPARI      climb.PPP.NOM.SG.M      DEM.GEN.SG      Aṣṭapada.GEN.SG

*pañca-mahāvaya-dhārau*

five-great\_rituals-performer.NOM.SG.M

‘The performer of five great rituals climbed on the Aṣṭapada peak.’

(PA 12.11.9)

These examples show how word order freedom is retained in principle until Apabhramsha (if becoming more uniform in quantitative terms). For instance, example (145) is taken from Haribhadra’s *Sanatkumāracaritam*, a text composed as late as the twelfth century—a period where scholars in parallel locate the birth of the modern New Indo-Aryan varieties (cp. *Sanatkumāracaritam* (Jacobi 1921: viii)), which by and large have postpositional phrases.<sup>19</sup> The figures for MADHYE and UPARI are summarized in Tables 6.1 and 6.2.

<sup>19</sup> Some identify an even earlier period. Shapiro (2007: 254–5) summarizes several proposals for the rise of Hindi, the earliest of which is the eighth century as proposed by Dwivedi, while many others point to the

**TABLE 6.1. Continuous and discontinuous relational noun expressions with MADHYE<sup>20</sup>**

MADHYE		Vedic Sanskrit (66)	Pali (43)	Apabhramsha (34)
<b>Continuous</b>	Σ	55	33	33
	GEN MADHYE	25 <sup>a</sup>	21 <sup>b</sup>	22 <sup>c</sup>
	MADHYE GEN	30 <sup>d</sup>	12 <sup>e</sup>	11 <sup>f</sup>
<b>Discontinuous</b>	Σ	11	10	1
	Type 1	0	3 <sup>g</sup>	
	Type 2	2 <sup>h</sup>		
	Type 3	4 <sup>i</sup>	2 <sup>j</sup>	
	Type 4	5 <sup>k</sup>	3 <sup>l</sup>	1 <sup>m</sup>
	comb. Type 2/3		2 <sup>n</sup>	

<sup>a</sup> RV 1.32.10, 7.89.4, 10.5.1, 10.11.2, 10.102.9, 10.139.2, YV 13.42, 13.49, 15.52, 17.59, 17.87, AV 10.7.38, 10.8.15, ŚaṅB 11.5, 15.3b, AiB 8.21.10, ŚaB 2.5.3.6, 3.5.2.10, 7.5.2.18, 7.5.2.34, 8.6.3.29, 9.2.3.17a, 10.2.3.1 10.5.4.15, 13.7.1.15.

<sup>b</sup> Digh II pp. 135, 274, Majjh II p. 196, III p. 128a, Sam I p. 156, II p. 229, V p. 219, Ud p. 84, Vim p. 122, 126, Peta p. 67, Jā I p. 412, 472, II p. 180, 228, III p. 380, 396, 508, IV p. 464, VI p. 117, Buddha p. 44.

<sup>c</sup> BH 42.6, 46.8, 82.8, 130.1, 207.16, 250.14, 284.8, 284.10, 294.8, KA 2.20.8, 6.16.1, 7.12.4, 8.5.11, 8.15.3, 9.8.4, 9.9.7, 9.14.13, 10.2.6, PA 5.11.9, SA 557.1, 713.7, 755.5.

<sup>d</sup> RV 1.108.12, 1.158.3, 1.182.7, 4.58.5, 5.47.3, 7.41.4, 9.65.23, 10.15.14, 10.102.5, YV 13.38, 17.60, 19.60, AV 4.15.14, 5.19.3, 6.106.2, 13.2.36, ŚaṅB 14.1a, 25.8, AiB 2.33.2, 2.33.3, 4.18.1, 4.22.1, ŚaB 1.7.4.20, 5.1.2.19, 5.4.3.9, 7.5.2.11, 9.2.3.18a, 9.3.1.10a, 9.3.4.18, 11.4.1.15.

<sup>e</sup> Digh II p. 337, Majjh I p. 337, II p. 146b, Aṅg IV pp. 201, 205, Ud pp. 52, 55, Sutta p. 108, Thera p. 78, Jā III p. 413, Jā V p. 223, Apa I p. 326.

<sup>f</sup> BH 46.5, 193.8, 207.17, PA 1.11.3, 6.3.6, 15.13.8, SA 484.3, 504.8, 514.6, 689.5, 707.4.

<sup>g</sup> Jā III p. 158, V p. 169, VI p. 282.

<sup>h</sup> AiB 2.19.1, ŚaB 8.6.3.4b.

<sup>i</sup> ŚaṅB 15.3a, 16.7b, ŚaB 6.7.4.11, 9.3.1.10b.

<sup>j</sup> Nidd I p. 353a, Apa II p. 387.

<sup>k</sup> RV 1.105.10, 7.49.3, 10.138.3, AV 19.13.9, ŚaB 12.5.1.17.

<sup>l</sup> Sutta p. 180, Paṭi pp. 167a, 167b.

<sup>m</sup> SA 566.5.

<sup>n</sup> Sutta p. 60, Thera p. 113.

These figures make clear that even Apabhramsha does not approximate the modern state of affairs, and that freedom of word order, albeit less exploited, has not changed in its essentials until this point for combinations of GEN + MADHYE/UPARI.

twelfth to fourteenth centuries. (Hindi in a modern sense as contrasting with Urdu is a much more recent phenomenon.)

<sup>20</sup> Throughout this chapter, I only take into account formally unambiguous genitives and exclude from consideration syncretic forms of genitive and locative (for the figures see section 5.2). The references given in the footnotes here and in the remainder of this study follow the order as given in the appendix.

**TABLE 6.2. Continuous and discontinuous relational noun expressions with UPARI**

UPARI		Pali (9)	Apabhramsha (35)
<b>Continuous</b>	$\Sigma$	9	33
	GEN UPARI	7 <sup>a</sup>	28 <sup>b</sup>
	UPARI GEN	2 <sup>c</sup>	5 <sup>d</sup>
<b>Discontinuous</b>	$\Sigma$	0	2
	Type 1		1 <sup>e</sup>
	Type 2		
	Type 3		
	Type 4		1 <sup>f</sup>
	comb. Type 2 & 3		

<sup>a</sup> Digh I p. 95, III p. 85, Majjh I p. 231, Saṃ I p. 144a, Apa I p. 273, 276, II p. 608.

<sup>b</sup> BH 222.3, 228.8, KA 2.19.10, 3.8.14, 3.12.10, 3.13.8, 4.1.13, 4.3.8, 4.4.4, 4.6.1, 4.12.1, 4.15.2, 5.4.7, 5.13.6, 6.7.6, 7.6.7, 7.11.13, 9.4.6, 9.15.5, 10.23.2, PA 1.3.3, 13.2.6, SA 489.5, 522.9, 594.1, 598.1, 650.5, 674.5.

<sup>c</sup> Saṃ V p. 445, Aṅg IV p. 404.

<sup>d</sup> BH 110.1, KA 2.9.1, 5.7.9, PA 13.1.10, SA 589.8.

<sup>e</sup> SA 693.2.

<sup>f</sup> PA 12.11.9.

Specifically, even when adjacent, we continue to find a considerable number of examples displaying the ‘reversed’ order N GEN. For MADHYE, these cases even outweigh GEN N order in Vedic and later form a third of the continuous cases. The cases for UPARI are a little less frequent, but just as well attested. The assumption in the literature of a basic GEN N order (and OV), while not contradicted but weakened by these data, leaves unexplained how we get from this stage to the fixed postpositional syntagm. Also, it is not clear how such tendencies in Vedic, whether internally grown or externally triggered, should bear on developments between late Middle and Early Indic. Importantly, note that this variety is not only attested for the poetic texts, but also for the prose texts of Vedic Sanskrit and Pali. What is more, MADHYE and UPARI in Apabhramsha do not even obligatorily take genitives yet:

- (149) *tihiṃ phalahiṃ majjhe ekkaho phalāsu /*  
 three.LOC fruit.LOC.PL MADHYE one.GEN fruit.GEN.SG  
*ṇirahariyau riṇu maiṃ maivarāsu*  
 remove.PPP.NOM.SG debt.NOM.SG INS.1SG noble\_minded.DAT.SG  
 ‘I have paid off the debt of one fruit out of three to the noble-minded.’ (lit.  
 ‘The debt of one fruit in the middle within three has been removed by me, for  
 the noble-minded one.’) (KA 2.18.2)

- (150) *uppari dasa-joyaṇehiṃ divāyaru / puṇu asīhiṃ*  
 UPARI ten-joyana.LOC.PL sun.NOM.SG again eighty

*lakkhijai*                      *sasaharu*  
 perceive.PASS.3SG      moon.NOM.SG  
 ‘The sun (is) at ten *yojanas* [= measuring unit] **beyond**; the moon can be seen at eighty (*yojanas*).’ (PA 2.3.4)

- (151) *vaddhu*                      *kaḍilli*                      *alakkhiyanāmaū* / *uppari*  
 bind.PPP.NOM.SG      hip\_cloth.LOC.SG      petticoat.NOM.SG      UPARI  
*pīṇiu*                              *rasaṇā-dāmaū*  
 buckle\_around.NOM.SG      rope-garland.NOM.SG  
 ‘A petticoat is bound to the hip-cloth; **above** a rope garland is buckled around.’ (BH 167.2)

Moving on in time, I skip the *ādi kāl* ‘first period’ (eleventh to fourteenth century) of Hindi, where we have almost solely attestations of poetry and Sanskrit and the Prakrits continue to be heavily drawn upon (see McGregor 2001). I turn directly to the renowned sixteenth-century text *Rāmacaritamānasa*, a retelling of the *Rāmāyaṇa* in Old Awadhi by Tulsidas, equally metrical but voluminous, with good secondary literature. Here, the postpositional slot can be considered nearly established. Despite metrical constraints which might enforce another order, only four examples out of 282 attestations involving MADHYE show N GEN order, and none discontinuity (see Reinöhl 2009).<sup>21</sup>

In view of the data presented in this section, I draw the following conclusions. First, the syntax of GEN + N expressions as evaluated here for MADHYE/UPARI remains essentially unchanged across Old and Middle Indo-Aryan with both possible orders and various types of discontinuity attested. Second, the propensity for GEN N order can be confirmed and increases with time, but appears less pronounced than assumed in the literature. Third, the shift that turns MADHYE and UPARI into adpositions with all the properties of phrasal heads, selecting dependents and constraining their position, takes place between late Middle and early New Indo-Aryan, and no earlier.<sup>22</sup> Overall, and this is the fourth conclusion, traditional proposals that rely on the—essentially correct—observation of a propensity for GEN N and OV order cannot answer the question of why New Indo-Aryan has developed adpositions which *obligatorily* stand in the postnominal slot. Therefore, I do not evaluate any further here language-internal or language-external scenarios that have been suggested, as these do not go beyond pointing out a tendency for GEN N (and OV). What is necessary is to explain the fixation of this pattern.<sup>23</sup>

<sup>21</sup> Prose works are only well attested again from around 1800 (see McGregor 1967).

<sup>22</sup> The spoken varieties may have been ahead in the development.

<sup>23</sup> Contact with Dravidian and other languages may well have played a role in the fixation. However, studies regarding Dravidian influence such as those cited at the beginning of this chapter are almost always directed at Vedic or pre-Vedic times. What seems necessary here is research into contact in the Middle Indic and early New Indic periods.

## 6.2 The Wackernagel position across time

In search of an explanation for the fixation of order, I now turn to the loss of P2. In section 2.4.1.2, I pointed out that several other types of discontinuity exist and not all strongly non-configurational languages have a P2. In addition, I argued that the 'splitting' analysis must be rejected for Indo-Aryan. Still, the loss of the Wackernagel position does indeed run very roughly in parallel with the rise of configurationality in several branches of Indo-European including Indo-Aryan. By the stage of New Indo-Aryan, not only have postpositions emerged, but P2 has also vanished. Therefore, in this section I assess whether there might be a connection between the two developments.

Taylor (1990) discusses the development of the placement of pronominal clitics in P2 across historical varieties of Ancient Greek, which bears on what I refer to as Type 2 discontinuity. She argues that Ancient Greek, having always been underlyingly configurational in her GB analysis, undergoes a 'stylistic' change from Homeric to Koine Greek in the form of a general decrease of extractions of clitics out of their 'normal' position and into P2 (or, in Taylor's terminology, an increase in the blocking of extractions), alongside the decrease of other extractions. In the post-Homeric Greek period, clitics increasingly often remain 'in situ', attaching to lower phrasal levels rather than to IP. As Taylor puts it, '[i]t is the movement option dying out which leaves the clitic in its base-generated position' (1990: 169). An underlying shift from OV to VO syntax is adduced as another factor in bringing about the loss of the Wackernagel position, as Taylor argues that head-final VPs do not block extractions, whereas head-initial ones do (1990: 171–2). Evidence from post-Homeric stages of Greek shows, however, that the Wackernagel position remains a possible 'landing site', if less frequently used, which is proposed to be due to independent processes of topicalization out of an underlying OV structure (1990: 178 *et passim*).<sup>24</sup>

Is there evidence for a connection between the loss of the Wackernagel position and the establishment of the postpositional phrases in Indo-Aryan? In the early and middle stages of Middle Indo-Aryan, the Wackernagel position shows some changes regarding the kinds of word which may enter it, as well as regarding its internal

<sup>24</sup> Taylor's study relates the loss of Type 2 discontinuity to the rise of configurational structure and is thus akin to my discussion of the development of P2 in Indo-Aryan. Other authors, by contrast, discuss a different type of connection between P2 and the rise of configurational structure, which involves a reanalysis of pronominals while placed in P2. For example, Fontana (1993) studies changes in the syntax of Old Spanish which involve discontinuities arising from object clitics placed in P2 'splitting' clause-initial strings. Fontana argues that P2 clitics were reanalysed as verbal prefixes in the history of Spanish as a result of the loss of topicalizations of lexemes into clause-initial position. Because the necessary hosts in initial position vanished, P2 clitics instead came to lean on finite verbs to their right. Vincent (1999) argues that the origin of the prepositional syntax in Latin lies in local particles being reanalysed as prepositions when standing in clause-initial position and hosting enclitic pronominal forms placed in P2.



complexity in comparison to Old Indo-Aryan. As regards Type 2 discontinuity, object pronouns now commonly remain in adjacency to the verb and less frequently move to P<sub>2</sub> (Thommen 1905: 533), matching Taylor's findings for Ancient Greek. Otherwise, there still is a P<sub>2</sub> filled by non-pronominal elements. Jacobi (1967 [1886]) outlines how in Mahārāṣṭrī (a Prakrit from the middle period of Middle Indic), P<sub>2</sub> can host, for instance, the copula, the conjunctions *ca* and *ya*, or the emphatic particles *pi* and *vi* (both from Skt. *api* 'also') (Jacobi 1967 [1886]: LVI, LXVI–LXIX). These particles are old inherited elements, in part already occurring in P<sub>2</sub> in Sanskrit and especially in Pali, but new elements also develop into P<sub>2</sub> discourse markers, for example *dāni* (from Skt. *idānim* 'now') or *pana* (from Skt. *punar* 'again'). Regarding Type 3 discontinuity, Thommen (1905: 535) notes that, from the Sanskrit epics onwards (i.e. early Classical Sanskrit), the constraint regarding a single lexeme only in P<sub>1</sub> is relaxed and groups may enter it as a string: for example *thairānaṃ dasane ca* (elder.GEN.PL sight.LOC.SG and) 'and at the sight of the elders' (Ashoka inscription, cited in Thommen 1905: 535).

Apabhramsha continues the developments of the early and middle periods of Middle Indo-Aryan in that the Wackernagel position is weakened further and it is debatable whether one should still assume a P<sub>2</sub> position at this stage. First of all, I did not find any case of Type 2 or 3 discontinuity in clauses containing either MADHYE or UPARI for this period. With regard to Type 2 discontinuity, clitic pronouns, while still attested for Prakrit varieties (cp. Jacobi 1967 [1886]), have generally fallen out of use (but see an instance of an archaic example quoted in Bubeník and Paranjape 1996: 118).<sup>25</sup> In order to test for accented pronouns which could appear in P<sub>2</sub> in earlier times, I went through all genitive *sa-/ta-*stem demonstratives and 1st person personal pronouns in the *Sanatkumāracaritam*, none of which occurs in P<sub>2</sub>. There are fifty-seven instances of genitive *sa-/ta-*stem demonstratives. Of these, in twelve<sup>26</sup> cases the genitive demonstrative unambiguously functions as a possessor in a nominal expression, with one case of Type 4 discontinuity (SA 680.9).<sup>27</sup> Of forty-five 1st person genitive forms, thirty function as adnominal possessors with twenty-eight<sup>28</sup> showing adjacency, one case of Type 1 discontinuity (SA 659.1), and one case of Type 4 discontinuity (SA 482.3). Based on this sample, I conclude that pronominal possessors do not or hardly ever enter P<sub>2</sub> in Apabhramsha anymore, which removes Type 2 discontinuity for late Middle Indo-Aryan.

<sup>25</sup> Modern Kashmiri has developed pronominal clitics which attach to the verb (e.g. Emeneau 1980: 137ff).

<sup>26</sup> SA 498.9, 508.3, 509.9, 537.8, 561.4, 614.9, 661.1a, 680.9, 777.5, 450.1, 728.1, 650.5.

<sup>27</sup> This excludes cases where the pronoun (i) directly depends on the verb, (ii) agrees with a lexical noun that is itself a genitive, (iii) is used in dative or ablative function, or (iv) is used in a genitivus absolutus construction.

<sup>28</sup> SA 485.7, 486.4, 489.5, 499.9, 500.2, 506.9, 507.2, 507.6, 508.6, 514.8, 515.2, 518.4, 523.1, 561.2, 574.2, 575.3, 598.1, 601.1, 607.7, 617.2, 627.6, 647.5, 690.2, 706.6, 709.6, 722.2, 722.6, 773.9.

Turning to Type 3 discontinuity, what about the initial slot delimited by P2 elements? As the constraint regarding one word only in P1 has already started relaxing in earlier periods, this type of discontinuity has diminished in significance. Indeed, I found no attestation of Type 3 discontinuity in my corpus for MADHYE/UPARI. In addition, it appears that ‘Wackernagel elements’ are not strictly constrained to appear in P2 any more. Compare the following examples: the particles *vi hu* from Skt. *api* (‘also’) and *khālu* (‘indeed’), as well as *puṇu/-a* occur in second position (exx. (152)–(154)), but also in positions later in the clause as in (153) and (155), and in initial position as in (156).<sup>29</sup>

- (152) *āyariṣi* *vi hu niya-gahīratta-avahatthiya-duddh’uyahi*  
 royal\_wise\_man.NOM.SG.M PRT PRT own-profundity-repelled-milk\_ocean  
*sayala-rōga niru samma-karaṇiṇa ahiyāsai*  
 entire-sickness certainly right-doing.INS endure.3SG  
 ‘And the royal wise man, whose profundity was greater than the milk ocean,  
 endured all the sickness through his right action.’ (SA 778)
- (153) *rāya-riṣiṇa vi tiyasagiri-sihara-thira-cittiṇa*  
 royal-wise\_man.INS.SG.M PRT divine\_mountain-peak-steady-mind.INS.SG  
*siha-avalōiṇa vi hu ti na nirikkhiya*  
 lion-glance.INS.SG PRT PRT QUOT NEG look\_at.PPP.NOM.SG.F  
 ‘The royal wise man whose mind was like the steady peak of a divine  
 mountain did not look at (her) even with a lion’s glance.’ (SA 766)
- (154) *kamiṇa puṇu aṇudiṇu vi parigamiru sampattu*  
 step\_by\_step again daily PRT wander.3SG reach.PPP.NOM.SG  
*mahāḍavihī kaha-vi*  
 great\_forest.LOC.SG somehow  
 ‘When he wandered every day step by step, he somehow reached a great  
 forest.’ (SA 537)
- (155) *aiugga-pavaṇiṇa saṃchāiya diṣi bhariya*  
 terrible-blowing.INS.SG cover.PPP.NOM.SG all\_around fill.PPP.NOM.SG  
*kumara-nayaṇa puṇa rēṇu-pasariṇa*  
 prince-eye.NOM.SG again dust-advancing.INS.SG  
 ‘By the terrible storm all around, the prince’s eyes were again filled with  
 advancing dust.’ (SA 579)

<sup>29</sup> In rare cases, some of the accented discourse markers like *khālu* also occur in initial position in Old Indo-Aryan. In the case of *puṇu* from *punar*, one might argue that it has not yet become a proper P2 clitic. However, the constraints do not seem to be so strict in general any more, whether involving old P2 elements or newer ones. Future research may shed more light on the details of P2 in this period.

- (156) *punu vi paṇamivi bhaṇiu tiyasehī*  
 again PRT bow.CONV speak.PPP.NOM.SG god.INS.PL  
 ‘And again the gods, bowing down, spoke:...’ [lit. ‘it was spoken by the  
 gods’] (SA 735)

Summing up the situation in Apabhramsha, Type 2 discontinuity seems to have virtually ceased to occur. In a similar way, Type 3, while still a possibility, is not attested in my corpus and has ceased to play an important role for discontinuity, as entire groups are allowed into initial position. In general, P2 elements do not appear to be so strictly constrained to appear in second position any more. At the same time, however, Type 1 and Type 4 discontinuity, which are not connected with the Wackernagel position, continue to be attested, if rarely. Therefore, while it may have contributed to the decrease in frequency, the weakening of the Wackernagel position appears to have had no direct effect on the loss of discontinuity and the establishment of the postpositional phrases. This is in accordance with the finding that, while a well-established P2 is frequently found in strongly non-configurational languages, it is not a necessary property. Note moreover that the loss of discontinuity itself could not in any case account for the establishment of the new phrasal pattern, as relative order must also become fixed.

### 6.3 Pronominal possessors

Turning from the question of juxtaposition to relative order, I focus in the following on expressions with pronominal possessors which deviate from the general set of possessors in prefiguring the postpositional pattern in a great majority of examples from earliest times (section 6.3.1). I then address the question of why these constructions show the pattern  $\text{GEN}_{\text{pro}} \text{N}$  rather than the reverse. I suggest that an association between the left edge of the clause and functions of anaphoric discourse linkage performed by a subclass of the pronominal possessors is an important factor in the distribution (section 6.3.2). This is a main source of the postpositional order.

#### 6.3.1 Data

While we usually find adjacent syntagms in Apabhramsha, both  $\text{GEN} \text{N}$  and  $\text{N} \text{GEN}$  still occur. When we separate nominal and pronominal possessors, however, a different picture emerges. Compare the figures extracted from Table 6.3 showing only expressions with pronominal possessors (*acc* indicating accented, *n-acc* non-accented pronominal possessors).

At first glance, the figures seem not too different from those for all possessors with both relative orders and discontinuity attested. However, when there is continuity, we see in virtually every case the order  $\text{GEN}_{\text{pro}} \text{N}$  in all historical stages from Vedic Sanskrit to Apabhramsha, rather than the reverse (four out of five in Vedic, four out

**TABLE 6.3. Patterns of MADHYE/UPARI with pronominal possessors across time**

MADHYE/UPARI		Vedic Skt. <sup>a</sup>	Pali <sup>b</sup>	Apabhramsha <sup>c</sup>
<b>Continuous</b>	Σ	5	4 (all <i>acc</i> )	19 (all <i>acc</i> )
	GEN N	4 ( <i>acc</i> )	4	18
	N GEN	1 ( <i>n-acc</i> )		1
<b>Discontinuous</b>	Σ	3	2	
	Type 1			
	Type 2	2		
	Type 3			
	Type 4	1		
	comb. Type 2/3		2	

<sup>a</sup> MADHYE: RV 4.58.5, 7.49.3, ŚaṅB 15.3b, AiB 2.19.1, ŚaB 2.5.3.6, 3.5.2.10, 8.6.3.4b, 10.2.3.1.

<sup>b</sup> MADHYE: Majjh II p. 196, Saṃ V p. 219, Jā III p. 508, Buddha p. 44, Sutta p. 60, Thera p. 113.

<sup>c</sup> MADHYE: KA 7.12.4, 8.15.3, 10.2.6, PA 1.11.3. UPARI: BH 222.3, 228.8, KA 2.19.10, 3.13.8, 4.1.13, 4.3.8, 5.4.7, 5.13.6, 7.6.7, 7.11.13, 9.15.5, SA 489.5, 522.9, 598.1, 650.5.

of four in Pali, and eighteen out of nineteen in Apabhramsha). I concentrate on this stable pattern here and will comment on discontinuous cases in the next section. I begin by discussing the two exceptions to GEN<sub>pro</sub> N order, first the Apabhramsha example which is an actual exception, and afterwards the Vedic example, which is a special case as it involves a clitic pronominal.

(157) APABHRAMSHA N GEN

*tailokku*                      *pariṭṭiu*                      *majjhe*    *tāsu* /  
 three\_worlds.NOM.SG    stay.PPP.NOM.SG    MADHYE    DEM.GEN.SG  
*caudaha*    *rajjuya*                      *āyāmu*                      *jāsu*  
 fourteen    [measuring unit]    stretch.3SG    REL.GEN.SG.M

‘The Triloka is lodged in the middle of it, (the area) of which spreads to fourteen rajus.’  
 (PA 1.11.3, translation Nagar)

In example (157), the demonstrative *tāsu* forms the end rhyme with the relative *jāsu*. Such an end rhyme is paralleled many times with nominal possessors, as in the following example with *mahā-saraho* and *divāyarahō*.

(158) APABHRAMSHA N GEN

*uppajjavi*                      *majjhe*    *mahā-saraho* /  
 come\_about.CONV    MADHYE    great-lake.GEN.SG  
*ṇaliṇiu*                      *viyasanti*                      *divāyarahō*  
 Nalin\_plant.NOM.PL.F    bloom.3PL    sun.GEN.SG

‘It happened in the middle of the great lake; the Nalin plants bloom through the sun.’  
 (PA 6.3.6)

While the exceptional pattern in (157) can be accounted for by metrical constraints, the following Vedic example (159) is a different case, involving an enclitic pronominal form, that is, a clitic remaining ‘in situ’ rather than migrating into P2. Note that clitic pronominals do not form part of the groups that develop into postpositional phrases—in that only accented pronominal forms are involved. Therefore, while this is a counter-example on a strictly syntagmatic level, it does not bear on the question of the emergence of a postpositional rather than of a prepositional template.

- (159) VEDIC N GEN  
*ghṛtāsya dhārā abhī cākaśimi hiraṇyāyo*  
 clarified\_butter.GEN.SG.N stream.ACC.PL.F LP see.INT.1SG golden.NOM.SG.M  
*vetasó mādhyā āsām*  
 cane.NOM.SG.M MADHYE DEM.GEN.PL.F  
 ‘(I) see the streams of ghee [clarified butter]. The golden cane is **in their middle**.’  
 (RV 4.58.5)

Apart from these two cases, all twenty-six other attested examples display the order GEN MADHYE/UPARI.<sup>30</sup> Compare the following examples for illustration:

- (160) VEDIC SANSKRIT (= ex. (92))  
*tasya madhye nividam dadhāti*  
 DEM.GEN.SG MADHYE Nivid.ACC.SG.F put.PRS.3SG  
 ‘**In the middle of it** [= a recitation] (he) places a Nivid [= a short invocation].’  
 [Śā 15.3a]
- (161) PALI  
*seyyathāpi bhikkhave nadi*  
 seyyathā pi bhikkhave nadi  
 as PRT monk.VOC.PL river.NOM.SG.F  
*pācīnaninnā pācīnapoṇā*  
 eastward\_bent.NOM.SG.F eastward\_sloping.NOM.SG.F  
*pācīnapabbharā. Tassā majjhe dīpo*  
 eastward\_bending.NOM.SG.F DEM.GEN.SG.F MADHYE island.NOM.SG.M  
 ‘Suppose, monks, a river flows east, slopes east, bends east; **in the middle of it** (is) an island.’  
 (Saṃ V, p. 219)

<sup>30</sup> There is one more example in Pali which shows the reversed order, which I did not include here, however: *Majjhe mayham bhavā assu* (MADHYE GEN.1SG existence.NOM.PL be.OPT.3PL) ‘Existences shall be within me’ (Apā I, 326). This example comes from the *Apadāna*, a text that has not received the same careful critical editing and analysis as other Pali texts discussed here. It is also a special case because it involves a 1st person pronominal form, while almost all other cases involve the demonstrative forms serving as 3rd person possessors. In the next section, I outline specific properties of these 3rd person forms on the level of information structure which may account for the leftward tendency, properties which do not in the same way apply to 1st and 2nd person forms. Lacking these properties, it is possible that the latter did not show the same stable tendency towards the left edge as the 3rd person forms.

## (162) APABHRAMSHA

*tahaṃ majjhi ṇivīṭṭhau so sahei /*  
 DEM.GEN.PL MADHYE settle\_down.PPP.NOM.SG.M DEM.NOM.SG.M shine.3SG  
*chaṇaindaho līlā ṇam vahei*  
 full\_moon.GEN.SG beauty.ACC.SG so bear.3SG  
 ‘Sitting in their middle, he shines as if bearing the beauty of the full moon.’  
 (KA 8.15.3)

This is a distribution markedly different from the figures we saw in Tables 6.1 and 6.2 that included all types of possessors, and where the figures for MADHYE in Apabhrāmsha were roughly 2:1 GEN N VS. N GEN and where several cases of the ‘reversed’ order are also attested for UPARI. The sample is of course too small to allow for any kind of generalization. Therefore, I examined *sa-/ta-stem* pronominal possessors<sup>31</sup> in a Mantra Vedic text (*Rigveda*) and a Vedic prose text (*Śatapathabrāhmaṇa*), as well as the set of *sa-/ta-stem* genitives and genitives of 1st person pronouns in the Apabhrāmsha text *Sanat-kumāracaritam* (already analysed in section 6.2 in the context of Type 2 discontinuity). Of 143 genitives in the Vedic sample, seventy-six encode pronominal possessors, thirty-eight<sup>32</sup> of which are discontinuous. In the remaining thirty-eight<sup>33</sup> continuous cases, we find in each and every case the order GEN<sub>pro</sub> N. The same situation holds for the *Sanatkumāracaritam*. All the eleven *sa-/ta-stem* genitives functioning as pronominal possessors and juxtaposed (out of a total of fifty-seven genitives) show the order GEN<sub>pro</sub> N.<sup>34</sup> There is one discontinuous example. Of forty-five genitive forms of 1st person pronouns, thirty function as possessors, twenty-eight of which are adjacent, all but one displaying the order GEN<sub>pro</sub> N.<sup>35</sup> In short, GEN N order is entirely systematic with pronominal possessors in general.<sup>36</sup> Compare the following small selection of examples:

## (163) VEDIC SANSKRIT

*tāsya vratāny ānu vaś carāmasi*  
 DEM.GEN.SG.M command.ACC.PL.N after DAT.2PL move.1PL  
 ‘(We) want to do his bidding for you.’  
 (RV 8.25.16)

<sup>31</sup> I included the singular and plural forms listed in Viśva Bandhu’s concordance.

<sup>32</sup> RV *tasya* 1.12.8, 1.164.13, 1.164.22, 2.23.12, 2.24.8, 3.17.5, 3.33.6, 4.21.2, 6.7.6, 6.49.13, 8.19.6a, 8.31.3, 8.45.15, 10.10.14, 10.87.10b, 10.88.1, 10.117.6, 10.139.1, 10.160.3, *tasyāḥ* 6.53.9, *teṣām* 1.47.2, 1.139.9a, 1.188.9, 2.30.10, 7.55.6, 7.66.13, 7.90.4, 8.20.14, *tāsām* 2.5.6, 6.17.12, 10.97.18, 10.114.2, 10.169.4 ŚaB *tasya* 1.1.2.22, 1.1.4.1, *teṣāṃ* 1.1.1.11, *tāsām* 1.2.1.22, 3.1.3.9.

<sup>33</sup> RV *tasya* 1.11.6, 1.22.6, 1.59.3, 1.93.8, 1.100.13, 1.141.7, 1.164.1, 2.25.2, 2.35.9, 3.3.9, 4.4.10a, 4.4.10b, 6.50.15, 7.31.11, 7.40.3, 8.25.16, 8.41.3, 10.87.10a, 10.91.11, 10.99.1, 10.133.5 *teṣām* 1.11.7, 1.125.7, 1.139.9b, 1.164.15, 10.82.2, 10.87.16, 10.94.8, 10.115.5, *tāsām* 3.2.9, 10.5.6 ŚaB *tasya* 1.1.1.17a, 1.1.1.17b, 1.1.2.23a, 1.1.2.23b, *tasyāḥ* 1.1.4.3, *teṣāṃ* 1.1.4.7, 1.3.1.2.

<sup>34</sup> See fn. 26 above for the list of attested examples.

<sup>35</sup> See fn. 28 above for the list of attested examples.

<sup>36</sup> In the case of the single exception, we are again dealing with a possessor that is a speech act participant, namely of the 1st person singular. Possibly, possessors of the 1st and 2nd person behave slightly differently compared to 3rd person possessors; see fn. 30 above.

- (164) VEDIC SANSKRIT  
*gārhapatye*                      *tāsya*                      *pātrāṇi*                      *sāṃsādayanti*  
 fire\_of\_the\_household.DAT.SG   DEM.GEN.SG.M   dish.ACC.PL.N   place.3PL  
 ‘(They) place **his dishes** in the fire.’                      (ŚaB 1.1.2.23a)
- (165) APABHRAMSHA  
*tassu*                      *antari*   *puvva-disa-<sa>muhu*                      *sīhāsaṇu*                      *saṃṭhavivi*  
 DEM.GEN.SG.M   inside   east-direction-facing.ACC.SG   throne.ACC.SG   place.CONV  
 ‘**In its interior** (they) placed the throne facing to the east.’                      (SA 728.1)

I conclude that  $\text{GEN}_{\text{pro}}$  N order has been a stable pattern since Vedic times and remains thus until Apabhramsha. This is not to say that the postpositional phrase as such already exists from early on with pronominal possessors. Discontinuity is attested here also, in the same way as reversed order in rare cases. However, I propose that the postpositional phrases have important roots in constructions with pronominal possessors, a claim which will be substantiated in the next chapter, where I show that they also played a central role in the obligatorification of genitive dependents. Before I move on to discussing obligatorification, however, I first enquire into the reasons why  $\text{GEN}_{\text{pro}}$  N is found so frequently, despite the strict meter in some texts and the principally free word order, the freedom of which is exploited by nominal (i.e. non-pronominal) possessors.

### 6.3.2 Why $\text{GEN}_{\text{pro}}$ N order?

Asking for the reason behind a certain word order pattern may appear futile in view of the types of text before us—highly stylized and metrical, partly archaizing, with contrived content, in short highly unnatural. In view of the conditions, it may be considered best not to attempt to account for word order patterns at all. Quite on the contrary, however, I think the fact that we do find a stable pattern in these texts, despite all adverse circumstances, indicates even more persuasively that this must be an actual regularity.

Note that, while adjacency of semantically connected elements can be accounted for on discourse-pragmatic grounds, relative order does not seem to have the same kind of prior motivation, at least not on a semantic level (cp. Himmelmann 1997: 142). And indeed, in the languages of the world, pronominal possessors are attested as both preceding and following the possessed noun (see Rijkhoff 2002: 197–200). In a free word order language like Vedic Sanskrit, one would expect pronominal possessives to be essentially free in their placement, just like other elements. The possibility of either pre- or postnominal placement has in fact been described for strongly non-configurational Nunggubuyu, as pointed out by Rijkhoff (2002: 200).

I propose that the surprisingly stable pattern of  $\text{GEN}_{\text{pro}}$  N order in Vedic Sanskrit can be linked to the strong tendency of Vedic demonstratives used as possessives and

(166) i. *tathaivāsīd* *vidarbheṣu* *bhīmo*  
*tathā eva āsīd* *vidarbheṣu* *bhīmo*  
*likewise* PRT *be*.IMPF.3SG *Vidarbha*.LOC.PL.M *Bhīma*.NOM.SG.M  
*bhīmaparākramaḥ*  
*bhīmaparākramaḥ*  
*great\_prowess*.NOM.SG.M  
‘Like there was among the Vidarbhans Bhīma, of terrible prowess...’

ii. *sa* *prajārthe* *param*  
DEM.NOM.SG.M *offspring\_benefit*.LOC.SG.M *utmost*.ACC.SG.M  
*yatnam* *akarot* ...  
*effort*.ACC.SG.M *do*.IMPF.3SG  
He made the utmost effort for the sake of progeny...

iii. *tam* *abhyagacchad* *brahmarṣir*  
DEM.ACC.SG.M *go\_near*.IMPF.3SG *Brahmin\_seer*.NOM.SG.M  
*damano* *nāma* ..  
D.NOM.SG.M *named*  
To him came a brahmin sage, named Damano...

iv. *taṁ* *sa* *bhīmaḥ* ... *toṣayāmāsa* ...  
DEM.ACC.SG.M DEM.NOM.SG.M *Bhīma*.NOM.SG.M *gladden*.PRF.3SG  
Him that Bhīma gladdened ...’ (MBh. 3.50.5–7, Hock 1994: 117)

(167) i. *prajāpatir* *ha vā idam agra eka evāsa /*  
*prajāpatir* *ha vā idam agra eka eva āsa /*  
*Prajāpati*.NOM.SG.M PRT PRT D *before* *one*.NOM.SG.M PRT *be*.PRF.3SG  
‘Prajapati was all alone here before.’  
*sa* *aikṣata* ...  
DEM.NOM.SG.M *consider*.IMPF.3SG  
He reflected: ...

ii. *sa* *aikṣata* *prajāpatiḥ* ...  
DEM.NOM.SG.M *consider*.IMPF.3SG *Prajapati*.NOM.SG.M  
Now, Prajapati reflected ...’



By contrast, members of the other most frequent demonstrative class, *eṣa-/eta-*stem forms, appear later, often at the end of P2, where they serve cataphoric reference (Hock 1982: 14):

- (168) *sadadī vai eṣaḥ dadāti yāḥ agnihotrām juhōti.*  
 generally PRT DEM.NOM.SG.M give.3SG REL.NOM.SG.M fire\_sacrifice offer.3SG  
 ‘He/THAT PERSON gives continuously who offers a fire-sacrifice.’  
 (MS 1.5.12, Hock 1982: 14, emphasis added)

The tendency of *sa-/ta-*stem forms to occur toward the left edge also shows in lexicalized uses which figure in the literature under the label of ‘*sa-* figé,’ that is ‘petrified *sa-*’. *Saḥ* (NOM.SG.M), and in part other inflectional forms, double as discourse linkers, making no reference to a participant in the clause. A usage at the threshold between pronominal function and that of a discourse marker is found, according to Hock (1982), in cases where *saḥ* occurs in P1 in usages parallel to those of clause-introducing elements like *átha* ‘now’, as can be seen in example (169). In a second step, *saḥ* does not have referential function any more, as in example (170).

- (169) *sá hovāca śvetáketuḥ...*  
*sá ha uvāca śvetáketuḥ*  
 DEM.NOM.SG.M PRT say.PRF.3SG Śvetáketu.NOM.SG.M  
*átha hovāca sómaśuṣmaḥ*  
*átha ha uvāca sómaśuṣmaḥ*  
 now PRT say.PRF.3SG Somaśuṣma.NOM.SG.M  
 ‘He, Śvetáketu, said..., then, Somaśuṣma said...’  
 (ŚaB 11.6.2.2–3, adapted from Delbrück 1878: 53)

- (170) *sá yáthā iva ha tát agnéḥ bhávati*  
 so as PRT PRT DEM.NOM.SG.N fire.GEN.SG.M be.3SG  
 ‘In that case, that (feud) is [= would be] (on the part) of the fire.’  
 (ŚaB 1.1.1.21, adapted from Schäufele 1991b: 73)

This anaphoric usage of *sa-/ta-*stems also concerns genitive forms encoding adnominal possessors. In the *Rigveda*, in thirty-nine out of sixty-four cases a genitive encoding a possessor stands at the beginning of a verse or at the beginning of a line break, which typically marks a syntactic division between clauses. Out of these, in turn, we find a complete group comprising the possessor and its nominal head in fifteen cases and otherwise often only a P2 element or another type of clitic intervening. Compare the following, with a continuous example in (171), and a verse with both one continuous example (but with part of the nominal expression constituting the head being placed apart) as well as a discontinuous one, where, however, only a vocative intervenes in second position, in (172):

- (171) *apām*                      *nāpātam*                      *āvase*                      *savitāram*                      *ūpa*  
 water.GEN.PL.F      son.ACC.SG.M      favour.DAT.SG.N      Savitr.ACC.SG.M      LP  
*stuhi* /  
 praise.IMP.2SG  
*tāsya*                      *vratāny*                      *uśmasi* //  
 DEM.GEN.SG.M      service.ACC.PL.N      desire.1PL  
 ‘Sing to the waters’ offspring, to Savitr, for favour! We are desiring his services.’  
 (RV 1.22.6)
- (172) *nṛcākṣā*                      *rākṣaḥ*                      *pāri* *paśya*                      *vikṣú*  
 leader’s\_stare.NOM.SG.M      demon.ACC.SG.N      LP      see.IMP.2SG      dwelling.LOC.PL  
*tāsya*                      *trīṇi*                      *prāti* *śṛṇīhy*                      *āgrā* /  
 DEM.GEN.SG.N      three.ACC      LP      break.IMP.2SG      tip.ACC.PL.N  
*tāsyāgne*                      *prṣṭīr*                      *hārasā*  
*tāsya*                      *agne*                      *prṣṭīr*                      *hārasā*  
 DEM.GEN.SG.N      Agni.VOC.SG.M      rib.ACC.PL.F      fire.INS.SG.N  
*śṛṇīhi*                      *tredhā*                      *mūlam*                      *yātudhānasya*  
 break.IMP.2SG      three\_parts      root.ACC.SG.N      demon.GEN.SG.M  
*vṛśca* //  
 cut.IMP.2SG  
 ‘With a men-leading stare, see the demon in the dwellings, break off his three spikes! Agni, break his ribs with your fury, cut into three parts the demon’s root.’  
 (RV 10.87.10)

The number of occurrences of MADHYE/UPARI with pronominal possessors is small, as shown in Table 6.3. But here, too, we find a similar distribution. Regarding MADHYE, in four out of five cases, the pronominal possessor is placed clause-initially in Vedic (in three cases, the entire group is placed so); in four out of five cases in Pali we find clause-initial placement of the group with the pronominal first, and the same in three out of four cases in Apabhramsha. For UPARI, cases with pronominal possessors are only found in Apabhramsha, where nine out of fifteen show clause-initial placement, all continuous. Compare the following illustrations with MADHYE:

- (173) *taho*                      *majjhi*                      *ṇivesiya*                      *divva*                      *devi* /  
 DEM.GEN.SG.M      MADHYE      install.PPP.SG.F      divine.SG.F      goddess.SG.F  
*pomāvai*                      *ṇāmem*                      *thira*                      *karevi*  
 Padmavati      name.INS.SG      firmly      do.CONV  
 ‘In the middle of it [= a circle] she installed firmly the divine goddess named Padmavati.’  
 (KA 7.12.4, translation Jain)

- (174) (= ex. (162))  
*tahaṃ majjhi ṇivattṭhau so sahehi /*  
 DEM.GEN.PL.M MADHYE settle\_down.PPP.NOM.SG.M DEM.NOM.SG.M shine.3SG  
*chaṇaṇḍaḥo līlā ṇaṃ vahehi*  
 full\_moon.GEN.SG beauty.ACC.SG so bear.3SG  
 ‘Sitting in **their middle**, he shines as if bearing the beauty of the full moon.’  
 (KA 8.15.3, cp. translation Jain)
- (175) *taho majjhi parittṭhiu ekku pomu /*  
 DEM.GEN.SG.M MADHYE remain.PPP.NOM.SG one lotus.NOM.SG  
*uddaḍaṇaḥim susohi nāim somu*  
 group\_of\_stars.LOC.SG adorned.NOM.SG like moon.NOM.SG  
 ‘**In the middle of it** stood one lotus, like the moon adorned by a group of stars.’  
 (KA 10.2.6)

Over time, this leftward tendency is relaxed and other elements come to be used as clause-introductory items, as observed by Canedo regarding post-Vedic Old Indo-Aryan and Middle Indo-Aryan (cp. also Bloch 1965: 309):

Now, the combination of ‘*atha kho*’ (almost exclusively) and some other particles stand in place of the age-old linkage through demonstrative pronouns (*sa-*, *ta-*, and derivations), which was predominant in the time of the Brāhmaṇas [i.e. Vedic prose, UR]. (Canedo 1937: 15–16, translation UR)<sup>37</sup>

However, the syntagm GEN<sub>pro</sub> N as such remains and shows this relative order even when not clause-initial. In addition, with Type 2 discontinuity vanishing, which concerns constructions with pronominal possessors, as well as Type 3 discontinuity receding, some of the factors which in earlier times led to discontinuity are lost, and we are left with the great majority of examples showing the pattern GEN<sub>pro</sub> N by the stage of late Middle Indo-Aryan.<sup>38</sup>

In sum, I propose that the order of  $\text{GEN}_{\text{pro}} \text{N}$  is linked to the generally strong leftward tendency in the placement of *sa-/ta-*stem demonstratives. From early on and especially after the loss of Type 2 and Type 3 discontinuity, we find at least the pronominal possessor and often the entire group  $\text{GEN}_{\text{pro}} \text{N}$  at the left edge of clauses. In the next chapter, I explore how this order spreads to constructions with all types of possessors, which I propose is triggered by specific semantic changes.

This scenario raises the question of universality. Pronominal possessors by their nature make anaphoric reference (and more rarely cataphoric reference), so that a

<sup>37</sup> 'An die Stelle der uralten Anknüpfung durch die Demonstrativpronomen (*sa-*, *ta-* und Ableitungen), die in der Brāhmaṇa-Zeit vorherrschte, tritt jetzt die Verbindung durch "*atha kho*" (fast ausschließlich) und einige andere Partikeln.'

<sup>38</sup> Thus, if not the sole factor, and certainly not a directly causal one, the loss of the Wackernagel position contributed to the rise of the postpositional phrases to a certain degree.

propensity of possessors for preceding the possessee would also be expected in other languages. However, as mentioned above, pronominal possessors are attested as occurring both to the left and to the right of possessee in the languages of the world. Still, Ultan (1978: 23) as well as Rijkhoff (2002: 307), on the basis of cross-linguistic comparison, notes a clear preference for  $GEN_{pro} N$  order over the reverse for pronominal possessors (and, less clearly so, also for nominal possessors: see also Dryer 1992: 91). Ultan adds that, whenever order is different for pronominal and nominal possessors in a language, pronominal ones always precede the possessee. Based on this evidence, I hypothesize that a language that allows for a degree of word order freedom will show  $GEN_{pro} N$  order by default, in accordance with the frequent anaphoric function of such possessors, where already topical participants are taken up early in the clause. This can also be expected if the class of possessives forms a morphosyntactic class of its own even when the language does not have free word order. Only when the marker of possession belongs to a larger class of function words which also mark other relations may it be the case that the marker of possession is syntactically constrained to follow the possessee, along with the elements with which it forms a paradigm. For instance, in German, we have a variation of *sein Buch* 'his book' vs. *das Buch von ihm* 'the book of his'. The latter example involves *von* 'of, from', constrained to occur after the modified noun (and before the noun it forms a phrase with) in the same way as other elements of the paradigm of simple prepositions. Similarly, in English, when possession is marked by a former pronominal genitive marker, we find a preposed possessor (*the boy's book* from *the boy his book*), but a postnominal one when the marker of possession is *of*, which belongs to the paradigm of prepositions, that is *the book of my grandmother*. In French we find *le livre du garçon*, involving a (former) preposition, and in German *das Buch des Jungen*, where the element in question forms part of the paradigm of articles.

## 6.4 Summary

In this chapter, I have explored the origins of the syntagmatic pattern  $GEN N$ . At first, I took up the strands of discussion in the literature on whether the existence of postpositions rather than prepositions in New Indo-Aryan should be accorded to language-internal or language-external factors. Claims that pertain to this issue typically assume a relatively stable  $GEN N$  (and OV) order in earlier Old and Middle Indo-Aryan times. The data in my corpus, however, while confirming the tendency, show that this pattern is less stable than has been assumed and, in general, it seems not enough to say that it simply became obligatory at some point. There is variation with regard both to relative order and to discontinuity even in late Middle Indo-Aryan, the fixation of which requires explanation.

Approaching the question of how discontinuity is lost, I turned to the Wackernagel position, which has often been linked to discontinuity. While it is true that

different patterns of discontinuity involve P2 in one way or another, it is not the case that the disappearance of the Wackernagel position directly entails the emergence of phrases. Instead, other types of discontinuity remain possible even when the Wackernagel position is considerably weakened in Apabhramsha and no longer impacts on the placement of nominal expressions. Thus syntactic changes cannot on their own account for the loss of discontinuity. In the next chapter, I will argue that semantic changes entail the fixation of continuity.

Subsequently, I turned to a specific subgroup of nominal dependents, pronominal possessors, which almost always show the order  $\text{GEN}_{\text{pro}} \text{N}$  when in juxtaposition from early onwards. The rare exceptions from this pattern typically involve clitic pronouns obeying placement rules of their own or other factors, such as metrical constraints. I argued that the very stable  $\text{GEN}_{\text{pro}} \text{N}$  order can be linked to the usage of genitive forms of the *sa-/ta-*stem demonstratives functioning as 3rd person pronominal possessors. This demonstrative class has a general strong tendency to appear at the left edge of the clause in anaphoric function, and the same applies to its genitive forms. Taking into account additional data sets of these pronominal possessors in Vedic and Apabhramsha, I found that they indeed almost always directly precede their nominal head and that, in addition, the group frequently stands in clause-initial position. I concluded that this sub-type of relational noun expression is an important root of the postpositional syntagm. Linking this syntagmatic pattern to discourse-functional aspects of anaphoric reference suggests a universal dimension. Indeed, quantitative typological studies suggest that there is a strong tendency to precede their nominal head for genitive possessors in general, and for pronominal possessors in particular. I proposed that this order might only be reversed in cases where the possessive forms part of a larger class of function words, which also cover other types of relations. In such cases, the paradigm of elements may be constrained to the position after the possessee, as in the case of a paradigm of prepositions or a paradigm of articles that head a possessor phrase following the possessee. In the next chapter, where I address the other major domain in the establishment of  $\text{GEN}_{\text{MADHYE/UPARI}}$ , the obligatorification of genitive dependents, I will argue that pronominal possessors play a crucial role there also.



## From group to phrase

In this chapter, I explore the transition from group to phrase. I propose that the shift of MADHYE and UPARI from autosemantic to synsemantic, that is their becoming function words, entails this change. The semantic shift takes place in two steps. First, we can observe metaphorical transfers from concrete spatio-temporal grounds to non-basic ones, that is, to such grounds as do not easily allow for a literal interpretation of IN MIDDLE<sup>1</sup> or ABOVE. Second, we see metonymic shifts of MADHYE and UPARI from encoding a complex spatial (or metaphorically spatial) relation, that is, from referring to spatial areas such as a middle (IN MIDDLE) or a space above (ABOVE), to encoding a simple spatial (or metaphorically spatial) relation, such as IN or OVER/ON. While the first development, the metaphorical expansion, begins in Vedic, the second development of metonymic shifts does not take place before the transition from late Middle to early New Indo-Aryan. These semantic changes will be shown to entail the obligatorification of the dependent genitive and the fixation of order. When both formal constraints are in place, the postpositional phrases are born, and the relational nominals have developed into postpositions.

As in the previous chapter, pronominal possessors play an important role in the development. Non-basic grounds will be shown to require the overt expression of possessors in order for the construction to be semantically interpretable. The increasing usage of such grounds across the historical stages entails an increase in overt possessors, optionally pronominally realized if activated at the respective point of discourse. Concomitantly, this leads to an increase in constructions showing the order GEN N, which is the default with pronominal possessors as shown in the

<sup>1</sup> As in Chapter 5, I use shorthands for different types of simple and complex spatial relations, namely IN MIDDLE, IN, AMONG, ABOVE, ON, and OVER. Simple spatial relations directly relate the figure to the ground (e.g. IN, ON), while complex spatial relations locate the figure in a spatial area with respect to the ground (e.g. IN MIDDLE (of sth.)). These expressions are used as corresponding to their basic usage in English, i.e. IN MIDDLE referring to a centre of some entity (e.g. *in the middle of a room*), IN expressing containment (e.g. *in the room*), AMONG referring to an entity within a group of entities irrespective of location (e.g. *the tallest one among them*), ABOVE referring to a spatial area high up in a three-dimensional space (e.g. *the bird is flying above, in the sky*), ON expressing vertical support with contact (e.g. *the cup is on the table*), and OVER expressing vertical support without contact (e.g. *the lamp hangs over the table*). Apart from these basic usages, no other meanings that these adpositions and relational nouns may have in English play a role here.

previous chapter. These developments pave the way for the reanalysis of this group as a phrase. Evidence that the metonymic shifts which turn MADHYE/UPARI from denoting complex spatial relations into denoting simple spatial relations take place precisely in this group, that is with an overt possessor which directly precedes MADHYE/UPARI, corroborates the hypothesis that grammaticalization processes pertain to constructions that have a fixed word order (see section 2.6).

This chapter is structured as follows. In section 7.1, I depict metaphorical shifts from concrete spatio-temporal to non-basic grounds, which accrue in the course of Middle Indo-Aryan. The latter, I propose, require overt possessors for the construction to be semantically interpretable. I discuss MADHYE in section 7.1.1 and UPARI in section 7.1.2. I then take a more detailed look at a subclass of non-basic grounds: animate participants, which have a special connection with periphrastic constructions of local cases (section 7.2). Subsequently, I focus on pronominal possessors and the several ways in which they are involved in the emergence of the postpositional phrases, both regarding obligatorification and syntagmatic fixation (section 7.3). I then turn to the metonymic shifts that mark the threshold from autosemantic to synsemantic (section 7.4.), first discussing MADHYE (section 7.4.1) and then UPARI (section 7.4.2). In section 7.5, I outline how the analysis presented in this chapter bears on a general theory of how grammaticalization may induce the rise of phrase structure in a strongly non-configurational language, and compare this to grammaticalization phenomena in languages that already have phrasal structures.

## 7.1 Metaphorical change and the obligatorification of possessors

While in Vedic and Pali, MADHYE and UPARI typically make reference in a literal sense to concrete spatial or temporal grounds, a steady increase in non-basic grounds can be observed towards late Middle Indo-Aryan. The original literal sense is found in those cases where a ground has an intrinsic centre, in the case of MADHYE, such as a room or a day (i.e. noon), or easily allows for the interpretation of something located above it, in the case of UPARI, for example something located in the air above some type of surface. Non-basic grounds, by contrast, consist in such entities, spatio-temporal or abstract, as lack an intrinsic centre (MADHYE) or do not lend themselves to having a spatial area vertically above them identified (UPARI). For example, regarding MADHYE, a sky does not have an intrinsic centre in the way a room does. Even harder is it in the case of abstract concepts, such as a state or activity, to identify a middle. Furthermore borderline cases, where an entity may have a centre but not necessarily so, fall on the side of non-basic grounds, as will be shown. For instance, I discuss examples below (exx. (191) and (192)) that make reference to the middle of a group of people. While a group of people may be arranged in such a way that it has a centre (e.g. when the people form a circle), they do not necessarily do so and may not form a single spatial structure at all. In cases



of non-basic grounds, overt possessors ensure semantic interpretability; covert ones imply a concrete spatio-temporal ground with an intrinsic centre (MADHYE) or with a space vertically above (UPARI).

### 7.1.1 MADHYE from Vedic Sanskrit to Apabhramsha

In my corpus of Vedic Sanskrit, (GEN +) MADHYE<sup>2</sup> identifies the centre of a spatial entity (ex. (176)), of a temporal entity (ex. (177)), or of an entity which is not primarily temporal or spatial, but which has a spatial and/or temporal dimension (ex. (178)), and in a single case of a state (ex. (179)).<sup>3</sup>

(176) VEDIC SANSKRIT

*mádhye hradásya plavasva vigṛhya catúrah padāḥ*  
MADHYE lake.GEN.SG.M swim.IMP.2SG spread.CONV four.ACC.M foot.ACC.PL.M  
‘Swim in the middle of the lake, spreading (your) four feet!’ (AV 4.15.14)

(177) VEDIC SANSKRIT

*ekaviṁśam etad ahar upayanti viṣuvantam*  
Ekaviṁśa.ACC.SG.N DEM.ACC.SG.N day.ACC.SG.N perform.3PL equinox.ACC.SG  
*madhye saṁvatsarasya*  
MADHYE year.GEN.SG.M  
‘They perform the Ekaviṁśa day, the equinox, in the middle of the year.’  
(AiB 4.18.1)

(178) VEDIC SANSKRIT

*madhya etasyai<sup>4</sup> nividaḥ sūktam*  
MADHYE DEM.DAT.SG.F short\_invocation.GEN.SG.F hymn.ACC.SG.N  
*śamset*  
recite.OPT.3SG  
‘He should recite the hymn in the middle of that invocation.’ (AiB 2.33.2)

(179) VEDIC SANSKRIT

*iṣṭásya mádhye áditir ní dhātu naḥ*  
wish.GEN.SG.N MADHYE Aditi.NOM.SG.F down put.IMP.3SG ACC.1PL  
‘Aditi should place us in the middle of (our) wish.’ (RV 10.11.2)

<sup>2</sup> Because I am dealing with the question of obligatoriness here, I consider cases with both overt and covert possessors: hence the notation of ‘GEN +’ in round brackets. While the possessor is not in all cases overt, a ground is always present in the semantic structure. Only some instances of UPARI are ambiguous between relicts of its old zero-valent usage and a relational usage with a covert possessor, as will be discussed below.

<sup>3</sup> The descriptive part of the outline of the semantic spaces of (GEN +) MADHYE in this section and in section 7.4.1 are loosely based on Reinöhl (2009). The presentation at hand differs as it is based on a larger data corpus and involves a partially revised analysis.

<sup>4</sup> The dative here is used in place of a genitive, a common usage in Vedic prose. Although not showing formal agreement, it modifies *nividaḥ* and does not encode an independent possessor.

Of these meanings, it is likely that the concrete spatial reading is the original one, as cross-linguistic evidence widely attests to the metaphorical construal *TIME IS SPACE* (e.g. Lakoff and Johnson 1980; Haspelmath 1997).<sup>5</sup> Also, there is some etymological evidence that *madhya-* may in origin have meant ‘waist’ (see section 3.1). The transfer to concepts that have a spatial or temporal extension, even if not primarily spatial or temporal as for instance in example (178), is a typical further metaphorical step. Similarly, states and activities are known to be construed as spatio-temporal entities cross-linguistically, as in (179).

MADHYE in Pali shows a similar semantic range as in Vedic, so I turn directly to Apabhramsha. For Apabhramsha, I identify four usages in my corpus. The first usage, represented by (180), covers most of the semantic space in Vedic Sanskrit where MADHYE makes reference to an actual middle in a spatial, temporal, or derived spatio-temporal sphere, i.e. involving basic grounds. Second, we find essentially the same usage, but in reference to such grounds, spatial or non-spatial, as can be *construed* as having a middle, but which lack an intrinsic centre; that is, we are in the domain of non-basic grounds (exx. (181) and (182)).<sup>6</sup> A third type is only attested once in my corpus, where MADHYE designates a time span. This example is the only case in Apabhramsha which shows a clear departure from a meaning *IN MIDDLE* and on towards expressing general containment (ex. (183)). Fourth, MADHYE can also express a relation of *AMONG* with reference to a plural-marked, individuated<sup>7</sup> referent (ex. (184)).

(180) (= ex. (173))

<i>taho</i>	<i>majjhi</i>	<i>ñivesiya</i>	<i>divva</i>	<i>devi /</i>	<i>pomāvai</i>
DEM.GEN.SG.M	MADHYE	install.PPP.SG.F	divine.SG.F	goddess.SG.F	Padmavati
<i>ñāmeṃ</i>	<i>thira</i>	<i>karevi</i>			
name.INS.SG	firmly	do.CONV			

‘In the middle of it [= of a circle, UR] she installed firmly the divine goddess named Padmavati.’ (KA 7.12.4)

(181) *annaha majjhi vasuṃdharaha / niya-nāmu vi na vahēmi*  
 otherwise MADHYE earth.GEN.SG own-name.ACC.SG PRT NEG carry.1SG  
 ‘Otherwise, in the middle of the earth, I do not carry my own name!’  
 (SA 504.7)

<sup>5</sup> See Sweetser (1988) on general theoretical aspects of a metaphorical transfer based on image schemas.

<sup>6</sup> Already ex. (179) is an early instance of a non-basic ground in Vedic. Since all other examples in Vedic clearly refer to a centre, I assume that we are dealing here with an emphatic assertion of containment by referring to a middle; see the discussion following ex. (184).

<sup>7</sup> I use ‘individuated’ in the sense that the plural-marked entity consists of individual sub-entities which can be clearly distinguished and where one or several of these sub-entities can be focused on as in ex. (184). An example for a plural-marked, but non-individuated entity are the streams in ‘in the middle of the streams’ (ex. (159)), where sub-entities are hard to identify.

- (182) *‘tuha pāyajuyalaha / patthāviṇa sēva haū /*  
 GEN.2SG foot\_pair.GEN.SG occasion.INS.SG service.ACC.SG 1SG  
*karisu’ dharivi ihu majji hiyayaha*  
 make.FUT.1SG carry.CONV DEM.ACC.SG MADHYE heart.GEN.SG  
 ‘He carried in the middle of his heart this: “I will, at the occasion, do service  
 to your pair of feet.”’ (SA 707.4)
- (183) *sō tihū diṇahā majjhe nau āvai*  
 DEM.NOM.SG.M three.GEN.PL day.GEN.PL MADHYE NEG come.3SG  
 ‘He does not come within three days.’ (BH 284.8)
- (184) *majji mahattarāṇa na kayāi vi /*  
 MADHYE best.GEN.PL NEG sometime PRT  
*vankavivanku vuccaē*  
 crooked\_contradicting.NOM.SG be\_spoken.3SG  
 ‘Among the best ones, nothing dishonest is ever spoken.’ (BH 207.17)

How do we get from the literal reference to a middle in (180) to the non-basic ground in (181)? The answer lies in pragmatics. The idea in (181) is not that an actual centre be identified, but that an emphatic expression of containment is asserted. The context for the exclamation is that a man promises to find his friend’s lover and, if unsuccessful, declares that he will not carry his own name any more on earth, that is, he will be most ashamed. It is the literal reference to a centre which lends the exclamation emphasis. A similarly emphatic expression is found in (182). A heart as a seat of emotions does not have an intrinsic middle. However, the reference to a middle conveys how deeply something is felt. Here, we encounter the oft-found construal of the human mind (or heart or soul) as a space, a metaphor of the ‘embodied mind’. Indeed, I propose that human referents, having both a spatial and an abstract reality as bodies and as minds, are an important stepping stone in the transferral from the realm of concrete spatio-temporal grounds to non-basic ones (the role of animates will be further discussed in section 7.2). When the emphatic effect wears off, the way is paved towards a literal expression of simple containment, that is, the metonymic transition from IN MIDDLE to IN. Example (183), indicating a span of time, is the earliest case foreshadowing this transition, which is completed towards early New Indo-Aryan, as will be outlined in more detail in section 7.4. In (184), too, an abstract domain is involved. The group of people is not construed as a concrete spatial entity, but someone being part of a certain group is accorded some property, encoding a meaning of AMONG. Thus, we are dealing with an abstraction from an actual geometric middle to identifying simply a sub-entity without any construal of a centre.

Along with the extension to more and more types of non-basic grounds and their increasingly frequent usage, the relative number of overt possessors grows towards late Middle Indo-Aryan (Table 7.1; selected figures from Table 5.1).

TABLE 7.1. Historical attestations of MADHYE (a)

MADHYE	alone	+GEN
Vedic Sanskrit	63	66
Pali	41	43
Apabhramsha	0	34

I claim that there is a connection between both findings—the increase in non-basic grounds and the increase in overt possessors—and propose the following hypothesis: any usage of MADHYE relating to a non-basic ground, that is to a ground which lacks an intrinsic centre, requires its overt expression. Only MADHYE expressing IN MIDDLE in relation to a spatial or temporal ground with an intrinsic centre may be used without an overt possessor. Thus, even before MADHYE shifts to profiling a simple spatial relation of IN, the genitive possessor is required in certain usages.

This hypothesis is justified as follows: If (GEN +) MADHYE refers to a non-basic ground in the sense outlined above, an overt possessor is necessary in order to ensure interpretability. This effect can also be illustrated in German with the cognate form *Mitte* ‘middle’:

- (185) *Sie spielen auf dem Fußballfeld. In der Mitte steht der Schiedsrichter.*  
‘They play on the soccer ground. In the middle, the referee stands/is standing.’
- (186) *Das Schiff segelt auf dem Meer. ?In der Mitte liegt eine Insel.*  
‘The ship is sailing on the ocean. In the middle, there is an island.’
- (187) *Es herrscht großer Wohlstand in unserer Gesellschaft. ??In der Mitte finden sich viele, die dennoch Existenzängste haben.*  
‘There is great prosperity in our society. In the middle, there are many who nevertheless live in existential fear.’

While a soccer field clearly has a centre, an ocean, while spatial, does not obviously have a centre, as does even less so an abstract concept such as a society that lacks either a spatial or a temporal dimension. Still, an ocean and even a society may be construed as having a centre—which, however, requires their overt expression, for instance as adnominal possessors. This holds even if such non-basic grounds are referred to in the previous utterance, that is, when they have a high degree of activation in discourse.

Evidence to test the hypothesis can be drawn from two domains. First, instances of MADHYE in isolation tell us about the possibilities when covert reference is possible. Second, instances where MADHYE takes a pronominal possessor give clues about where a ground which is inferable from context is nevertheless expressed overtly, even though covert reference is a syntactic possibility.

The results of the first test are clear. All of the 104<sup>8</sup> Vedic Sanskrit and Pali examples attested in my corpus where MADHYE is used without an accompanying genitive or a local case form<sup>9</sup> are references to an actual centre of some spatial or temporal ground, such as the middle of an altar, of a knot, of a mound, of a day, of an invocation, etc. In many of the examples, a centre is contrasted with corners, right and left, above and below, beginning and end, etc., which highlights the literal usage of MADHYE. Compare the following two examples for illustration:

(188) VEDIC SANSKRIT

*atha madhya āghārayati*

now MADHYE pour.3SG

‘Now, he pours (ghee) **onto the middle** (of the altar, i.e. as opposed to its corners).’ (ŚaB 3.5.2.13)

(189) VEDIC SANSKRIT

*tasyārdhāḥ śastvārdhāḥ pariśiṣya madhye*

tasya ardhāḥ śastvā ardhāḥ pariśiṣya madhye

DEM.GEN.SG.M half.ACC.PL.F recite.CONV half.ACC.PL.F leave\_over.CONV MADHYE

*nividaṃ dadhāti*

invocation.ACC.SG.F put.3SG

‘Having recited half its (verses), leaving half over, (he) places an invocation in **the middle** (of the recitation).’ (AiB 3.19)

There is a single borderline case of MADHYE in isolation referring to a ground that lacks an intrinsic centre, namely the ‘nations’ in the next example. However, in this example, MADHYE only lacks an overt possessor in the embedded participial clause, while the possessor is expressed in the matrix clause. Therefore, this example does not qualify for covert reference in a strong sense:

<sup>8</sup> RV 5.1.6, 5.44.3, 8.27.20, YV 32.2, AV 4.15.15, ŚānB 7.4a, 7.4b, 11.3, 12.3, 14.1b, 14.3, 16.7a, 18.9, 21.5, 23.8, AiB 3.10.2, 3.19.5, 3.35.6a, 3.35.6c, 3.41.4, 4.15.6, 4.15.15, 4.18.4, 4.18.5, 4.19.7, 4.23.5, 6.27.13, ŚaB 1.2.5.16a, 1.2.5.16b, 2.6.1.11b, 3.2.1.9, 3.3.2.20, 3.5.1.34, 3.5.2.13, 3.6.1.2b, 3.7.1.12a, 3.7.1.12b, 4.2.2.16, 5.2.4.5, 5.5.1.1, 6.1.1.2, 6.2.1.24a, 6.2.1.24b, 7.5.1.15, 7.5.1.23a, 7.5.2.14, 8.1.2.7, 8.1.3.6, 8.1.3.10, 8.2.4.19, 8.5.1.12, 8.6.1.9, 8.6.1.20, 8.6.1.23, 8.6.2.8, 8.6.3.4a, 8.6.3.4c, 8.6.3.11, 9.4.3.6, 11.7.3.3a, 11.7.3.3b, 13.6.1.9, 13.8.3.9, Digh II p. 235, Majjh I p. 279, II p. 21, Majjh III p. 178, Saṃ III p. 93, IV p. 179a, 180a, 180b, 181, 195, 199, Aṅg II p. 95, III pp. 399a, 399b, 399–400, 400a, 400b, 400c, 401a, Dhamm 98, 118, Iti p. 90, Sutta pp. 99, 122, 184, 199a, 202, 213b, Peta 11a, Jā IV p. 156, V pp. 233, 417, VI p. 23, Nidd I pp. 353c, 433, 445, 449, Nidd II pp. 9, 17, 35, Apa II p. 579.

<sup>9</sup> I exclude combinations of MADHYE/UPARI with local case forms from the category of MADHYE/UPARI in isolation, as before in Chapter 5, even though there is no syntactic difference. MADHYE/UPARI do not form a construction with a local case form, but stand in apposition. On the content side, however, such a local case form can in some cases encode the respective ground, i.e. show semantic dependency, e.g. *madhye samudre* (MADHYE ocean.LOC.SG.N) ‘in the middle, at/in the ocean’ is virtually commensurate with *madhye samudrasya* (MADHYE ocean.GEN.LOC.N) ‘in the middle of the ocean’. While this is not necessarily so (see the discussion of ex. (124)), such pragmatic ties can exist. Therefore, these examples are not taken into consideration in a discussion of MADHYE/UPARI occurring in isolation.

## (190) VEDIC SANSKRIT

yúvā kavīḥ puruniṣṭhā ṛtāvā  
 youthful.NOM.SG.M sage.NOM.SG.M excelling.NOM.SG.M proper.NOM.SG.M  
 dhartā kṛṣṭínām utá mādhyā iddhāḥ  
 supporter.NOM.SG.M people.GEN.PL.F and MADHYE inflame.PPP.NOM.SG.M  
 ‘(Agni,) the youthful sage, who excels and is proper, preserver of the nations,  
 inflamed in (their) **middle**.’ (RV 5.1.6)

I turn to the second test for the hypothesis put forward, which relates to constructions involving pronominal possessors. In some cases, a pronominal encodes a ground with an intrinsic centre. Even though semantically recoverable, this may be triggered by a weak activation status. Certain such factors on the level of information structure will be discussed below in section 7.3. In other cases, however, we see usages where the pronoun refers to entities that lack an identifiable centre, a usage never found with MADHYE in isolation, which is further confirmation for my hypothesis. Compare the following illustrations (see too ŚaB 8.6.3.4, Sutta p. 60, Apa I p. 326, or Buddha p. 44):

## (191) VEDIC SANSKRIT (= ex. (140))

dāsyāḥ putraḥ kitavo ’brāhmaṇaḥ  
 slave.GEN.SG.F son.NOM.SG.M cheat.NOM.SG.M non\_Brahmin.NOM.SG.M  
 katham no madhye ’dikṣiṣṭeti  
 katham no madhye adikṣiṣṭa iti  
 how GEN.1PL MADHYE consecrate.AOR.MID.3SG QUOT  
 ‘The child of a slave woman, a cheat, no Brahman; how has he been  
 consecrated in our middle?’ (AiB 2.19.1, cp. Keith’s translation)

## (192) APABHRAMSHA (see (162))

tahim jāivi ṭiṇṭahim gau turantu /  
 DEM.LOC.SG go.CONV gambling-house.LOC.SG go.PPP.NOM.SG.M in\_haste  
 jūvārahaṁ sayalahaṁ maṇu harantu / taham  
 gambler.GEN.PL all.GEN.PL mind.ACC.SG grab.PPA.NOM.SG.M DEM.GEN.PL  
 majjhi ṇivittṭhau so sahei /  
 MADHYE settle\_down.PPP.NOM.SG.M DEM.NOM.SG.M shine.3SG  
 chaṇaindaho līlā ṇam vahei  
 full\_moon.GEN.SG beauty.ACC.SG so bear.3SG  
 ‘Reaching there, (the king) quickly went to the gambling house attracting the  
 mind of all the gamblers. Sitting **in their middle**, he shines as if bearing the  
 beauty of the full moon.’ (KA 8.15.3, cp. translation Jain)

In both these examples, MADHYE refers to a group of people, that is a non-basic ground, whether or not arranged in a constellation with an actual middle. The pronominal possessor, however, ensures interpretability despite this lack of an

intrinsic centre. If the pronominal possessor was lacking, another interpretation, that of referring to a basic spatio-temporal entity, would be the default. This can be inferred from the fact that all 104 examples of *MADHYE* in isolation involve basic grounds with intrinsic middles. For instance in the example (192), if there were no pronominal, *MADHYE* could in theory be referring either to the middle of the house or to the middle of the group of people. Based on the evidence, it seems safe to assume that the former interpretation would be the default.

### 7.1.2 *UPARI* from Pali to Apabhramsha

I discussed the semantics of (GEN +) *UPARI* in Pali in Chapter 5. At that historical stage, we find the two meanings of *ABOVE* with or without contact in relation to an object or surface (cp. Brugman 1983 and Lakoff 1987 on English *over*, cognate with *UPARI*):

(193) (= ex. (121))

<i>Saccakassa</i>	<i>Nigaṇṭhaputtassa</i>	<i>upari</i>	<i>vehāsaṃ</i>
Saccaka.GEN.SG	son_of_Jains.GEN.SG	UPARI	air.ACC.SG
<i>ṭhito</i>	<i>hoti</i>		
stand.PPP.NOM.SG.M	be.3SG		

‘The Yakṣa came to stand in the air above Saccaka, the son of Jains.’

(Majjh I, p. 231)

(194) (= ex. (96))

<i>seyyathāpi...</i>	<i>payasotattassa</i>	<i>nibbāyamānassa</i>	<i>upari</i>	<i>santānakam</i>
just_as	boiled_milk.GEN.SG	cool.PPM.GEN.SG	UPARI	scum.NOM.SG
<i>hoti</i>	<i>evam evaṃ</i>	<i>pātur</i>	<i>ahosi</i>	
become.3SG	just so	manifest	be.AOR.3SG	

‘Even as scum forms on top of boiled milk that is cooling, so did (the earth) appear.’

(Digh III, p. 85)

In Apabhramsha, *UPARI* still denotes relations of *ABOVE* in relation to a spatial entity (ex. (195)). Additionally, it has expanded to also cover cases of non-vertical contact, as in (196) (this example being on the verge between spatial and abstract interpretation). Furthermore, we find an abstract usage where *UPARI* marks a stimulus (ex. (197), cp. also (129)), a usage especially frequent in the corpus (see section 5.4). Similar to example (182) above, this is a metaphor where a human mind is construed as a spatial entity. Note that some examples in Apabhramsha already transgress the line regarding the transition from encoding a complex to a simple spatial relation, mirrored in the occasional translation by ‘over’ or ‘on’. However, *UPARI* can still be used in isolation and, when used spatially, typically continues to denote a space high up in relation to an object or surface. Thus, the shift to encoding a simple spatial relation is not yet completed.

- (195) *taho layaṇaho uppari girivarammi*  
 DEM.GEN.SG cave.GEN.SG UPARI great\_mountain.LOC.SG  
 ‘above the cave, on the great mountain’ (KA 4.4.4)
- (196) *campāhiu calliu taho uvāri*  
 king\_of\_Campa.NOM.SG.M move.PPP.NOM.SG.M DEM.GEN.SG UPARI  
 ‘The king of Campa marched against it [= a hostile army].’ (KA 4.1.13)
- (197) *mā tila-tusa-tibbhāgiṇa vi / mah’ uvāri*  
 NEG seed-chaff-third\_part.INS.SG like GEN.1SG UPARI  
*kōvu karēsu!*  
 anger.ACC.SG make.IMP.2SG  
 ‘Don’t be angry with me, not (even) by the third part of the chaff of a seed!’  
 (SA 522.9)

Now, as in the case of MADHYE, in parallel to the increase in non-basic grounds, we find a relative rise in overt possessors (Table 7.2; selected figures from Table 5.2).

Against this background and in parallel with the hypothesis formulated for MADHYE, I propose that UPARI allows for a covert ground only when literally meaning ABOVE something else with or without contact.<sup>10</sup> Most examples of UPARI in isolation corroborate this hypothesis, as do the following two examples:<sup>11</sup>

- (198) *bhīsaṇa-rayanihim bhīsaṇa aḍai / khāi va*  
 fear-night.LOC.SG fearful forest.NOM.SG as\_if PRT  
*gilai va uvāri va paḍai*  
 swallow.3SG PRT UPARI PRT fall.3SG  
 ‘In the horrible night, the forest was terrible, as if swallowing or falling on top of (you).’ (PA 19.3.2)

TABLE 7.2. Historical attestations of UPARI (a)

UPARI	alone	+GEN
Vedic Sanskrit	29	0
Pali	11	9
Apabhramsha	5	35

<sup>10</sup> In contrast with MADHYE, UPARI, when used literally, does not have an application to the temporal dimension.

<sup>11</sup> Recall that, even though these examples are already translatable by ‘on’, UPARI still requires in this period, when used in its literal spatial sense, a rise along the vertical axis. Thus, it does not occur in usages such as ‘(There is water) on the floor’. This is one reason why I do not consider it as yet expressing a simple spatial relation at this point.



- (199) *jahiṃ kuhiniu ravikanta-ppahau / sihi-saṃkae*  
 REL.LOC.SG street.NOM.PL sun\_crystal-ray.NOM.PL fire-fear.INS/LOC.SG  
*uvari ṇa dei pau*  
 UPARI NEG give.3SG foot.ACC.SG  
 ‘Its [= the monkey island’s] foot paths (were illumined [sic!] like) the gems  
 issuing the dazzling rays of the sun; no one placed his foot **on top of** (them)  
 for fear of burning.’ (PA 6.6.2, translation Prasad)

In the remaining cases, we also see UPARI referring to a space above, but it is not clear whether in relation to some ground. Instead, we sometimes appear to be dealing with vestiges of the old zero-valent UPARI. For instance, example (200) is ambiguous between an interpretation of ‘above, i.e. somewhere around the upper body’ and ‘above/on the petticoat’. In any case, whether it refers to a ground as in the examples above or not, it encodes a meaning ABOVE in a three-dimensional space; that is, it still shows its original concrete meaning:

- (200) (= ex. (151))  
*vaddhu kaḍilli alakkhiyanāmaū / uppari*  
 bind.PPP.NOM.SG hip\_cloth.LOC.SG petticoat.NOM.SG UPARI  
*pīṇiu rasaṇā-dāmaū*  
 buckle\_around.NOM.SG rope-garland  
 ‘A petticoat is bound to the hip-cloth; **above**, a garland is buckled around.’  
 (BH 167.2)

Regarding usages with pronominal possessors, ten out of the fifteen attested examples involve a derived meaning and only five express ABOVE in relation to an object or surface.<sup>12</sup> The ten derived usages involve non-vertical contact<sup>13</sup> twice (cp. ex. (196)), and in eight cases a stimulus of a feeling or action, as in (197)<sup>14</sup> and as shown also in examples (201) and (202). While being low in absolute terms, this evidence, too, supports the hypothesis that derived usages necessitate overt possessors.

- (201) *so tumhahaṃ uvari ṇivaddhaṇehu / saṃcallahu*  
 DEM.NOM.SG.M GEN.1PL UPARI bound\_affection.NOM.SG.M move.2PL  
*tumhaim tāsu gehu*  
 NOM.2PL DEM.GEN.SG.M house.ACC.SG  
 ‘He has his affection fixed **on you**. You come to his house.’ (KA 7.6.7)

<sup>12</sup> KA 2.19.10, 5.4.7, 5.13.6, 9.15.5, SA 650.5.

<sup>13</sup> KA 3.13.8, 4.1.13.

<sup>14</sup> BH 222.3, 228.8, KA 4.3.8, 7.6.7, 7.11.13, SA 489.5, 522.9, 598.1.

- (202) *maho uvāri bhaḍārā ṇaravarasārā karuṇa kari*  
 GEN.1SG UPARI lord.VOC.SG.M best\_of\_men.VOC.SG.M pity.ACC.SG.F do.IMP.2SG  
 ‘Have pity on me, oh my master, the best of men.’ (KA 7.11.13)

## 7.2 Animacy and localization

Before I proceed to the metonymic shifts, I first go into more detail regarding two subgroups of dependents that stand out: namely constructions with animates, and with pronouns. I focus on animates here and turn to pronouns in the next section. Fahs notes for Pali that animates (as well as pronouns) are especially frequent in periphrastic constructions of case:

The periphrastic expression of case is used much more often with animates and pronouns than for things (cp. the usage in Latin of simple cases with things, but of prepositions with persons, e.g. instrument-marking ablative with things, but *per* + Acc with persons). (Fahs 1989: 105, translation UR)<sup>15</sup>

Compare the figures for animates featuring in constructions of GEN + MADHYE/UPARI in Tables 7.3 and 7.4.

The number of animates is low for Vedic MADHYE. In Pali and Apabhramsha, however, about half of the examples involve animates, and the same goes for UPARI.

TABLE 7.3. Historical attestations of MADHYE (b)

MADHYE	+GEN	of which animate
Vedic Sanskrit	66	4
Pali	43	25
Apabhramsha	34	16

TABLE 7.4. Historical attestations of UPARI (b)

UPARI	+GEN	of which animate
Pali	9	4
Apabhramsha	35	16

<sup>15</sup> ‘Die Kasusumschreibung wird bei Lebewesen und beim Pronomen viel häufiger verwendet als bei Sachen (vgl. im Lateinischen die Verwendung von einfachen Kasus bei Sachen, aber von Präpositionen bei Personen, z.B. instrumentaler Ablativ bei Sachen, aber *per* + Akk. bei Personen).’

These findings are in accordance with Fahs' observation. On the one hand, this may simply be accounted for by my claim that non-basic grounds require overt possessors. Since people generally talk about animates and especially about humans a lot, typically conceptualized as non-basic grounds,<sup>16</sup> they are bound to figure prominently as overt possessors in the corpus. There is another reason, however, why animates specifically figure in periphrastic expressions of case, rather than being encoded by morphological case forms such as the locative or ablative, even though the latter are still functional in Pali and Apabhramsha (if not formally distinctive in each declensional class).

What we are observing here is a watershed between animates and inanimates in what concerns the formal encoding of semantic roles: while animates and especially humans are typical agents or patients, they are untypical instruments, places, goals, sources, etc. (cp. Silverstein 1976 for the animacy hierarchy, and Levin and Rappaport Hovav 2005: 154ff on versions of the semantic role hierarchy). On the formal side, this is matched by animates frequently figuring in core case forms, which typically encode high semantic roles, while they are infrequently encoded in local case forms. When however used in a low ranking semantic role, more lexically concrete expressions, which give more semantic detail tend to be used, rather than morphological case forms. This ensures interpretability in view of the unusual usage of a highly animate participant in a semantically low-ranking role.

While the general increase in animates can be affirmed for GEN + MADHYE/UPARI, it is beyond the scope of this study to detect systematic constructional choices of the kind Fahs mentions for Latin, as this would require a larger number of attestations.<sup>17</sup> Nonetheless, there are a few usages of GEN MADHYE/UPARI which almost solely involve animates in my corpus and are thus good candidates for such a constructional choice.

For instance, UPARI encodes a stimulus of a feeling typically when this is directed at a person rather than at an inanimate object. In particular, UPARI in most cases expresses strong feelings of anger or love.<sup>18</sup> Consider once more the following example:

(203) (Apabhramsha, = ex. (197))

<i>mā</i>	<i>tila-tusa-tibbhāgiṇa</i>	<i>vi /</i>	<i>mah'</i>	<i>uvari</i>
NEG	seed-chaff-third_part.INS.SG	like	GEN.1SG	UPARI

<sup>16</sup> An exception is a usage where a body is viewed as an object, above which something is located, e.g. ex. (121).

<sup>17</sup> Another example from Latin is *cum* 'with' which marks comitative role with humans and *per* 'through, by' which encodes human agents in passive constructions, whereas both semantic roles are expressed by the morphological ablative when involving inanimates (Hofmann 1965: 114).

<sup>18</sup> In these usages, we can also already see one of the several typical domains where pronominals are used, discussed in the next section. Expressions of strong feelings often involve speech act participants, pronominally encoded.

*kōvu karēsu!*  
 anger.ACC.SG make.IMP.2SG  
 ‘Don’t be angry **with me**, not (even) by the third part of the chaff of a seed!’  
 (SA 522.9)

Another such special usage appears to hold for MADHYE in the meaning of AMONG; in this meaning it always refers to animates in my corpus. This is expected if we think of the fact that AMONG selects an individual entity as contrasting with other members of the group it is a member of, individuated members tending to be animate, cp. example (204):

(204) PALI  
*na tassa hetu visamaṃ careyyaṃ /*  
 NEG DEM.GEN.SG.M because misconduct.ACC.SG go.OPT.1SG  
*majjhe sivīnaṃ usabhomhi jāto*  
 MADHYE devotees\_of\_Shiva.GEN.PL bull.LOC.SG live.PPP.NOM.SG.M  
 ‘I, born a leader [= born in form of a bull] **among the Shiva-devotees**, would not commit a wrongdoing for that reason.’  
 (Jā V, p. 223)

(205) (= ex. (184))  
*majjhi mahattarāṇa na kayāi vi /*  
 MADHYE best.GEN.PL NEG sometime PRT  
*vankavivanku vuccaē*  
 crooked\_contradicting.NOM.SG be\_spoken.3SG  
 ‘**Among the best ones**, nothing dishonest is ever spoken.’ (BH 207.17)

In sum, enquiring into the various types of non-basic grounds, one can identify a special sub-class, namely animate (and typically human) participants. Fahs’ observation that there is an increasing correlation between periphrastic constructions of case and reference to animates is supported by evidence of MADHYE/UPARI. Accounting for this finding, I identify a special propensity for periphrastic expressions of local cases, because animates are untypical locations, instruments, stimuli, etc. Rather than being encoded by local case forms, more elaborate periphrastic constructions are used that ensure semantic interpretability. While the empirical basis does not allow for observations of structural choices such as those that have been proposed for Latin, for instance, there are certain usages that appear only or almost only with animates in my corpus. This subgroup of non-basic grounds, I propose, with its correlation with periphrastic expressions of local cases, was an important factor in the proliferation of non-basic grounds from early Middle Indo-Aryan onwards.

### 7.3 The role of pronominal possessors

In the quotation by Fahs cited in the previous section, not only animates, but also pronouns are observed to appear especially often in periphrastic expressions of case.

This has also been noted for early New Indo-Aryan Old Awadhi: ‘In the old texts we find postpositions employed oftener [sic!] with pronouns than with nouns’ (Saksena 1971: 214). Saksena adds that one and the same case ending, oblique *-hi* (an old genitive formation) appears distinctive enough for nouns, but pronouns in *-hi* often receive additional support from a postposition (1971: 214–15).

This effect is the expected consequence of the hypothesis formulated and substantiated with data in section 7.1, that non-basic grounds require overt possessors. Activated concrete spatial or temporal grounds which allow for a literal reading of IN MIDDLE or ABOVE do not necessarily require an overt possessor. Non-basic grounds, by contrast, require one to ensure interpretability. Over time, their usage becomes increasingly frequent and, when activated at a given moment, can be expressed by a pronominal possessor.

I here take a closer look at the phenomenon. While the number of examples is small, an increase in the usage of pronominal possessors can indeed be observed at least for UPARI, from no attestation in Pali to almost half of the cases in Apabhramsha (Tables 7.5 and 7.6).

At first glance, the observation that pronominal possessors are especially frequent in these constructions might appear to conflict with the cross-linguistic finding that (synthetic) pronominal case forms often survive for a longer time than inflected nominal forms. For instance, whereas English still has small case paradigms of inflectional pronominal forms, case distinctions are lost in the domain of nouns and adjectives. Thus, one might expect that synthetic pronominal forms are quite resilient, and not quickly replaced by periphrastic expressions involving genitive pronouns.

TABLE 7.5. Historical attestations of MADHYE (c)

MADHYE	alone	+GEN	of which GEN <sub>pro</sub>
Vedic Sanskrit	63	66	8
Pali	41	43	7
Apabhramsha	0	34	4

TABLE 7.6. Historical attestations of UPARI (c)

UPARI	alone	+GEN	of which GEN <sub>pro</sub>
Pali	11	9	0
Apabhramsha	5	35	15

The phenomenon can be accounted for by the theory of Preferred Argument Structure which states that, in natural discourse, at most one argument is typically expressed lexically, while other arguments are expressed pronominally, continuing already activated participants (Du Bois 1987). Thus, pronouns are highly frequent in core case forms, encoding core participants, much more so than nouns. This high frequency is an important factor in the longer survival of core case distinctions in pronominal paradigms (see Detges 2009). By contrast, pronominal forms of peripheral cases are much less frequent and are lost earlier. Here, productive formation of periphrastic expressions using pronominal (genitive-marked) possessors is found, which take over the functions of formerly used synthetic forms.

There is a gap in my corpus bearing on the topic of pronominal realization, namely that there is not a single example where *MADHYE/UPARI* occur in apposition with a pronominal *local* case form. When combining with a pronominal form, it is always a genitive. I suggest that this gap is due to the primarily discourse-functional purposes of pronouns, situating referents in the discourse world. In the case of *MADHYE/UPARI* and their possessors, the pronominals identify the former as some already activated discourse referent. This structural relation is mirrored in the morphosyntactically and semantically dependent role of the adnominal genitive, in contrast with a morphosyntactically independent and semantically charged *local* case form. A frequent pattern we find, which links back to the previous section on animates, is that a certain animate or even human participant is introduced in a first utterance and is encoded as a location with *MADHYE/UPARI* in the next utterance. For instance, some of the examples involving *MADHYE* expressing a relation of *IN MIDDLE* with pronominal possessors relate to the middle of a circle of people. A typical case in the Pali scriptures involves the Buddha or another venerable person speaking to a congregation of monks or nuns while standing or sitting ‘in their middle’ (e.g. Sutta p. 60).

In sum, pronominal possessors are relevant for the emergence of the postpositional phrases on several levels. In the previous chapter, I pointed out a preference for the pattern *GEN<sub>pro</sub> N*. In the present chapter, I have outlined in section 7.1 how they grow more frequent as an effect of the increase in non-basic grounds. When activated, a pronominal possessor may be chosen. A particular correlation holds with the special subgroup identified in the previous section—animate participants, which require more semantic detail when figuring as locations, for instance, and therefore often figure in periphrastic expressions of case rather than in the basic inflectional forms. In such cases, the correlation with pronominal realization is high, for 1st and 2nd person forms, but strong also for 3rd person animates. Last but not least, whenever *MADHYE/UPARI* combine with a pronominal form, it is always a genitive and never a *local* case form. This shows the pronominals’ function of reference tracking, rather than of independent reference, in contrast with nominal *local* case forms. On all these levels, functional and formal, constructions with pronominal possessors are a driving force in the development of the postpositional phrases.

## 7.4 From complex to simple spatial relations

I return to the question of the covert vs. overt expression of possessors. Whereas in Vedic a ground with an intrinsic centre can remain unexpressed, Hindi *mē* has a slot for an oblique-marked argument which must be filled, no matter how easily inferable. Thus, what used to be a semantic constraint restricted to non-basic grounds has become a syntactic constraint. Compare the following examples from Vedic and Hindi which both involve activated, basic spatio-temporal grounds. In such cases, covert reference is possible in Vedic, but impossible in Hindi:

(206) Vedic Sanskrit (= ex. (3))

*átha somopanáhanasya samutpāryántān*  
*átha somopanáhanasya samutpārya ántān*  
 now Soma\_cloth.GEN.SG.N gather.CONV end.ACC.PL.M  
*uṣṇīṣeṇa vígrathnāti... atha madhye 'ṅgūlyākāśam*  
 head\_band.INS.SG tie\_together.3SG now MADHYE finger\_hole.ACC.SG.M  
*karoti*  
 do.3SG

‘Having gathered up the ends of the Soma-cloth, he ties them together with a head-band... he then makes a finger-hole **in the middle** (of the knot).’

(ŚaB 3.3.2.19–20)

(207) Hindi (= ex. (5))

*hamār-e ghar ke pās nadī hai*  
 POSS.1PL-OBL.SG.M house.OBL.SG POSS proximity river.DIR.SG.F be.3SG  
*us /\*Ø mē tairnā acchā lagtā hai*  
 DEM.OBL.SG in swim.INF.DIR good.DIR.SG.M strike.IPFV.SG.M be.3SG

‘There is a river near our house; it’s nice to swim **in it**.’

Before we reach the New Indo-Aryan state of affairs, in Apabhramsha, even though MADHYE/UPARI already often combine with non-basic grounds that require overt possessors, they have not yet become adpositions—not formally, but not semantically either. At this point, they do not yet obligatorily require overt possessors and they also still show their basic meanings of IN MIDDLE and ABOVE in the majority of cases. In the next two sections, I explore their shift into elements denoting simple spatial relations and the way in which this completes the obligatorification of possessors as well as their syntagmatic fixation. These formal effects will be shown to go hand in hand.

### 7.4.1 MADHYE from Apabhramsha to Hindi

In Hindi, the meaning of IN MIDDLE is lost, and *mē* primarily expresses simple spatial or temporal containment and also encodes different kinds of abstract containers such

as minds or states: see examples (208)–(211).<sup>19</sup> A special usage that builds on the image schema of containment is that of marking a standard of comparison: see (212). In the temporal domain, *mē* in combination with punctual verbs can now mark the endpoint of a period of time elapsed, as illustrated in example (213). To this usage I connect an idiosyncratic expression where *mē* also marks an endpoint of a certain space (i.e. amount of money) covered, that is where it is used in naming a price: see (214). Reference to containers has also further developed with *mē* marking spheres of interest or objects of feeling as stimuli in contexts where predicates trigger such interpretations: see (215). Finally, the meaning of AMONG already identified in Apabhramsha is found in Hindi as well, as shown in (216).

- (208) *maī śahar mē rahtī hū*  
 1SG city.OBL.SG in live.IPFV.SG.F be.1SG  
 ‘I live **in the city**.’
- (209) *garmiyō mē maī pahār par jātā hū*  
 hot\_season.OBL.PL in 1SG mountain.OBL.SG on go.IPFV.SG.M be.1SG  
 ‘I go to the hills **in the hot season**.’ (Sandahl 2000: 32)
- (210) *netā mē himmat nahī hai*  
 leader.OBL.SG.M in courage.DIR.SG.F NEG be.3SG  
 ‘The leader has no courage.’ /lit. ‘In the leader, there is no courage.’  
 (Sandahl 2000: 30)
- (211) *vah pareśānī mē hai*  
 DEM.DIR.SG trouble.OBL.SG.F in be.3SG  
 ‘S/he is **in trouble**.’
- (212) *merā bhāī parivār mē sab se choṭā hai*  
 POSS.1SG.M brother.DIR.SG.M family.OBL.SG in all ABL small.DIR.SG.M be.3SG  
 ‘My brother is the youngest **in the family**.’
- (213) *maī ek ganṭe mē āuṅgā*  
 1SG one hour.OBL.SG in come.FUT.1SG.M  
 ‘I will come **in an hour**.’
- (214) *pacās rupaye mē to bahut sastī hai kamīz*  
 fifty rupee.OBL.PL in FOC very cheap.SG.F be.3SG shirt.SG.F  
 ‘For 50 rupees, this shirt is indeed very cheap.’ (Agnihotri 2007: 144)

<sup>19</sup> Remarks on the semantic range of *mē* in Hindi can be found in Kellogg (1972 [1875]), Singh (1977), or Khan (2009). The last offers an overview of spatial markers in several modern Indo-Aryan languages. A quantitative-typological study of the grammaticalization of spatial markers is Svorou (1988). See also Yamaguchi (2004) on adpositions in typological perspective.



- (215) *vah paṇḍit ke baccō mē bahut*  
 DEM.DIR.SG Pandit.OBL POSS.PL.M child.OBL.PL.M in very  
*dilcaspī letī hai*  
 interest.DIR.SG.F take.IPFV.SG.F be.3SG  
 ‘She takes great interest in the Pandit’s children.’ (cp. Sandahl 2000: 32)
- (216) *ham mē kaun hai?*  
 1PL in who be.3SG  
 ‘Who is there among us?’ (Kellogg 1972 [1875]: 419)

The range of OBL *mē* in Hindi having been illustrated, what takes us from Apabhramsha to this current state of affairs? In order to pin down what has changed, I contrast the semantic spaces in Apabhramsha and Hindi in Table 7.7. To simplify the exposition, I use abridged English translations of the Apabhramsha examples in section 7.1.1 for the first three usages and of the Hindi examples above for the latter five usages.<sup>20</sup>

TABLE 7.7. Semantic space of MADHYE in Apabhramsha and Hindi

	IN MIDDLE	IN	
Apabhramsha	‘he is sitting <b>in the middle of the circle</b> ’	‘he is sitting <b>in the circle</b> ’	Hindi
	‘he proclaims <b>in the middle of the world</b> ’	‘he proclaims <b>in the world</b> ’	
	‘he carries it <b>in the middle of his heart</b> ’	‘he carries it <b>in his heart</b> ’	
	‘s/he is <b>in the middle of trouble</b> ’	‘s/he is <b>in trouble</b> ’	
	? ‘my brother is the eldest <b>in the middle of the family</b> ’	‘my brother is the eldest <b>in the family</b> ’	
	* ‘ <b>in the middle of the leader</b> , there is no courage’	‘ <b>in the leader</b> , there is no courage’	
	* ‘she takes great interest <b>in the middle of his children</b> ’	‘she takes great interest <b>in his children</b> ’	
	* ‘I will be ready <b>in the middle of an hour</b> ’	‘I will be ready <b>in an hour</b> ’	

<sup>20</sup> The fourth example ‘in (the middle of) trouble’ representing localization in a state is not attested in Apabhramsha, which is why I use an example adapted from the one in Hindi. I expect, however, based on the Vedic ex. (179) ‘placed in the middle of a wish’ that a spatial conceptualization of a state of mind or similar condition was possible also in Apabhramsha. I exclude from the schema the naming of prices and disregard for the moment the usage of AMONG.

As emerges from the table, several usages are compatible with IN MIDDLE as well as IN and are also attested with the respective meaning in both historical stages. This covers spatio-temporal containers with intrinsic centres (e.g. a circle), without intrinsic centres (e.g. the world), embodied mind metaphors (e.g. the heart), and also localization in states (e.g. in trouble). Other usages in Hindi, however, are only compatible with an interpretation as IN. This involves the usage as a standard of comparison, the attribution of a property, the identification of topics of interest, and the identification of an endpoint of a time span. These usages, then, are not attested in Apabhramsha.<sup>21</sup> Logically, the shift from IN MIDDLE to IN must have preceded any extensions into abstract domains incompatible with the identification of a centre (incompatible meanings are shaded grey in the table).

I suggest that the semantic shift from Apabhramsha to Hindi proceeded along two paths, one taking off with IN MIDDLE in reference to a non-individuated entity (singular or plural), which is at the origin of almost all usages in Hindi illustrated above, the other taking off with IN MIDDLE in reference to a plural-marked, individuated entity giving rise to the remaining usage of AMONG. I only outline the first in detail; the second is already attested in Apabhramsha and was commented on in section 7.1.1.

I have already described the transition from a literal reference to a centre to an emphatic assertion of containment in section 7.1.1. Now, an increase in such emphatic usages eventually leads to the disappearance of the semantic feature of a centre, which leaves us with an expression of simple containment, an early attestation of which is example (183). Reference can then be made even to such containers as do not even allow the metaphorical construal of a middle, such as standards of comparison, possessors of a certain property, or objects of interest. A special usage arises in those cases where we are dealing with a temporal stretch of time combining with a punctual verb, where an endpoint is focused on. If someone arrives within a certain period, he or she will have arrived by the end of it. This semantic shift can also be found in English, e.g. *I will arrive in three hours*.

What about the formal side of the process? As a result of the semantic change to synsemantic, MADHYE changes into an adposition, as it cannot be used with an unfilled argument slot any more. Regarding word order, I proposed in section 2.6 that grammaticalization processes pertain to syntagmatically fixed constructions. This means that the shift of MADHYE from IN MIDDLE to IN—the crucial shift from autosemantic to synsemantic—could only have occurred in the construction GEN N. Corroborative evidence that this indeed is the case comes from sixteenth-century Old Awadhi. Here, MADHYE has come to denote the simple spatial relation IN, *nagara mahā* ‘in the city’, *mana mahā* ‘in the mind/by oneself’, etc. (see Reinöhl 2009 for

<sup>21</sup> Since it is only ex. (183) which suggests an actual shift to IN in Apabhramsha, I exclude it from consideration for the schema in Table 7.7.

details). At the same time, despite the strict metrical constraints in the *Rāmācarita-mānasa*, which might manipulate order, we almost always find the pattern OBL MADHYE (the oblique being the syncretic successor of the genitive and other case forms), namely in 278 out of 282 cases. The four cases with reversed pattern are shown below—the string *mājha ura* is found in two more passages apart from the example cited (6.93, 6.97):

- (217) *bharata bacana suni mājha tribenī*  
 bharata.OBL speaking hear.CONV MADHYE Triveni.OBL  
 ‘having heard Bharata’s words in the middle of the Triveni’ (RA 2.205)<sup>22</sup>
- (218) *mājha ura lāgā / maratī bāra kapaṭu*  
 MADHYE chest.OBL strike.PFV dying.IPFV moment fraudulent\_intent  
*saba tyāgā*  
 all give\_up.PPP  
 ‘(He defiantly fitted an arrow to his bow and shot with such steady aim that) it struck (Meghananda) full in the breast and (the demon) abandoned all false appearances at the moment of death.’ (RA 6.75, translation Prasad)

In these examples, reference is made to an actual centre; we witness vestiges of the old meaning IN MIDDLE before the metonymic shift. A voice emerges in example (217) not simply from somewhere in the water, but from right in the middle of the confluence of the Ganges, Yamuna, and Sarasvati, a location highly sacred in Hindu belief. Similarly, someone is not only struck somehow by an arrow in example (218), but the bow is aimed so precisely that he is struck right into the middle of the chest. Thus, the only four instances of MADHYE OBL in the entire set of examples are precisely such examples as do not show the new usage as a simple spatial relation meaning IN, in contrast with the many cases of OBL MADHYE where MADHYE does so.<sup>23</sup>

This supports the prediction that phrasal constructions only emerge from specific groups, that is groups with a specific word order. The increase in overt possessors and specifically of the subgroup of pronominal possessors had already led to an increase in the pattern GEN (> OBL) MADHYE/UPARI in the course of Old and Middle Indo-Aryan. And it is only in this pattern that the metonymic shift is completed, leading from autosemantic to synsemantic and entailing the reanalysis of the group as a phrase.

<sup>22</sup> The examples from the *Rāmācaritamānasa* shown here are cited after [http://gretil.sub.uni-goettingen.de/gret\\_utf.htm#TulRcm](http://gretil.sub.uni-goettingen.de/gret_utf.htm#TulRcm), last accessed 9 October 2013.

<sup>23</sup> This functional split between the two patterns also has a phonological correlate. *Mājha* is comparatively close to the older sound shape of MADHYE, while other examples show more reduced sound shapes, such as *mahā*. Phonetic erosion is one of the typical correlates of grammaticalization. However, it does not seem to be the case that the different phonological shapes indeed reflect different functions in this case (see Reinöhl 2009). Strnad (2013: 347) counts as many as sixteen versions of MADHYE in the poems of Kabir. See section 1.2.3 on the factors regarding borrowings as well as orthographic variation in these periods.

7.4.2 *UPARI from Apabhramsha to Hindi*

Post-Vedic *UPARI* had syntactically adapted to relational nouns by acquiring a slot for a genitive, but it did not acquire ‘nominal’ semantics, for example turn into a noun that means ‘top’ or ‘upper side’. However, the metonymic shift from profiling a region to profiling a basic spatial relation can be observed here also.

Hindi *par* extends some old usages and has also entered new semantic domains (cp. Kellogg 1972 [1875]: 421–2 and Khan 2009: 29 for descriptions). An extension from the meaning of *ABOVE*, it can mean *ON* or *OVER* in relation to a spatial object or surface (ex. (219)), whether raised or not (ex. (220)), and now also expresses non-vertical contact in a concrete spatial sense (ex. (221)). As already in Apabhramsha, it can encode a stimulus of a feeling (ex. (222)). It is used in reference to a point at a certain spatial distance (ex. (223)) and a point in time (ex. (224)). Finally, it can also mark possession in some contexts (ex. (225)) and figures in the expression for ‘at home’ as a general locative (ex. (226)):

- (219) *kitāb mez par hai*  
 book.DIR.SG.F table.OBL on be.3SG  
 ‘The book is **on the table**.’
- (220) *khilaune zamīn par paṛe hai*  
 toy.DIR.PL.M floor.OBL.SG.F on lie.DIR.PL.M be.3PL  
 ‘Toys are lying **on the floor**.’
- (221) *darvāze par ek baṛā sign hai*  
 door.OBL.SG.M on one big.DIR.SG.M sign.DIR.SG.M be.3SG  
 ‘There is a big sign **on the door**.’
- (222) *vah is bāt par prasann ho gayā*  
 DEM.DIR.SG DEM.OBL.SG matter.OBL on happy be.INF go.PFV.SG.M  
 ‘He became happy **at this**.’ (Singh 1977: 17)
- (223) *ghar ek kilometer par hai*  
 house.DIR.SG one km on be.3SG  
 ‘The house is **at a kilometer’s distance**.’ (Singh 1977: 17)
- (224) *vah thīk samay par āyī*  
 DEM.DIR.SG right time on come.PFV.SG.F  
 ‘She came **right on time**.’
- (225) *us par bahut qarz hai*  
 DEM.OBL.SG on much loan be.3SG  
 ‘S/he has too great a loan.’ (adapted from Khan 2009: 29)

- (226) *vah*                *ghar*                *par*    *hai*  
 DEM.DIR.SG    house.OBL.SG    on    be.3SG  
 ‘S/he is at home.’

From what has been outlined for UPARI, it may appear as if the semantic shift to a spatial relation was smaller than with MADHYE, as it basically only acquired a ground through reanalysis (which had already happened in the course of the post-Vedic genitive shift: see Chapter 5), but otherwise did not change much semantically, at least in its spatial usages. Does this mean that it became a postposition at a much earlier stage? As pointed out in section 7.1.2, UPARI in Apabhramsha, when referring to spatial grounds, still designates a space vertically above something else. Also, it may at this stage still occur on its own even when clearly relating to a ground. At the same time, we see such derived usages as in examples (196) or (197). Overall, UPARI is clearly in a transitory phase in Apabhramsha. The essential shift to actually denoting a simple spatial relation is when the feature of a vertical axis in the case of concrete spatial grounds is lost, with only the notion of contact remaining, foreshadowed in Apabhramsha in example (196), and completed in Hindi, for instance examples (220) and (221). It can then freely expand further to abstract grounds, such as stimuli, distances, points in time,<sup>24</sup> possessors, and it also figures in the phrase *ghar par* ‘at home’, not ‘on the house’. Since this crucial shift does not take place before the transition to New Indo-Aryan, I conclude that UPARI is not or not much ahead of MADHYE.

Did UPARI, too, come to denote a simple spatial relation only in groups where it follows the genitive? Evidence can be drawn from Apabhramsha, where almost all usages involving non-basic grounds are found in the form of the group GEN N. Of eighteen examples in my corpus,<sup>25</sup> only (227) shows the reversed order. The pattern GEN N can be seen for instance in examples (196), (197), or (201).

- (227) *karakaṇḍo*                      *uppari*    *kheyarāsu* /                *aipauru*  
 Karakaṇḍa.NOM.SG.M    UPARI    Khecara.GEN.SG.M    most\_excellent.ADV  
*pavaḍḍhiu*                      *ṇehu*                      *tāsu*  
 increase.PPP.NOM.SG    affection.NOM.SG    DEM.GEN.SG.M  
 ‘For Karakanda, the affection of the Khecara increased very much.’  
 (KA 2.9.1)

On the other hand, out of seventeen literal spatial usages, twelve show GEN UPARI, four show UPARI GEN, and one Type 4 discontinuity. Thus, though a small data set, these figures also support the hypothesis that a shift from ABOVE to OVER/ON took place

<sup>24</sup> This may be categorized as a basic ground, but does not figure in the corpus at all before the transition to New Indo-Aryan. Possibly, the transfer of the image schema from something spatially above to temporally above is not as straightforward as, for instance, from spatial to temporal containment.

<sup>25</sup> BH 222.3, 228.8, KA 2.9.1, 3.8.14, 3.12.10, 3.13.8, 4.1.13, 4.3.8, 6.7.6, 7.6.7, 7.11.13, 9.4.6, 10.23.2, PA 1.3.3, SA 489.5, 522.9, 598.1, 693.2.

in the group GEN UPARI. While this shift is not yet completed in Apabhramsha, the metaphorical extensions seen in the eighteen cases of non-basic ground similarly pave the way for metonymic shifts as in the case of MADHYE.

### 7.5 Grammaticalization and the emergence of phrase structure

How does phrase structure develop in a strongly non-configurational language? I have outlined in this chapter how semantic changes entail the establishment of the fixed syntagms in question. First, metaphorical shifts were shown to lead to an increase in overt possessors. Note that this is a phase prior to grammaticalization *sensu strictu*. Bybee et al. (1994: 5) observe that, typically, elements that develop into TAM markers already have a very general meaning from the start, for example 'go' and 'come' (if the element in question has a more specific meaning, then it still denotes a meaning central to human experience, as argued by Bybee et al.). In my corpus, we can observe how MADHYE/UPARI grow into semantically ever more general elements, though still lexical, through the expansion of their semantic host classes by way of metaphorical transfers. This increase is accompanied by an increase in overt possessors, necessary in the case of non-basic grounds, and as a result also of pronominal possessors. The increase of the latter concomitantly leads to an overall increase of the pattern GEN N. By the stage of late Middle Indo-Aryan, we have reached a point where, as a result of these processes, possessors are expressed overtly most of the time, and typically directly precede MADHYE/UPARI. Due to the already much generalized meanings including cases where a middle, for instance, is only construed but not intrinsically present, metonymic shifts become possible, changing MADHYE/UPARI into elements denoting simple spatial relations. At this point, we are indeed dealing with grammaticalization *sensu strictu*. Importantly, these metonymic shifts appear only to take place in groups GEN MADHYE/UPARI, that is, only in such cases where the possessor is overt and only in cases where it directly precedes. In this way, grammaticalization processes applying entire, specific constructions lead to the emergence of phrasal structures.

How do the developments observed for Indo-Aryan relate to grammaticalization processes cross-linguistically and to the emergence of phrases? In the case of strongly non-configurational languages where phrases—whether in the nominal domain or in another syntactic domain—develop from asymmetric groups, we can expect similar developments. In general, the historical changes must involve the establishment of a certain group that can undergo grammaticalization. I hypothesize that it is always semantic changes that entail obligatorification and fixation of order. Non-configurationality was acceptable to speakers at one point and there is no inner-syntactic reason why this would change. Instead, external reasons such as semantic shifts can render the expression of elements in certain usages necessary. Pro-forms too can be assumed to play an important role in general, since they are central in the

development of obligatorification. In the case of derived usages (as in the usage of non-basic grounds with *MADHYE/UPARI*), pro-forms become necessary in order to ensure interpretability. Furthermore, in a free word order language, pro-forms may have specific word order properties that can bear on the syntagmatics of constructions developing into phrases, as I argue in the present case study.

When, however, phrases arise from symmetric groups in a non-configurational language as in other branches of Indo-European, the scenario is a different one. Here, it is the reanalysis of surface strings, consisting of syntactically independent elements co-occurring as a syntagmatic group, which are reanalysed as syntactic units. Indeed, the Latin evidence, for instance, suggests that the shift from local particle to preposition was primarily a syntactic one. On the semantic side, the local particles acquired an argument slot for a ground. However, they did not otherwise undergo marked metaphorical or metonymic changes. The changes that did ensue on the semantic level, for example *ad* 'to, at' turning into a marker of indirect objects, developed when they had already become prepositions (e.g. Vincent 1999).

Last but not least, what about the much more widespread case, where we have a phrase from the start, that is, where the language in question is not strongly non-configurational?<sup>26</sup> Semantically, the same changes are involved in the grammaticalization processes. However, since the constructions in question are already phrasal, semantic changes do not entail the formation of the latter, at least not where it concerns the syntagmatic template. While syntagmatic order is already in place, semantic changes may however lead to the obligatorification of a dependent in cases where no relation of government or other obligatory dependency relation existed beforehand, as with a body part noun that did not originally require an overt possessor. In all cases, no matter what the syntactic origins, non-configurational or configurational, as soon as the shift from autosemantic to synsemantic has occurred, the well-known sub-processes characterizing advancing grammaticalization may set in—for example further functional extensions on the expression side, phonetic erosion and univerbation on the formal side, etc.

Now, synchronically, while not each and every specific phrasal construction in a language must have function words, I claim following Himmelmann (1997) that all phrase structure comes about through the grammaticalization of function words. For example, an article is not obligatory in each and every English noun phrase as in for instance plural indefinites: *Red roses are beautiful*. However, whether an article is present or not, we find one and the same phrasal template with slots for the head

<sup>26</sup> It is unclear how to place a language that already has phrasal organization in one domain, but not in another, and where phrases arise in the latter. It is plausible that, for instance in Romance, the prior existence of adpositional phrases influenced the formation of 'DPs', i.e. nominal phrases headed by articles, since the processes applied to the same syntactic domain of nominal expressions. However, it is not so clear whether the existence of phrasal nominal organization has a role to play in the formation of a phrasal verbal expression, only because phrasality already existed in the language.

noun and for the different types of adnominal elements. Such a phrasal template is the result of the grammaticalization of function words such as articles or adpositions in the nominal domain, or auxiliaries in the verbal domain. I will show that this claim can be supported by the recent history of the Hindi postpositions in the next chapter. Also such nominal expressions in Hindi as lack a postposition are organized on the basis of the postpositional phrasal template, the latter being the result of MADHYE/UPARI and other elements grammaticalizing. Thus, while it is individual constructions with specific lexemes that grammaticalize, they generate structural templates that apply to the syntactic organization of the respective syntactic domain more generally, above and beyond the specific constructions that the grams themselves occur in.<sup>27</sup>

## 7.6 Conclusion

In this chapter, I have dealt with the eventual transition from group to phrase, brought about as I argue by semantic shifts of GEN (+) MADHYE/UPARI, first metaphorical, then metonymic. The former are typical for a period leading up to or early in grammaticalization, while the latter mark the core of (primary) grammaticalization, that is, the transition from autosemantic to synsemantic. First, extensions to non-basic grounds make overt possessors a requirement. The increasingly frequent occurrence of such usages entails an increase in the frequency of overt possessors, required for semantic interpretability. This furthermore leads to an increase specifically in pronominal possessors, 'filling in' in cases where a non-basic ground is activated, but requires an overt anaphorical resumption. These developments lead to a situation in late Middle Indo-Aryan where we already almost always find the group GEN MADHYE/UPARI, that is, with an overt possessor and in this specific order. The stage is now set for grammaticalization. Through the many derived usages that we already find at this point, for example references to non-basic grounds that lack an intrinsic centre, the way is paved for metonymic shifts changing MADHYE and UPARI from referring to complex spatial regions to encoding simple spatial and non-spatial relations. Evidence from Apabhramsha and Old Awadhi supports the hypothesis that these metonymic shifts did indeed only occur in the group GEN MADHYE/UPARI, and did not apply to any other string. These semantic shifts finally result in the reanalysis of these groups as phrases.

I proposed that the development will be similar in other cases where phrases develop in a strongly non-configurational language from asymmetric groups, since here, too, fixed syntagms must develop involving the overt expression of dependents as well as fixed order. I hypothesize that it is always semantic shifts which entail the

<sup>27</sup> Niches of older syntactic patterns and strategies may remain (Himmelman 1998), as shown for Hindi in section 8.3.



ever more frequent formation of specific groups and their reanalysis as phrases. The formations of phrases from symmetric groups in non-configurational languages, as well as from constructions that are already phrasal in other languages, proceed along different lines. However, after the shift from autosemantic to synsemantic, and thus from group to phrase, is completed, the situation no longer differs in principle, being subject to the same sub-processes of grammaticalization, whether or not having started off in an already phrasal construction.



## Paradigmatization

### *A process sui generis?*

In outlining the rise of the Hindi postpositional phrases, a final step remains to be addressed, which is the formation of the paradigm<sup>1</sup> of seven simple postpositions.<sup>2</sup> In Chapter 5, I discussed adaptations of elements that would later become postpositions, namely the reanalysis of UPARI (and other adverbs) in analogy to relational nouns like MADHYE in the context of the post-Vedic genitive shift, through which they acquired a slot for an adnominal genitive. In this chapter, I study how these and the remaining simple postpositions come to be organized in a single new class of function words.

On a theoretical level, I focus on the question of whether paradigmization exists as a process *sui generis*. This focus—apart from its general interest for a theory of syntactic change—arises from the specifics of the data at hand. In Chapter 3, I showed that the postpositions originate in elements of at least three different syntactic classes. This raises the question of how they could end up as members of a single category. Some type of convergence seems the obvious answer. However, in the grammaticalization literature, the path of development of a grammaticalizing element is typically derived from the semantic and formal properties of its source construction, rather than by interactions with, or even adaptations to, other grammaticalizing elements (see section 2.1). In the following, I outline and test two different scenarios of paradigmization where this process is either denied or accorded the status of an independent

<sup>1</sup> Here, the notion ‘paradigm’ is not restricted to morphologically bound elements, as in the traditional understanding of the inflectional paradigm. Instead, a paradigm is defined by distributional similarity and paradigmatic complementarity: see section 8.1. Recent explorations into language change affecting morphological paradigms can be found in Robbeets and Bisang (2014).

<sup>2</sup> By implication, this also extends to complex postpositions (see Ch. 2, fn. 37). These are formed on the model of OBL *kā/-e/-ī* with a nominal form that is itself morphologically oblique, e.g. OBL *ke pās* ‘in the vicinity of/near sth.’ where *pās* descends from Old Indo-Aryan *parśve* (side.LOC) ‘by the side’. These constructions involve a younger layer of relational noun expressions, structurally parallel to older elements such as MADHYE, etc. Functionally, at least one complex postposition, namely *ke dvārā* ‘through’, is involved in grammatical case marking, namely of passive agents. Because I focus here on the initial emergence of the paradigm and because these constructions are structurally covered by the discussion of *kā/-e/-ī*, I do not discuss them separately. The other type of more elaborate postpositional constructions found in Hindi, ‘compound postpositions’, combinations of two simple postpositions, are touched on in section 8.1.

type of syntactic change. It will become clear that neither scenario can fully account for the emergence of the postpositional paradigm, but that we find aspects of both epiphenomenality (i.e. where paradigmaticization is no independent process, but only derivative of individual grammaticalizations) and interaction (i.e. where paradigmaticization is an independent process affecting grammaticalizing constructions).<sup>3</sup>

This chapter is structured as follows: I start off from a synchronic angle, discussing what constitutes paradigmaticity and subsequently outlining the paradigmaticity of the various simple postpositions in Hindi (section 8.1). I then turn to a discussion of two very different scenarios of paradigmaticization (section 8.2). In the first scenario (section 8.2.1), which I refer to as the *epiphenomenality scenario*, and which can be found in Bybee and Dahl (1989), paradigmaticization is denied the status of a process sui generis. Instead, it is viewed as an epiphenomenal product of the grammaticalization of constructions of a similar type in a similar historical period. In the second scenario (section 8.2.2), the *slot-formation scenario* as I call it, paradigmaticization is assumed to be an independent process. A forerunner element grammaticalizes first and establishes a syntactic slot in a specific construction. Other elements follow suit and wander into this slot, one by one. A scenario along these lines is found in van de Velde (2010). In both sections 8.2.1 and 8.2.2, I apply the respective scenario to the Hindi postpositions. It will turn out that neither of the scenarios fully matches the developments in Indo-Aryan. In 8.3, I therefore take a closer look at the recent history of the Hindi postpositions and show how there are both epiphenomenal and interactive facets to the emergence of the postpositional paradigm, both regarding its distributional environment and regarding relations of complementarity between the individual elements. I propose a unified, frequency-based account for both these aspects of paradigmaticization in section 8.4 and present a conclusion in section 8.5.

## 8.1 Paradigmaticity and the Hindi postpositions

Before I turn in the next sections to paradigmaticization, the process by which an element acquires (greater) paradigmaticity (Lehmann 2002 [1982]: 118–23),<sup>4</sup> I first adopt a synchronic perspective, exploring the notion of paradigmaticity and applying

<sup>3</sup> In this chapter, I prefer the more neutral term ‘interaction’ to ‘convergence’; see section 8.2.2.

<sup>4</sup> In Lehmann’s terminology, the notion of paradigmaticity can refer both to the integration of a paradigm as a whole and to the integration of one element into a paradigm: ‘What is meant here by paradigmatic cohesion or paradigmaticity is the formal and semantic integration both of a paradigm as a whole and of a single subcategory into the paradigm of its generic category. This requires that the members of the paradigm be linked to each other by clear-cut paradigmatic relations, especially opposition and complementarity’ (Lehmann 2002 [1982]: 118). I adopt the second understanding here, since I concentrate on the individual elements growing into members of the paradigm we find today. For this process I accordingly choose the term ‘paradigmatization’ which can refer either to an element or to a paradigm, while ‘paradigm formation’, which is also found in the literature, has only the holistic sense.

it to the Hindi postpositions. With this preliminarily ‘final’ state in mind, we can then turn to the development in recent centuries leading up to it.

The notion of paradigmaticity has many facets, but I consider defining properties to be distributional similarity, which relates to the syntagmatic axis, and complementarity, which refers to the paradigmatic axis. These properties are formal in nature with certain typical semantic correlates.<sup>5</sup> Note that these two axes are partly interdependent. Maximal distributional similarity entails complementarity, and the choice of a specific gram over another in turn hinges on the larger distributional context, such as the choice of verb and the respective semantic role. I tease apart the two domains for analytical and expository purposes.

Elements that form a paradigm are maximally similar regarding their syntactic distribution, occurring in the same slot in a certain construction. Typically, elements sharing a slot express values within a single domain of grammatical meaning, or within closely related and intertwined domains, such as modality, or modality and tense. On the paradigmatic level, the elements are complementary in the sense that they compete for the same slot and cannot concatenate. As a correlate, the elements in question often express semantic values that are contrastive, e.g. *+/- definite* in the realm of articles, or *+/- progressive* in the realm of auxiliaries, etc. Note that the ideal of maximal distributional similarity and complementarity is perhaps only reached in the case of inflectional paradigms. Often, function words do not show absolute paradigmaticity on all these levels, but they do so to a high degree, which sets them apart from content words. Regarding the Hindi postpositions, we will notice imperfect paradigmaticity with respect to both distribution and complementarity.<sup>6</sup>

In section 2.4.2, I outlined certain distributional properties of the Hindi postpositions. Formally, they share the property of occupying the right-most slot in the syntactic template of nominal expressions.<sup>7</sup> This template is further characterized by specified slots for a (typically referentially used) nominal(ized) dependent to its left and further adnominal categories at the left edge. Another distributional characteristic is that the postpositions trigger oblique marking on their nominal dependent (but only in part on adnominal elements, depending on inflectional class). Functionally,

<sup>5</sup> Elements and constructions that contrast functionally do not necessarily constitute a formal paradigm, as illustrated by strategies for tense and modality marking in English, consisting of such different devices as *be going to*, *will*, or the past tense suffix. Vice versa, elements forming a paradigm do not necessarily express the meaning of a single grammatical domain either, e.g. Hindi *ko* can express relations of case as well as of definiteness.

<sup>6</sup> Note that the idea of the ‘paradigm’ implies a stable constructional context, a notion at the heart of a structuralist view of grammar. Accordingly, I will refer in this chapter to the paradigmaticity of the postpositions, rather than of the entire postpositional constructions. However, in section 8.3.1.3, it will become clear that the distributional environment even within the immediate constructional context is not as stable as it might seem at first glance.

<sup>7</sup> I consider the postpositional slot as part of the phrasal template of nominal expressions, rather than external to it, as in an outer adpositional shell. My reasons for doing so are spelled out in section 8.3.2.2.

they share the property of marking structural and local cases.<sup>8</sup> In accordance with this shared behaviour, they are commonly regarded as a single category of function words. For example, Butt and King (2004) assign what I refer to as postpositions to a category K(ase) which heads Kase Phrases; syntactic generalizations are formulated with respect to this category (see also Bubeník's 2006a syntactic take on the postpositions). This is matched by their treatment in historical-morphological perspective where they are considered the middle layer of case marking, framed by an older layer consisting of inflectional remnants and a younger layer consisting of complex postpositions (e.g. Masica 1991: 230ff).<sup>9</sup>

While the elements seem to have a relatively homogeneous constructional environment, we find a certain amount of variation regarding complementarity. Only the more 'grammatical' of the postpositions stand in strict complementarity to each other where stacking (i.e. the combination of two postpositions) is impossible. Their selection is triggered by grammatical rules of argument realization (see Mohanan 1994) as well as differential case marking (e.g. de Hoop and Narasimhan 2005). For example, if a clause contains a perfective form of a transitive verb, the postposition *ne* marks the more agent-like nominal expression, while *ko* marks the more patient-like nominal expression if human and/or definite. The situation is quite different, however, in the realm of local case relations, where elements may combine into 'compound postpositions' when semantically felicitous. Here, the selection of a postposition is motivated by the choice of a specific meaning, rather than being structurally triggered. For instance, *gārī mē* 'in the car' contrasts with *gārī par* 'on the car' and with *gārī se* 'by/from the car'. An example of a compound postposition is *gārī mē se* 'from inside the car', formal complementarity being relaxed.

Notice that the distinction between the more grammatical and the more local postpositions is not clear-cut. For instance, it is correct that *ne* typically marks the more agentive argument in a perfective transitive clause, irrespective of the specific semantic role; for instance, example (228) involves an agent, example (229) an experiencer, and example (230) an instrument. Sometimes, however, *ne* is chosen on concrete semantic grounds, for example when expressing volition with the semelfactive verb 'cough': see examples (231) and (232).

- (228) *maī ne āj ek svādiṣṭ ḍosā khāyā*  
 1SG ERG today one delicious Dosa eat.PRF.SG.M  
 'I ate a delicious Dosa today.'

<sup>8</sup> To consider the postpositions as marking case is not uncontroversial. See the discussion in section 8.3.2.2.

<sup>9</sup> Also on other levels do we find characteristics of paradigmaticity. On the phonological level, all simple postpositions display a CV pattern (with the exception of *tak*) and are united by a common prosodic structure. Butt and King (2004) classify the elements as clitics, not bearing a main stress of their own. On the morphological level, simple postpositions are monomorphemic, any former morphological structure having eroded.

- (229) *maī ne āj kaī sundar cīzē dekhī*  
 1SG ERG today several beautiful thing.DIR.PL.F see.PRF.PL.F  
 ‘I saw several beautiful things today.’
- (230) *is cābī ne darvāzā nahī kholā*  
 DEM.OBL.SG key ERG door NEG open.PRF.SG.M  
 ‘This key did not open the door.’
- (231) *Rām khāśā*  
 Rām.DIR cough.PFV.SG  
 ‘Rām coughed.’ (Butt and King 2004: 190)
- (232) *Rām ne khāśā*  
 Rām.OBL ERG cough.PFV.SG  
 ‘Rām coughed (purposefully).’ (Butt and King 2004: 190)

*Ko* and *se* have usages not only of argument marking, but also of adjunct marking. For instance, *ko* marks indirect objects, but also points in time, e.g. *śām ko* ‘in the evening’. *Se* marks addressees with verba dicendi (e.g. *X se bolnā* ‘say to x’), the passive ACTOR (with *ke dvārā* as an alternative), as well as sources (*ghar se laṭnā* ‘return from the house’) and instruments (*cābī se darvāzā kholnā* ‘to open the door with a key’). As expected, it may not be concatenated with other postpositions when marking passive ACTORS (ex. (233)), but may be concatenated like the other local postpositions when designating a source (ex. (234)).

- (233) *Anil se kitābē becī jāeṅgī*  
 Anil by book.DIR.PL.F sell.PRF.F become.FUT.PL.F  
 ‘The books will be sold by Anil.’ (adapted from Mohanan 1994: 107)
- (234) *larkā ghar mē se niklā*  
 boy.DIR.SG.M house.OBL in from emerge.PRF.SG.M  
 ‘The boy emerged from inside the house.’ (adapted from Sandahl 2000: 33)

In sum, the two defining aspects of paradigmaticity, distributional similarity and complementarity, apply overall to the Hindi postpositions. However, there is a certain variation especially with regard to complementarity and a general characterization of a certain postposition cannot be given—instead, the degree of complementarity depends on its respective usage. In section 8.3.1.3, we will see that distributional similarity is not quite as uniform as may appear at first either.

## 8.2 Scenarios of paradigmaticization

Much has been published on what is essential to grammaticalization processes and what is not. For instance, many have observed that phonetic erosion is not a

necessary component (think of the uncontracted version of *be going to*), or that structural scope may either increase or decrease, so that no general rule can be given. Paradigmatization (apart from cases where there is only one grammaticalizing element), the integration of a grammaticalizing element into a paradigm with other grams, is however a commonly acknowledged correlate. In Lehmann (2002 [1982]), paradigmization is considered one of the parameters of grammaticalization processes. Diewald and Smirnova (2012) argue that paradigmization (or ‘paradigm integration’ in their terminology)<sup>10</sup> sets off grammaticalization from other historical changes such as lexicalization. After an initial preparatory stage, a second stage where grammaticalization is triggered, and a third stage of consolidation where the gram separates from its lexical source, the authors posit a fourth stage of paradigm integration:

We introduce a fourth stage which is specific to grammaticalization as opposed to lexicalization (and other kinds of language change as well) and is designed to capture this essential feature of grammatical(ized) items. We call it the stage of *paradigmatic integration* of the newly grammaticalized item into a (relatively) closed paradigm. (Diewald and Smirnova 2012: 127)

While paradigmization is generally acknowledged as a correlate of grammaticalization, or even singled out as the defining component, outlines of how exactly it comes about are scarce. However, some proposals can be found and I discuss two quite different ones in the next two sections.

### 8.2.1 *The epiphenomenality scenario*

Consider the following quotation from Bybee and Dahl (1989):

[I]f each gram follows a path of development according to its original meaning, then it develops independently of other grams. It belongs to a structural class if other grams from structurally similar sources ... undergo grammaticization at approximately the same period of time. Its membership in a structural class, then, is not determined solely by its meaning, but at least in part by chronological coincidence. (Bybee and Dahl 1989: 61)

Here, Bybee and Dahl express the view that the existence of structural classes—that is paradigms—goes back to grammaticalizations of categorically similar grams in the same period. The individual grammaticalizations are independent. This view forms part of the general theoretical approach of Bybee and colleagues in several works: that

<sup>10</sup> That Diewald and Smirnova have in mind a process similar to that of Lehmann can be gathered from the following definition, which invokes complementarity and a shared grammatical domain: ‘The last stage [of paradigm integration, UR] refers to a process whereby the new sign loses its independent status [i.e. its autonomy] as it comes to be associated with other members of the paradigm as well as with the paradigm/grammatical category as a whole. The newly grammaticalized sign comes to be confronted with opposing members of the same paradigm, on the one hand, and is gradually associated with a more abstract grammatical meaning which serves as a common denominator for the whole paradigm, on the other hand’ (2012: 128–9).



there is no such thing as an actual structuralist system of elements standing in clear-cut oppositions, where a change in one element will have an influence on the surrounding ones.<sup>11</sup>

We do not take the structuralist position that each language represents a tidy system in which units are defined by the oppositions they enter into. . . . In our view, then, language-internal systems, whether tidy or not, are epiphenomenal, and the clues to understanding the logic of grammar are to be found in the rich particulars of form and meaning and the dynamics of their coevolution. (Bybee et al. 1994: 1–2)

One of the examples given by Bybee and Dahl (1989) for the epiphenomenal nature of the genesis of structural classes are the English modals *may*, *can*, *must*, *will*, *shall*, etc. These compete for the same slot before non-finite verbs which lack the infinitive marker *to*. This is explained by the fact that they all grammaticalized at a time when the infinitival marker was the suffix *-an* and not yet *to*.<sup>12</sup>

How does this scenario apply to the Hindi postpositions? Given their many structural similarities, one would infer, based on Bybee and Dahl (1989), that they go back to structurally similar elements that grammaticalized in the same period. However, this does not fit the historical evidence. As we know, the Hindi postpositions neither descend ‘from structurally similar sources’ nor ‘undergo grammaticization at approximately the same period of time’. Instead, they derive from at least three different syntactic classes—nouns, an adverb, and a participle—and grammaticalize at different points in time. While *MADHYE* and *UPARI* were already used as periphrastic expressions of local cases at least from early Middle Indo-Aryan, the other postpositions are attested in such functions only from late Middle or early New Indo-Aryan onwards. Even if Bybee and Dahl’s notion of a grammaticalization in the same period is applied in a strict sense, that is as the transition from autosemantic to synsemantic—which is the transitional period between Middle and New Indo-Aryan for most elements including *MADHYE/UPARI*—at least *ne* is still lagging behind (see section 3.5). Since there is overlap neither in the type of syntactic class they descend from nor in the period of time when they grammaticalized, I conclude that there must be other reasons that explain the formation of the paradigm we find today.

<sup>11</sup> While generally advocating this approach, Bybee and colleagues do consider the possibility of interactions between grams in specific cases. For example, Bybee et al. (1994) report on a case where a gram in the semantic domain of anteriority expresses discontinuous meanings, namely an action just completed or an action completed on a previous day, while another gram expresses the intermediate meaning, namely an action completed earlier the same day. This may be explained by the fact that the second gram developed at a later point in time and came to express the intermediate meaning, which the old gram then ceased to express (see Bybee et al. 1994: 104).

<sup>12</sup> The fact that members of one category do not have to belong to a single functional domain is another piece of evidence for the epiphenomenality of paradigmization, as argued by Bybee and Dahl (1989: 60–1; see fn. 5 above).

### 8.2.2 *The slot-formation scenario*

Van de Velde (2010) presents a detailed scenario of how the determiner slot of the Dutch NP was formed. The terms 'paradigmatization' or 'paradigm formation' are not used, but van de Velde refers to the emergence of a syntactic slot and the subsequent recruitment of slot members ('lexical diffusion'). The question, however, is the same: how does a new paradigm of function words emerge?<sup>13</sup>

Van de Velde argues that the determiner slot was formed with the grammaticalization of a demonstrative into the definite article in Old Dutch. Since then, more and more elements have grown into this slot established by the definite article and the process is ongoing today. Notice that this is a relatively strong variety of 'interaction' between grammaticalizing elements. One might alternatively conceive of a mutual convergence, where the elements adapt to each other gradually. I adopt van de Velde's slot-formation scenario as one way in which paradigmaticization may be conceived of as a process *sui generis* for two reasons. First, the Hindi postpositions stem from different historical periods; therefore, the historical time spans of their development into grams overlaps only in part and they could thus have only partly mutually converged. Second, grammaticalization applies to syntagmatically fixed constructions (see section 2.6), and so does not leave much room for change and adaptation on a syntactic level. That this assumption is valid is supported by evidence discussed in section 7.4.1 involving data from early Hindi.

Returning to van de Velde's case study, he distinguishes between the syntactic slot and its slot-fillers, the determiner slot being filled by 'determinatives'. On a theoretical level, he assumes the reality of the determiner slot above and beyond its individual members:

[I]n order to fully grasp what is going on in Dutch (and English) NPs, a wider scope is needed: instead of taking the perspective of the individual lexemes, **the whole process must be looked at from the perspective of the abstract determiner slot itself**. All separate instances of grammaticalization of the individual lexemes are in fact part of a larger process, viz. the emergence of the determiner slot in Old Dutch and the subsequent growth of its membership figure. (van de Velde 2010: 291, emphasis added)

Van de Velde furthermore identifies a causal relation between the existence of the slot and elements migrating into it:

<sup>13</sup> In another article from 2009, van de Velde studies the emergence of the modifier slot in the Dutch NP. I will not discuss this study here, however, since it concerns the formation of a slot for lexical items. I would expect other factors to be at work than in the paradigmaticization of function words. However, I do believe that the developments of the formation of slots for grammatical and for lexical elements can (but need not be) connected, as I show below in section 8.3.1.2 for the emergence of separate slots for referential and modifying elements within Hindi nominal expressions, another indirect result of the grammaticalization of the postpositions.

The upshot [of this study] is that the emergence and accretion of the determiner projection in the NP is the underlying cause, rather than the epiphenomenal result of the array of lexically specific instances of grammaticalization that feature separately in scholarly studies ... (van de Velde 2010: 293)

Van de Velde presents a catalogue of semantic as well as formal criteria which define elements as determinatives. Criterion I is semantic in nature, referring to a part of speech encoding exclusively 'definiteness, referentiality, specificity or identifiability'. Criterion II invokes the complementary distribution of adnominal elements expressing these semantic features and Criterion III the obligatoriness of having overt elements encoding them. Criterion IV refers to the ability to function predicatively, evidence against determinative status. Criterion V involves a reference to the modifier slot, where an element occurring to the right of a modifier cannot be a determinative. In order to trace how elements grow into the determiner slot, van de Velde analyses a historical corpus of Dutch, studying the frequencies with which elements fulfil these criteria over time. For instance, while older elements such as the possessive pronouns have fully entered the determiner slot (i.e. fulfil the criteria in all or at least most usages), younger elements like the demonstrative manner adjective *dergelijk* 'suchlike' are at present under way, fulfilling the criteria in part, while further elements are only taking initial steps towards the slot.

To test van de Velde's scenario for the Hindi postpositions, we must ask whether there is one element taking the lead as in the case of the Dutch definite article, establishing the postpositional slot defined by certain semantic and formal criteria, which younger elements wander into in a piecemeal fashion. In analogy to the Dutch definite article, an element most integrated into the postpositional slot in Hindi would specify grammatical case relations (criterion I), stand in complementary distribution on a formal level (criterion II), and at least in some cases be obligatory in the sense of answering to rules of argument realization (criterion III).<sup>14</sup> (Van de Velde's remaining two criteria are relevant in stages up to late Middle Indo-Aryan, but not thereafter).<sup>15</sup> All these criteria refer to aspects of what I have defined as

<sup>14</sup> I adopt an understanding of 'obligatory' slightly modified from van de Velde's. He refers to the question whether a grammatical category must be instantiated by one of its members (in a certain environment): 'When a language uses an obligatory element for the expression of definiteness etc., these elements may be considered determinatives' (2010: 268). However, I concentrate here on the perspective of the gram, not the category, and ask whether the individual element fulfils the criteria of membership in the postpositional slot. Note that this is neither Lehmann's (2002 [1982]) parameter of 'transparadigmatic' nor of 'intraparadigmatic' obligatoriness. The former relates to the instantiation of a category as understood by van de Velde. Intraparadigmatic obligatoriness overlaps with my usage, but it also involves the question of whether other elements may be used alternatively in the same context.

<sup>15</sup> Criterion IV, the question of predicative function, has an equivalent in usages of *MADHYE*, *UPARI*, etc. in isolation. Criterion V can here relate to occurrences of the element elsewhere than to the right of nominal elements. Both topics were touched on in Chapters 5, 6, and 7 regarding *MADHYE* and *UPARI*. By the time of *Apabhramsha*, appearance in isolation has almost fallen out of use and there are only stray examples in Old

paradigmaticity, namely the expression of values within a shared grammatical domain, and formal as well as functional complementarity.

As in the case of the Dutch determiner slot, the various postpositions fulfil the criteria to a greater or to a lesser degree, as we saw in section 8.1 with regard to complementarity. There is, however, a fundamental difference between the Dutch determinatives and the Hindi postpositions. Those postpositions which appear on the scene first, *mē* and *par*, are the ones which least fulfil the criteria for membership in the postpositional slot. They express concrete spatial meanings, rather than being structurally selected, and do not stand in absolute complementary distribution, but may concatenate. Instead, the youngest element *ne* is the one which fulfils the criteria best, almost always being structurally selected and not allowing for concatenation. In Hindi, it is the synchronic system of argument realization and differential case marking which determines the paradigmaticity of the postpositions in their various usages, and not their age.

In sum, while the epiphenomenality scenario does not match the fact that the source constructions are more heterogeneous than expected, the slot-formation scenario would predict a different chronological order, where the elements grammaticalizing first are also the most paradigmatic. In addition, it will be shown below in section 8.3.1.3 that even the most paradigmatic postpositions *ne* and *ko* show more heterogeneity than would be expected if the slot had a causal effect on elements, inducing their grammaticalization and development into slot-fillers.

### 8.3 The emergence of a paradigm

In the previous sections, we have seen that neither the epiphenomenality scenario nor the slot-formation scenario match the Indo-Aryan data fully. In the following sections, I will therefore take a closer look at the recent history of the Hindi postpositions, first focusing on distribution, then on complementarity. Based on this descriptive foundation, I will subsequently turn to an interpretation of the data in section 8.4.

#### 8.3.1 *Distribution*

In the next two sections, I explore the development of the postpositional slot as defined by its immediate environment, the construction it is embedded in. While I focused in previous chapters on *MADHYE* and *UPARI*, I now expand the data set to include what is known about the histories of the other simple postpositions of Hindi. On the one hand, the constructional environment of the various emerging postpositions is very

---

Awadhi, where *MADHYE* precedes rather than follows its dependent. Data on the other postpositions are too scarce to allow for an application of these criteria.

homogeneous and gives rise to the shared structural pattern OBL POSTP (section 8.3.1.1). In fact, the grammaticalization of the postpositions leads to the fixation of further schematic slots inside nominal expressions, such as for modifiers (section 8.3.1.2). On the other hand, when taking a closer look at constructions containing pronominal forms, it becomes clear that the distributional similarity among the various postpositions is not in fact maximal (section 8.3.1.3).

8.3.1.1 *The selection of a construction* Of all the constructional environments that they occurred in across Old and Middle Indo-Aryan, MADHYE/UPARI grammaticalized in the construction GEN N and not in another one.<sup>16</sup> They could in principle have grammaticalized in apposition with local case forms, as did cognates of UPARI in other Indo-European branches, or in the reverse order,<sup>17</sup> or both, but they didn't (see Chapter 5 as well as sections 7.4–7.5). *Ko* is another likely candidate for a grammaticalization in a construction GEN N. As outlined in section 3.4, it is reconstructed as going back to *kakṣe* 'flank.LOC'. However, since it is attested in its modern adpositional function only from a time when the genitive and other local case forms had already collapsed into the general oblique, its constructional origins cannot be ascertained. Still, its likely descent from a relational noun makes an origin in a construction with genitives very probable, as this is by far the most frequent construction type in which relational nouns occurred at least from early Middle Indo-Aryan onwards (see section 5.2).

Additional research is needed to ascertain the origins of *kā/-e/-ī* regarding the exact participial form it descends from, *kṛta(ka-)* (perfective participle) or *kārya(ka-)* (future participle) (see section 3.6). Both forms could construe with genitives from early on. Whatever the ultimate source—an origin in *kṛta(ka-)* appears likely on semantic

<sup>16</sup> I remark in this section on all postpositions apart from *ne*, which was borrowed and therefore presents a special case, discussed below in section 8.3.2.1. Note that the example of *ne* is perhaps the best evidence for the reality of the phrasal structure above and beyond the number of individual grammaticalizations that gave rise to it, which are discussed here. It seems to have been borrowed into the postpositional slot in the early Hindi period, the slot being delineated by the postpositional syntagm, which had only gradually come into being in the course of the grammaticalization of the older members of today's postpositional paradigm.

Before I proceed, a note on the reanalysis of UPARI as outlined in Chapter 5 is also in order. Its reanalysis and that of other adverbs in the context of the post-Vedic genitive shift may appear like a prototypical interaction between constructions, having taken place by analogy to relational noun expressions like GEN + MADHYE. However, I do not consider this an interaction between constructions that would form part of a process of paradigmization as a correlate of grammaticalization. First, this reanalysis is a change long before grammaticalization sets in. Second, we are dealing with an abrupt change in word class, which altered the semantic and syntactic structure of the element in such a way that there is no identity between pre-reanalysis and post-reanalysis UPARI and where only the latter is the precursor of Hindi *par*. Accordingly, I only consider here reanalysed UPARI.

<sup>17</sup> Among the complex postpositions of Hindi, there are some that allow for variation of order, e.g. *ke binā* 'without', where *binā* may also precede the noun (see Strnad 2013: 360).

grounds—the modern construction is in place in Apabhramsha.<sup>18</sup> Here, the forerunner form *kera-* serves as a marker of possession,<sup>19</sup> agreeing with a head noun, the possessee, and combining with a genitive, the possessor. Compare the following example from Apabhramsha, cited by Bubeník, and the matching Hindi translation:

- (235) *kesari jasu keraem huṃkāraḍem*  
 lion.NOM.SG.M REL.GEN.SG.M *kera*.INS.SG roaring.INS.SG  
*muḥaḥuṃ paḍanti ṛṇāim*  
 mouth.GEN.PL.N fall.3SG leaf.NOM.PL.N  
 ‘the lion, by whose roar grass fell [sic!] from (your) mouths’  
 (Hemacandra 8.4.422.15, cp. Bubeník 1998: 74)
- (236) *vah śer, jis kī garaj se tumhāre*  
 DEM.DIR.SG.M lion REL.OBL.SG POSS roar.OBL.SG.F ABL POSS.2SG.OBL.PL  
*mūh se khānā gir gayā thā*  
 mouth.OBL.PL.M from food.DIR.SG.M fall go.PRF.SG.M be.PST.SG.M  
 ‘that lion, by whose roar food fell from your mouths’

Certainly, this construction differs in several respects from those containing denominal postpositions. The participial form serves as a linker between two nominal expressions, agreeing in case, gender, and number with the possessed noun, whereas other postpositions typically link a nominal expression to the verb, and there is no agreement process involved. Irrespective of these differences, the important aspect that *kā/-e/-ī* shares with the denominal forms is that it grammaticalized in a construction with a directly preceding genitive.

Turning to *se*, I showed in section 3.3 that its etymological origin is uncertain. Among the candidates put forward is a noun *saṅge* ‘in contact’, which could have taken a genitive, but there are also other suggestions involving a participle or an adverb, which might not have construed with genitives. In fact, the candidate usually favoured now is the adverb *samaṃ* ‘together’. As for the Apabhramsha evidence,

<sup>18</sup> Bubeník (1998: 74–6) suggests that adnominal usages of *kera-* might have originated in its being predicatively used and bearing focus (‘in the slot of the complement which receives sentential stress’). However, at least if it descends from *kṛta(ka)-*, which is the more likely scenario considering its semantic structure (see section 3.6), it seems more plausible that the forerunner of the modern construction was always an adnominal one. The resultative *ta*-forms, which *kṛta-* belongs to, were in origin adnominal modifiers. In the transition from early Old Indo-Aryan to younger stages of Old Indo-Aryan and early Middle Indo-Aryan they did develop certain functions as main verbs, in a construction with a nominative O-argument and an instrumental (or sometimes genitive) A-argument—the origin of the ergative construction. There may be a connection between the two constructions, that is the adnominal possessive and the verbal ergative, especially when the genitive is used for A-arguments in the verbal construction, and it has been discussed (controversially) whether the latter go back to the role of possessor (cp. Benveniste 1952; Cardona 1970). However, in the case of the construction we are dealing with here, it would seem unlikely that *kṛta(ka)-* started off in adnominal function, then developed a predicative usage, before it developed back into adnominal usage.

<sup>19</sup> Another form used in this function is *tana-* (Bubeník 1998: 74), cp. Skt. *tanu-* ‘body’.

Bubeník (1998: 83–4) points out that *samau* (from *samam*) typically occurs with instrumentals, which matches its instrumental usage in Hindi, and the same goes for *sahū* (which Bubeník deems to be at the origin of Hindi *se* and which continues Skt. *saha* ‘with’).<sup>20</sup> If *samau* (or *sahū*) should be the precursor form of Hindi *se*, it may have grammaticalized with a local case form rather than with a genitive and, if this scenario is correct, it would have aligned to the other grams only at a point of time when the genitive and the local case forms became fully syncretized into the general oblique.<sup>21</sup> The word order, in any case, is again the same.

Summarizing, all postpositions discussed here, possibly apart from *se*, are likely to have grammaticalized in a construction GEN N (N subsuming nominal and nominalized forms), and *se* joins this pattern at the latest after the genitive is replaced by the oblique. This is the case even though all elements also used to combine with other case forms and occurred in other word order constellations. In part, the overlap is indeed purely structural. The relation between *kā/-e/-ī* and the genitive, for instance, is entirely different from that between MADHYE/UPARI and the genitive. *Kā/-e/-ī* originates as an adnominal modifier, the genitive being subordinate to the possessed noun, and not to the precursor form of *kā/-e/-ī*. By contrast, MADHYE/UPARI directly head the genitive and do not modify any other nominal.

8.3.1.2 *The emergence of a hierarchical phrase* The reader may have noticed that I have been referring to the grammaticalization of postpositions as bringing about the phrasal organization of *nominal* expressions in Hindi. This may be surprising in view of the fact that noun phrases are typically seen as an inner shell of structure, possibly marked by determiners, around which there may be an outer shell, the adpositional phrases. However, in the case of Indo-Aryan, no articles have developed and, as I show in this section, it is indeed the case that it is the grammaticalization of postpositions which gives rise to the hierarchical phrasal organization of nominal expressions generally. To talk about the emergence of postpositional phrases in Indo-Aryan, then, is coextensive with talking about the emergence of the phrasal organization of nominal expressions. In this section, I study how further slots are formed in the constructional environment over time.

An important difference between the Old Indic state of affairs and Hindi is that the distinction between nouns and adjectives has strengthened through the fixation of a syntactic slot for referential elements (typically filled by nouns) and another slot for adnominal attributes (typically filled by adjectives), as pointed out in section 2.4.2.

<sup>20</sup> I believe that the full answer for the origin has not yet been found, since there is no clear evidence for the origin of the ablative usage, as well as certain other meanings of *se*.

<sup>21</sup> *Se* may thus be the exception to the rule that, in Indo-Aryan, the modern adpositions grammaticalized in asymmetrical rather than in symmetrical groups (see Chapter 5). However, the earliest elements to grammaticalize, and thus to establish the postpositional phrases in Indo-Aryan, do so in asymmetrical groups, with OBL *se* being a later addition.

This distinction also holds true on the morphosyntactic level: apart from masculine singular *ā*-stem adjectives, only elements in the referential slot change into the oblique form (if morphologically distinct), when governed by a postposition. Note that this distinction hinges on the syntactic slot rather than on the morphological potential. The morphological distinction between noun and adjective is still not very strong—they can both inflect for the same categories, but only the latter agree in gender. However, while adjectives may still be used in referential function without further derivative suffixes or other additional strategies as in Old Indo-Aryan (see Kellogg 1972 [1875]: 426; Kachru 2006: 65), nouns now require derivational morphology in order to be used as modifiers. I claim that the distinction between the two syntactic slots develops in the wake of the grammaticalization of the postpositions. Thus, not only does the postpositional slot itself become established, but so also do further slots inside nominal expressions.

In Apabhramsha, attributive adjectives and other adnominal elements may still precede or follow the noun they modify.<sup>22</sup> The following examples involve adjectives (exx. (237), (238)), a participle (ex. (239)), an indefinite modifying a referentially used adjective (ex. (240)), and a numeral (ex. (241)), all following the respective referentially used noun:

- (237) *niya-dhammiṇa* *patta-jasu* / *majjhi*  
 own-religion.INS.SG obtain.PPP.NOM.SG-fame.NOM.SG.M / MADHYE  
*vāla-tavasihī* *udaggahā*  
 naïve-Tāpasvin.GEN.PL haughty.GEN.PL  
 ‘According to his religion, he had obtained glory (lit. (he was one) with obtained glory) among the **haughty naïve Tāpasas** [a type of ascetic].’  
 (SA 689.5)
- (238) *jaha* *payadiu* *appu* *pañ* / *karivi*  
 as make\_evident.PPP.NOM.SG.M REFL.NOM.SG INS.2SG / make.GER  
*majjh’* *uvari* *karuṇa* *nimmala*  
 GEN.1SG UPARI pity.ACC.SG pure.ACC.SG  
 ‘As you have shown yourself, having **pure compassion** for me, (have pity on me!)’  
 (SA 489.5)
- (239) *ēha* ‘*kumāri*’ *bhaṇivi* *jā* *vutti*, ... *pañcahā*  
 one ‘princess’ call.GER REL.NOM.SG.F say.PPP.SG.F five.GEN.PL  
*sayahā* *jai* *vi* *maṇi* *ranjiya*, / *jai*  
 hundred.GEN.SG if PRT jewel.NOM.SG colour.PPP.NOM.SG then

<sup>22</sup> Cases of discontinuity exist also, e.g. in KA 9.4.6 or KA 6.16.1. However, since the Apabhramsha texts are metric, it is necessary to disentangle whether or not the elements placed apart have different syntactic roles and therefore may or may not constitute a discontinuous nominal expression. Therefore, here I only take into account continuous strings.



*vi kō vi guṇadōsu na najjai*  
 EMPH INDEF badness\_of\_character.ACC.SG NEG enter.3SG  
 ‘The one who is called ‘princess’ said...if she is the **coloured jewel** among five hundred,... then she will not enter great pride.’ (BH 193.8)

- (240) (= ex. (147))  
*ahaha ari jiya karisi mā rōsu /*  
 EXCL EXCL soul.VOC.SG do.IMP.2SG NEG anger.ACC.SG  
*iyarassu kassu vi uvari!*  
 other.GEN.SG.M INDEF UPARI!  
 ‘Oh soul, don’t be angry with **some other person!**’ (SA 693.2)<sup>23</sup>

- (241) *ettha antari paḥiṇa gayaṇassu...*  
 there between path.INS.SG heaven.GEN.SG  
*khayara-ku<ma>ra dō tattha āgaya*  
 Vidyādhara-prince.NOM.PL.M two there arrive.PPP.NOM.PL.M  
 ‘In the meantime, **two Vidyādhara-princes** [= supernatural beings, UR] came on a heavenly path...’ (SA 483.2)

This ‘reversed’ pattern is still attested in Old Awadhi. Saksena (1971: 326) lists cases of a noun followed rather than preceded by an adjective (*muni gyānī* ‘well-versed sages’), possessive pronoun (*pitā hamāra* ‘our father’), or numeral (*dīpa dui* ‘two lamps’). It was shown in Chapter 7 that the postpositional slot is by and large established at this point. However, in this period, there are also still cases where MADHYE precedes the oblique (see exx. (217), (218)). I want to propose a connection between both findings.

The postpositional slot does not emerge simply at the right-most edge of nominal expressions, but to the right of the respective referentially used element.<sup>24</sup> This is foreshadowed in Apabhramsha where, whenever we are dealing with a complex nominal expression that forms the dependent of MADHYE/UPARI and where MADHYE/UPARI appear at the right edge, we find the order *adnominal modifier-noun-postposition* (sometimes with an inserted emphatic clitic), rather than another pattern (ex. (240), potentially an exception, is a special case involving two pro-forms). Compare the following examples:

- (242) *niya-māhappu samaggaha vi / jayaha majjhi*  
 own-magnanimity.ACC.SG complete.GEN.SG PRT world.GEN.SG MADHYE  
*payaḍantu Saṇatukumā<ru kumā>ravaru / Hatthiṇāgapuri*  
 make\_evident.CONV excellent\_prince\_S. Hastināpura.ACC.SG

<sup>23</sup> Here, *vi* combines with *kassu* into an indefinite as a lexicalized combination.

<sup>24</sup> In exceptional cases, the dependent may be non-referential, as with certain pronouns or quantifiers. However, when we are dealing with combinations of a referentially used form and an adnominal modifier, this is the order we find.

*pattu*

reach.PPP.NOM.SG

‘Having made evident his magnanimity **in the middle of the entire world**,  
the excellent prince S. reached H.’ (SA 713.7)

- (243) *ahō kula-dhavalā! ēu darisāvahi! /*  
Oh family-beautiful.VOC.SG DEM.ACC.SG show.IMP.2SG /

*annaho jaṇaho majjhi mēlāvahi!*  
other.GEN.SG people.GEN.SG MADHYE join.IMP.2SG

‘Oh beauty of the family! Show this! Get together **in the middle of other people!**’ (BH 82.8)

This relative order becomes firmly established in the course of time. In Hindi, also nominal expressions that lack a postposition show the same syntagmatic template, as illustrated by (244) with a postposition and (245) without one. Demonstratives and quantifiers appear at the very left edge in accordance with their semantic scope.

- (244) *mere tīn choṭe baccō ko zukām hai*  
POSS.1SG.PL.M three small.OBL.M child.OBL.PL.M DAT cold.DIR.SG be.3SG  
‘My three small children have a cold.’ (lit. ‘A cold is to my three small children.’)

- (245) *mere tīn choṭe bacce miṭāi*  
POSS.1SG.PL.M three small.DIR.PL.M child.DIR.PL.M sweet.DIR.PL.F  
*khānā cāhte hai*  
eat.INF want.IPFV.PL.M be.3PL  
‘My three small children want to eat sweets.’

Thus, the grammaticalization of the postpositions results not only in the formation of the postpositional slot itself, but also in slots for the different referential and non-referential elements within nominal expressions. That Hindi did not develop articles, the other type of function word whose grammaticalization is attested as giving rise to phrasal nominal expressions (see Himmelmann 1997; van de Velde 2010), has no bearing on the structure of nominal expressions without postpositions—they, too, show the same hierarchical template.<sup>25</sup>

8.3.1.3 *Retention of earlier form* While I studied the emergence of the abstract hierarchical, phrasal schema of nominal expressions in the last two sections, I now

<sup>25</sup> Note, however, that grammaticalization does not always seem to give rise to a hierarchical phrase, since flat phrase structure with variable internal positions is also attested, for instance for Tagalog (cp. Himmelmann 1997: 160ff). This may have to do with the different types of function words involved and their respective semantic relations to the nominal elements. For instance, I touch on differences in the grammaticalization of articles and of adpositions in section 8.4.

concentrate on the opposite, namely on constructional sub-types where remnants of source constructions have survived into modern times, rather than adapting to the general abstract schema. Bybee et al. (1994) discuss how semantic properties of a source construction may survive a long time even when a grammaticalization process is already much advanced, the ‘retention of earlier meaning’ (1994: 15–17; see also section 2.1). For example, English *will* retains an old sense of volition in questions. In this section, I focus on the retention of earlier *form*. In some constructional subtypes involving pronouns, we find old pieces of form that are an exception to the constructional schema OBL POSTP: consider Table 8.1.

The Hindi pronouns, like nouns, have direct and oblique forms deriving from former morphological case forms. Even more so than nominal forms (compare the introductory lines of Chapter 5), the pronominal forms go back to a variety of different case forms (see Beames 1970 [1875]; Kellogg 1972 [1875]; Oberlies 1998). The direct forms can be reconstructed as deriving from nominatives (*tū*, *yah*, *vah*, *ye*, *ve*), oblique stems with NOM-ACC suffixes (*ham*, *tum*), a former instrumental (*maī*), and a noun (*āp* from *ātman* ‘self, soul’). The oblique stems, where they are morphologically distinct, descend from dative (*mujh*, *tujh*, later GEN-DAT) and genitive forms (*is*, *us*, *in*, *un*). The paradigm of *se* combining with pronominal forms in Table 8.1 displays the normal combination of a postposition with the oblique form, which is also the one found for the remaining local postpositions not listed here, *mē*, *par*, and *tak*. I will not discuss these varied sources further, but focus on deviating forms found with some of the other postpositions.

*Ko*, *ne*, and *kā/-e/-ī* do not simply combine with the pronominal oblique stems. *Ko* may optionally combine with the oblique forms, but there are also old inflectional variants for all forms apart from *āp*. These old variants are reconstructed by Oberlies (1998: 17–19) as deriving from accusative forms, while Kellogg (1972 [1875]: 211, 213, 217, 218) traces only the 1st and 2nd plural forms to accusatives, but the singular forms generally to different types of form extended by genitive suffixes. The combination of ergative *ne* with a pronominal referent shows the most variation. It is the exception among the postpositions (apart from *kā/-e/-ī* which shows again a different pattern) in taking the direct forms *maī* and *tū* rather than the oblique stems *mujh-* and *tujh-*. Furthermore, there are separate demonstrative plural forms *inhō ne* and *unhō ne*, traced by Oberlies (1998: 20) to extended genitive forms (see also Kellogg 1972 [1875]: 218). The remaining forms either do not distinguish between direct and oblique (*ham ne*, *tum ne*, *āp ne*), or the oblique is chosen (*is ne*, *us ne*). The inflectional possessive forms in the 1st and 2nd person are old genitive pronouns fused with participles of the root *kr* ‘do’, the same forms as underlie *kā/-e/-ī* (cp. Schwarzschild 1954).

So, on the one hand, the class of postpositions appears to be clearly demarcated by its stable constructional environment, occurring in the right-most position in nominal expressions, governing the oblique of a referentially used nominal element to its left. This is a result of their grammaticalization in similar environments as outlined in

TABLE 8.1. Hindi postpositions with pronominal dependents

	1SG	2SG (informal)	1PL	2SG/PL (semi-formal)	2SG/PL (formal)	DEM SG (prox.)	DEM SG (dist.)	DEM PL (prox.)	DEM PL (dist.)
DIR	<i>maĩ</i>	<i>tū</i>	<i>ham</i>	<i>tum</i>	<i>āp</i>	<i>yah</i>	<i>vah</i>	<i>ye</i>	<i>ve</i>
OBL	<i>mujh</i>	<i>tujh</i>	<i>ham</i>	<i>tum</i>	<i>āp</i>	<i>is</i>	<i>us</i>	<i>in</i>	<i>un</i>
<i>se</i>	<i>mujh se</i>	<i>tujh se</i>	<i>ham se</i>	<i>tum se</i>	<i>āp se</i>	<i>is se</i>	<i>us se</i>	<i>in se</i>	<i>un se</i>
<i>ko</i>	<i>mujh ko/mujhe</i>	<i>tujh ko/tujhe</i>	<i>ham ko/hamē</i>	<i>tum ko/tumhē</i>	<i>āp ko</i>	<i>is ko/se</i>	<i>us ko/use</i>	<i>in ko/inhē</i>	<i>un ko/unhē</i>
<i>ne</i>	<i>maĩ ne</i>	<i>tū ne</i>	<i>ham ne</i>	<i>tum ne</i>	<i>āp ne</i>	<i>is ne</i>	<i>us ne</i>	<i>inhōne</i>	<i>unhōne</i>
<i>kā/e/ī</i>	<i>merā/e/ī</i>	<i>terā/e/ī</i>	<i>hamārā/e/ī</i>	<i>tumhārā/e/ī</i>	<i>āp kā/e/ī</i>	<i>is kā/e/ī</i>	<i>us kā/e/ī</i>	<i>in kā/e/ī</i>	<i>un kā/e/ī</i>

the preceding section. This schematicity, however, does not extend to all subtypes of nominal dependent. In particular, deviating pronominal forms combining with the more grammatical ones of the postpositions limit their distributional similarity. Before I draw theoretical conclusions about how distributional similarity has arisen in recent centuries, including its limitations as sketched in this section, I first turn to complementarity.

### 8.3.2 Complementarity

I outlined above in section 8.1 how the postpositions show varying paradigmaticity with regard to complementarity, depending on usage. Here, I focus on the way in which elements entered the paradigm at all. Specifically, I focus on those elements that entered the paradigm as secondary additions, when the paradigm was already in place with the oldest of the postpositions as its members. I first study how ergative *ne* enters the system (section 8.3.2.1), and then turn to the integration of the direct as the unmarked form (section 8.3.2.2).<sup>26</sup>

**8.3.2.1 The integration of ergative *ne*** By early New Indo-Aryan, postpositions have become part of the grammatical system, namely precursor forms of *mē*, *par*, *ko*, *se*, and *kā/-e/-i*. It is only now that *ne* enters the picture. As outlined in section 3.5, the most plausible scenario is that it was borrowed some centuries ago (possibly between the fourteenth and sixteenth centuries) from a language spoken to the west of the modern-day Hindi belt, such as Rajasthani or Haryani. Importantly, *ne* enters the language at a time when the ergative-absolutive alignment with perfective verb forms is already long established, the ergative argument being encoded by the bare oblique. Thus, the postpositional phrase headed by *ne* did not arise in the same way as the other postpositions, deriving from a lexical element governing some nominal form. Instead, *ne* is borrowed into a pre-existing construction and into a pre-existing slot, reinforcing the oblique. This is an important piece of evidence that the postpositional slot as defined by its constructional environment acquired a reality independent of the elements whose grammaticalization gave rise to it.

Turning to the paradigmatic axis, *ne* enters a pre-existing set of elements that contrast formally and functionally (i.e. primarily the set of postpositions with argument marking functions, but since this is not categorically delineated, by extension also the complete set of simple postpositions). As already indicated in section 3.5, in the neighbouring languages which are candidates for being the donor of *ne*, it has a wider functional space than in Hindi, namely for marking dative, and in part ablative and locative functions, besides ergative function (see section 3.5). Its more restricted functional space in Hindi is attributed by Butt and Ahmed (2011) to the fact that Hindi already possessed *ko*

<sup>26</sup> Complex postpositions with structural case marking functions might also be discussed in this context, but they are beyond the scope of this study.

for dative functions and *se* for ablative functions when *ne* entered the language.<sup>27</sup> Accordingly, *ne* was only taken over for the function that as yet had no postpositional marking (Butt and Ahmed 2011: 568).<sup>28</sup> In Hindi, it becomes the most paradigmatic postposition, that almost exclusively answers to structural rules of argument realization (see the examples in the next section). In this scenario, complementarity is the factor that shapes the usages of *ne*.

8.3.2.2 *The paradigmaticity of zero* In the literature, there has been some debate about the correct label for the morphological remnants of inflectional case marking on the one hand, and the postpositions on the other hand. For instance, Butt and King (2004) consider what I call simple postpositions to be case markers, the old layer of inflections being hardly functional any more. Postpositions in their terminology are what I have been referring to as complex postpositions. Spencer (2005), in contrast, advocates the view that the inflectional layer is the domain of case marking. These different analyses result from the fact that, in a small area of the grammatical system, the morphological distinction between direct and (bare) oblique still marks a difference in case relation (or at least in a local case relation), while elsewhere—and this is the much larger domain—the direct instead functionally contrasts with postpositions (which in turn trigger oblique marking on the noun).<sup>29</sup> The small area in which the oblique marks a distinction is in encoding goals, for example *Kalkatte* ‘to Calcutta’ and it also figures in other adverbial expressions.<sup>30</sup>

This situation has not always been so, however. In early Hindi varieties, the morphological system was still a true inner shell of case marking, with an outer one of postpositions optionally used for local case meanings. In Old Awadhi, for example, the contrast in the domain of argument marking is still between direct and oblique (Saksena 1971: 120ff):<sup>31</sup> as in Hindi, the direct marks A- and (inanimate and/or indefinite) O-arguments in non-perfective clauses, and (inanimate and/or indefinite) O-arguments in perfective transitive clauses. In contrast to Hindi, the oblique also still expresses several relations of case including argument marking ones. It marks

<sup>27</sup> Hoernle (1880: 219–20) argues along these lines.

<sup>28</sup> More specifically, *ne* was only borrowed into the specific construction of the bare oblique encoding ergative arguments in clauses with perfective verb forms of transitive verbs. Outside of this specific construction, it does not generally contrast with *ko* or *se* (but see Bashir 1999).

<sup>29</sup> That the delineation between inflectional case markers and adpositions is not straightforward on a functional level is a very common phenomenon, especially since adpositions may fulfil very different functions of argument or adjunct marking, depending on the verb’s argument structure and other factors (see Asbury et al. 2008: 10–11; Hagège 2010: 23ff).

<sup>30</sup> In this study, I have been using the term ‘postpositions’ for the formal level, and ‘case’ for the functional level. Formally, the postpositions are a shell around the inner one of morphological case marking, even if the latter remains only weakly functionally productive. Functionally, the postpositions mark both arguments and adjuncts and I use ‘case’ for both in a wide sense.

<sup>31</sup> A similar situation can be found in the language of Kabir; see Strnad (2013) and the discussion of ex. (62) in section 3.5).

locations, but also arguments such as animate O-arguments in transitive clauses (ex. (246)) and A-arguments with perfective participles (i.e. ergative) (ex. (247)):

- (246) *saṭhanhi*                      *rāma-sammukha*    *kō*    *karata*  
 wicked.OBL.PL.M    Rāma-facing            who    do.IPFV.SG.M  
 ‘Who would have turned the wicked towards Rāma.’  
 (Tulsidas, Saksena 1971: 123)

- (247) *sakhinha*                      *kahā*: ...  
 friend.OBL.PL.M    say.DIR.SG.M  
 ‘The friends said: ...’  
 (Jayasi, Saksena 1971: 123)

Gradually, the bare oblique is replaced in these functions by postpositions (that in turn govern the oblique); in standard Hindi by *ko*<sup>32</sup> for animate and/or definite O-arguments, and by *ne* for A-arguments in perfective clauses. The direct, by contrast, continues to mark arguments and has functionally merged with the postpositions into a paradigm, presenting its formally unmarked<sup>33</sup> and functionally ‘basic’ member. Compare the following examples illustrating how argument marking is split among the direct, *ko* (appearing on O-arguments when animate and/or definite) and *ne*:

- (248) *maĩ*    *tasvīr*                      *dekh*    *rahī*                      *hū̃*  
 1SG    picture.DIR.SG.F    see    PROG.SG.F    be.1SG  
 ‘I am looking at the picture.’
- (249) *maĩ*    *laṛkī*                      *ko*    *dekh*    *rahī*                      *hū̃*  
 1SG    girl.SG.F    ACC    see    PROG.SG.F    be.1SG  
 ‘I am looking at a/the girl.’
- (250) *maĩ*    *ne*    *tasvīr*                      *dekhī*  
 1SG    ERG    picture.DIR.SG.F    see.PFV.SG.F  
 ‘I saw the picture.’
- (251) *maĩ*    *ne*    *laṛkī*                      *ko*    *dekhā*  
 1SG    ERG    girl.DIR.SG.F    ACC    see.PFV.SG.M  
 ‘I saw a/the girl.’

This integration of the postpositions and the direct case into a single paradigm of elements marks a difference from strategies of case marking in modern Germanic or

<sup>32</sup> I reserve the term ‘reinforcement’ for *ne*, which was borrowed into a pre-existing construction. *Ko*, by contrast, can be assumed to have grammaticalized with a genitive- or oblique-marked dependent, replacing the bare oblique, rather than being added on to it.

<sup>33</sup> Focusing on functional contrast, I am glossing over the morphological difference between the direct and the oblique, i.e. the form in which a nominal element appears when governed by a postposition.

Romance languages. Here, adpositions are much more restricted to local case relations and arguments are differentiated by inflecting articles, inflecting pronominal forms, and word order.<sup>34</sup> This takes us back to Chapter 4 and the question of what gram types are at the origin of adpositions in modern Indo-Aryan languages in contrast with other Indo-European branches. The local case forms in Germanic, Italic, and other western Indo-European branches were reinforced by the local particles. When these strings were reanalysed as prepositional phrases, the local case forms lived on inside them, eroding phonologically over the course of time. This may have been one factor in the far-reaching restriction of prepositions to local case functions. Instead, asymmetrical groups contrasting with morphological forms evolved into postpositional phrases in Indo-Aryan, functionally unrestricted by any old local case forms.<sup>35</sup>

#### 8.4 Converging scenarios

How to interpret the evidence on distribution and complementarity from recent centuries and from contemporary Hindi? Aspects both of epiphenomenality and of interaction can be identified. First, the formation of what seems like a single template became possible only through an overlap in possible constructional environments of the various grammaticalizing postpositions. This matches Bybee and Dahl's (1989) stance that structural similarities in the source constructions are also the reason for similarities after a period of grammaticalization. As shown in section 8.3.1.3, structural dissimilarities at the outset may also survive, at least in the case of high-frequency forms such as pronouns. Even after postpositions emerge and acquire argument marking functions, certain old pronominal forms remain in use, rather than merging into the general oblique otherwise governed by the postpositions. Remnants are found especially with 1st and 2nd person forms, which are highly frequent in argument function as speech act participants. For instance, old accusative forms remain in use in the pronominal domain even after *ko* enters the scene. Similarly, the old instrumental *maĩ* survives combining with its reinforcer *ne*, rather than giving way to the new oblique form *mujh*.<sup>36</sup> The remaining deviating forms

<sup>34</sup> Of course, this is not to say that the modern Germanic and Romance languages do not at all employ adpositions for argument marking. French *à* and English *to*, for instance, do mark arguments. However, this is mostly restricted to indirect objects and is not comparable to the role postpositions play in New Indo-Aryan, where sometimes every nominal expression in a clause is marked by a postposition (see ex. (251)).

<sup>35</sup> Butt and Ahmed (2011) also note a difference between the branches. In contrast with New Indo-Aryan postpositions, these authors do not consider prepositions in modern Germanic or Romance languages as case markers: 'Given the differing histories of Indo-Aryan versus Germanic and Romance languages with respect to case, a question that naturally arises is why a difference should exist: why does one language family innovate new case markers [i.e. Indo-Aryan, UR], while others do not [i.e. Germanic and Romance, UR]?' (Butt and Ahmed 2011: 548).

<sup>36</sup> The usage of *ne* with *tū* poses a challenge at first glance, as it was precisely the instrumental and not the nominative which encoded the more agentive argument in transitive perfective clauses. Kellogg (1972 [1875]: 209) offers a solution for this problem, hypothesizing that *tū* in combination with *ne* goes back to



almost all go back to genitives (or genitive-datives). Here, the pronominal forms do adapt to the general schema of appearing as genitive forms when governed by a postposition. However, being generally frequent as pronominal forms in core argument function or marking possessors, old forms survive, rather than giving way to the new syncretic oblique forms.

While the similarities and dissimilarities in the constructional context found today can be viewed as epiphenomenal of earlier similarities and dissimilarities, there are also signs of interaction and even slot-formation. Again, frequency effects are suggestive. The different elements also occur in constructions not shared by the others (e.g. only *UPARI*, but not *MADHYE*, is attested to have construed with accusatives besides genitives and locatives: see Chapter 5; and *kā/-e/-ī* as a participial has wholly different properties), but they grammaticalize precisely in the same constructional pattern. It appears that the model of relational noun expressions plays a role here, as it already did in the post-Vedic genitive shift. Possibly the best evidence for interaction is the insertion of *ne* into the postpositional slot, reinforcing the oblique. The emergence of slots (both the postpositional slot and further distinct slots inside nominal expressions) is reminiscent of the formation of the determiner slot in the Dutch NP as outlined by van de Velde. However, in contrast to van de Velde's scenario I cannot detect any type of causal relation that the slot is exerting on grammaticalizing elements.

In the realm of complementarity, we can equally find signs of epiphenomenality and of interaction. The complex and idiosyncratic semantic spaces of *MADHYE* and *UPARI*, for instance, can be derived from their lexical sources, as described in detail in Chapter 7. The functional spaces of *ne* and of the direct, by contrast, are characterized by their oppositions to other postpositions with argument marking properties. *Ne*, in fact, is likely to have been borrowed only with precisely the function that was still 'vacant,' that is, which did not yet have a postpositional expression.

To summarize, in the domains both of distribution and of complementarity, aspects of epiphenomenality and aspects of interaction appear to be closely interconnected. For instance, the formation of a shared constructional environment can be described as epiphenomenal, as it builds on earlier overlaps between the grams' syntactic properties; but interaction can also be identified, in that the most frequent pattern overall grammaticalizes. Thus, the paradigmaticization of the Hindi postpositions defies a clear-cut model based either only on epiphenomenality or only on slot-formation. Despite the fact that the notion of paradigmaticization has played a certain role in recent publications on grammaticalization, this two-fold dimension has gone largely unrecognized and unexplored to date.

the old genitive form of the 2nd personal pronoun *tava* instead of nominative *tvam*, which would bring it in line with the other old genitive forms found with *ne*.

Finally, one aspect of the development remains unexplained by either of the two scenarios, namely that the earliest elements to grammaticalize, *MADHYE* and *UPARI*, remain comparatively weakly 'paradigmatic', and elements that grammaticalize later, in particular *ne*, show much stronger paradigmaticity in terms of both distribution and complementarity. I propose that this difference in chronology in comparison to the case of the Dutch determiner has to do with the different semantics of adpositions and of articles. Adpositions originate in non-core case functions, typically adding semantic content to spatial relations, and only slowly spread to core case relations. Articles, by contrast, at first grammaticalize in argument expressions where definiteness and specificity add important information on the discourse-functional status of the respective participant. As expected, these arise in argument positions and only later spread to adjuncts (see Himmelmann 2005: 82).

## 8.5 Conclusion

In this chapter, I have taken into account evidence that can be gathered about the recent history of the other postpositions besides *mē* and *par* in order to construct a scenario of how the paradigm we find today came into being. A striking aspect of the Hindi postpositions is that they descend from several different syntactic classes and acquire functions of case marking at different points in time, which suggests a process of convergence. However, when testing two scenarios that either deny or accord an independent status to the paradigmaticization of function words, it turns out that there is no either-or, but that the developments of the different elements show evidence for both epiphenomenality and interaction. Effects of epiphenomenality can be detected in the overlap in constructional possibilities, making possible grammaticalizations within the same environment, and effects are also found in retentions of earlier high-frequency pronominal forms, as well as in the idiosyncratic semantic spaces of elements such as *mē* or *par*. On the other hand, there is evidence for interaction and true slot-formation. For example, the fact that the most frequent construction type is selected for grammaticalization suggests interaction between grams. Moreover, not only the postpositional slot was formed, but also the slots for referential elements and adnominal modifiers. Furthermore, the integration of new forms shows how complementarity determines the functional range of an individual element. It appears, then, that both epiphenomenality and interaction in the form of slot-formation occur, with frequency effects as a crucial factor in both domains.

Importantly, elements may appear to form a tightly integrated paradigm with other elements, but this does not allow for the conclusion that their origin is uniform, either regarding their structural source, or regarding their time of grammaticalization. Instead, their syntagmatic as well as paradigmatic properties may at least in part be determined by interaction among themselves. Such interaction, it seems, has a selective effect: for example the most frequent construction is typically the one

that grammaticalizes, while other constructional options do not survive. Note that this is always a selection of properties within the realm of the functional and formal properties that the gram in question already had at the outset, for instance occurring in a specific syntagm, or already possessing a certain function (as in the case of *ne*). The selection is interactive, and determined by frequency of usage, but the choice of options that can be selected from pre-exists.<sup>37</sup>

<sup>37</sup> Extreme cases, where the gram develops in a way which is beyond its own original properties, exist also, as in the reanalysis of *UPARI* and other adverbs in the post-Vedic genitive shift. Such developments, however, are likely to be very rare, needing a considerable systemic pressure. In the post-Vedic genitive shift, the proliferation of relational noun expressions exerted such pressure.



## Conclusion

In this book, I have examined the question of how configurational structures may arise for the first time in a branch of languages. Indo-Aryan offers a rare—and in such detail probably the only—documentation of this development, ranging from far-reaching non-configurationality in early Old Indo-Aryan, which lacks any phrasal structures, to low-level configurationality in New Indo-Aryan, involving postpositional phrases. The rise of phrasality, I argue, is a by-product of the grammaticalization of postpositions. In particular, semantic changes undergone both prior to and during their grammaticalization lead to the establishment of phrasal structure. In my corpus study, I focused on the two Hindi postpositions *mē* ‘in’ and *par* ‘on’, which can be traced back particularly far, to early Old Indic. Additionally, data on the remaining simple postpositions of Hindi were drawn upon.

The semantic changes leading to the emergence of phrase structure can be summarized as follows. First, metaphorical extensions of the types of ground that the precursor forms of *mē* and *par*—*MADHYE* (Skt. ‘in the middle’) and *UPARI* (Skt. ‘above’)—can refer to lead to an increase in the overt expression of these grounds as genitives. I showed that the ground can be dropped when it is a basic spatio-temporal one (e.g. ‘in the middle of a lake’) and when it is inferable from context, while it is always expressed in cases of derived usages. For example, when ‘middle’ refers to a spatio-temporal ground lacking an intrinsic centre (as in ‘in the middle of the world’), or to a non-spatial ground (as in ‘in the middle of his heart’), an overt possessor is necessary. Analogically, in the case of *UPARI*, when referring to some ground which does not allow for a literal reading, for example ‘be angry at s.o.’ (literally ‘above s.o.’), the ground is also always overt. The increasing occurrence of such derived usages across the historical stages led to a steady increase in overt possessors, paving the way for syntactic obligatoriness constraints—one of the characteristics of the emerging phrasal structures. By way of their non-literal readings, the increasing usage of these types of ground also paved the way for the second type of semantic change: metonymic shifts which turn *MADHYE* and *UPARI* from elements denoting the complex spatial relations *IN MIDDLE* and *ABOVE* into elements denoting simple spatial relations meaning *IN* and *ON/OVER*. These shifts could come about as an effect of emphatic usage, for instance. *MADHYE* started being used in Middle Indic

times for emphatic expressions of containment, such as ‘in the middle of my heart, I feel . . .’ rather than simply ‘in my heart’. Through frequency effects, emphasis wears off and we are left with a designation of simple containment, that is *IN*. Regarding the types of construction these changes apply to, there is a certain amount of evidence that these metonymic shifts took place only in the specific syntagm GEN MADHYE/UPARI (or OBL MADHYE/UPARI after syncretism). This underscores the assumption that grammaticalizations take place in constructions with a fixed order. Through these metonymic changes in a specific syntagm, the second crucial trait of the modern phrases besides the obligatorification of dependents is in place—the fixation of the postpositional order.

In order to reach this central part of the development discussed in Chapter 7, where the semantics–syntax interface is crossed, several issues and topics had to be clarified and discussed first (Chapters 4–6). To begin with, a different scenario of the rise of the Indo-Aryan postpositions has so far been assumed in the literature, paralleling the well-attested changes in other branches of Indo-European. The scenario involves an origin of the postpositions in the adverbial category of Proto-Indo-European local particles. However, the local particles vanish as free forms with nominal orientation (i.e. in usages foreshadowing adpositional status) at an early point in time in Old Indic, unlike in other Indo-European branches. I showed that the precursor forms of the modern postpositions cannot be linked to this old category of local particles, since there is hardly any temporal overlap and since the local particles prefer the prenominal slot, in contrast to the ancestor forms of the modern simple postpositions. In order to account for the early loss of the local particles with nominal orientation in Old Indic, I proposed that a prosody–semantics mismatch disfavoured survival in this function.

Instead of the local particles, relational nouns, nominalized adverbs, and a participle stand at the origin of the modern Hindi simple postpositions. Crucially, we are not dealing only with etymological sources different from those of modern adpositions in other Indo-European branches, but these elements combine with (different types of) dependents, whereas the local particles in other branches grammaticalized in apposition with local case forms. In order to identify the correct source constructions, it proved crucial to take into account data from Middle Indo-Aryan, as looking only at Vedic would suggest an origin in constructions with local case forms similar to those in more western branches of Indo-European.

The difference between the sources of adpositional phrases in Indo-Aryan and other branches of Indo-European attests to two alternative paths to phrase structure in a non-configurational language: either we may find an origin in symmetrical groups of syntactically co-ranking and semantically co-referential elements, as in the case of the local particles developing into adpositions in western branches of IE, or we find an origin in asymmetrical groups where one element is the formal and semantic head, as in Indo-Aryan. Of course, such asymmetric constructions

are known to give rise to adpositional constructions in the literature on grammaticalization; however, when it comes to non-configurational languages, only cases of symmetrical groups have so far been discussed (e.g. Vincent 1999 and Hewson and Bubeník 2006 on local particles in other branches of Indo-European). Note that 'symmetrical' also covers source constructions giving rise to other types of phrase. For example, demonstratives and nouns are equally co-ranking and co-referential, as discussed in Himmelmann (1997) and van de Velde (2010). Another type of asymmetrical construction besides relational noun expressions is the combination of an inflected verb with an infinitive developing into an auxiliary construction, for example the well-known case of Latin *cantare habeo* developing into a future construction. This dichotomy between symmetrical and asymmetrical source constructions marks a general distinction between two sorts of grammaticalization process. Formally, the reanalysis of symmetrical groups as phrasal units is concomitant with the birth of a construction, and there may be no major semantic changes involved. On the other hand, an asymmetrical group is a construction from the start characterized by internal dependency relations, and gradual semantic changes eventually lead to the rise of phrasality.

Having identified the source constructions, I inquired into the factors at play which gave rise to the postpositional syntagm, rather than to a prepositional one, in Indo-Aryan. I argued that it is not enough to point out that the pattern with the preceding genitive was already frequent in earlier times. The fixation of a certain pattern in a non-configurational language needs to be explained, because the speakers previously seemed to be content with a flexible word order. I argued that an important factor in the obligatorification were constructions with pronominal possessors and particularly genitive-forms of *sa-/ta-*stem demonstratives serving as possessives, which had had clause-linking functions since earliest times. Serving these functions, these elements (in the same way as other inflectional forms of this paradigm) have a general leftward tendency in Old and Middle Indic, virtually always preceding *MADHYE/UPARI*, and frequently clause-initial. Often, the entire string stands at the left clausal edge. The only major exception is when these elements are clitics and placed in 2nd position, but this insertion into P2 vanishes towards Apabhramsha. Through the semantic changes outlined in Chapter 7, pronominal possessors become more frequent over time due to the increase in derived metaphorical usages. In cases where the ground is already activated, it is often realized pronominally. The increase in pronominal possessors then induces a more frequent occurrence of the specific syntagmatic pattern of the genitive preceding—a pattern which finally becomes obligatory when the metonymic shifts from complex to simple spatial relations take place.

In the course of this book (especially Chapters 5–7), the three claims formulated in section 2.6 were substantiated: the individual trajectories of different postpositions bring about the earliest phrasal constructions in Indo-Aryan; the postpositional constructions arise from asymmetric groups; and semantic changes entail the

formation of syntactic phrases. In Chapter 8, I added observations regarding the individuality of the grammaticalization paths on the one hand, and the general rise of a new class of function words on the other. I addressed the question of whether, above and beyond the individual grammaticalizations, we can observe some effects of paradigmaticization of the various elements into a single functional class. This question arises because the different simple postpositions of Hindi constitute a functionally and formally very homogeneous class, but descend from heterogeneous syntactic classes and periods. Thus, in an account of the emergence of the postpositional forms, we must ask ourselves how these elements could end up forming an integrated paradigm. In several respects, it seems to have been an epiphenomenal overlap of constructional similarities (and dissimilarities) that gave rise to today's paradigm. However, certain dynamics of interaction and slot-formation can also be identified. As for epiphenomenality regarding similarities (such as grammaticalization in an environment with a preceding genitive) or dissimilarities (such as the retention of earlier pronominal forms), the high-frequency patterns are the ones which survive. As for interactions there are levels on which the postpositional slot has acquired a reality beyond the individual postpositions whose grammaticalization gave rise to it. This shows in its accommodation of ergative *ne* and of the unmarked direct case, and also in the secondary formation of a distinction between a slot for referential and for modifying elements. This sharpening of the noun–adjective distinction arises through the grammaticalization of the postpositions specifically to the right of referentially-used elements. The evidence of this sharpening gives support to the assumption of a connection between non-configurationality and a weak noun–adjective distinction—the loss of the former going hand in hand with a loss of the latter.

In conclusion, it is illustrated in this book how syntactic micro changes can bring about syntactic macro changes in language history. The grammaticalization of a small handful of specific lexemes has led to a reorganization of the nominal syntax of the Indo-Aryan branch of languages in the course of the three millennia of its attested history. All formal aspects of this development, be it the obligatorification of dependents, the fixation of order, or the survival and selection of specific constructions rather than others in paradigmaticization, have been shown to result from functional changes. In early stages this involves the selection of expressions for semantic or pragmatic effect; in later stages it involves the entrenchment and fixation of high-frequency patterns.



# Appendix

## Attested examples of MADHYE/UPARI

Abbreviations:

(page number/verse)<sup>C</sup> > occurrence in compound<sup>1</sup>

(page number/verse)<sup>R</sup> > repetition<sup>2</sup> of clause / sentence / line within the respective published text<sup>3</sup>

(page number/verse)<sup>U</sup> > unclear passage

<sup>1</sup> This includes cases where MADHYE appears as the second member in a compound. As an initial member, the uninflected stem form MADHYA- appears instead; such cases are excluded, as MADHYA- may express all kinds of semantic roles. There are some examples especially in Pali where a petrified MADHYE appears as initial member in a compound as well, especially in the phrase *So dhammaṃ deseti ādi-kalyāṇaṃ majjhe-kalyāṇaṃ pariyosāna-kalyāṇaṃ* . . . (and variants thereof), either written as a compound or separately, where the string parallels other uncontroversial compounds. Such cases are also classified as compounded and do not enter the analysis. Also UPARI can sometimes occur as the initial member in compounds. At times, it cannot be decided whether it is the first member of a compound or stands side by side with a spatial noun. In such cases, the interpretation of the editor is followed, i.e. whether it is spelled separately or as one word.

<sup>2</sup> In order to avoid too strong a skewing of the data through repetitions of the same clause, sentence, line, or even entire paragraph, especially frequent in the Brāhmaṇas and the Pali corpus, I only count a clause (allowing for minor variations such as an inserted emphatic particle, a different pronoun, or syntagmatic rearrangement) once per published work (see bibliography of primary sources). Repetitions of Rigvedic verses are marked as repetitions on a general basis, having a strongly quotative character.

<sup>3</sup> Instances where the gram appears both in a compound as well as in a repeated clause, are marked for compound only.

TABLE A1. MADHYE (all attested examples)

Language	Text	Abbr.	Vol./ Subsection	All occurrences (either verse no. or page no. given); examples in bold were included in the quantitative analysis
Vedic	Rigveda	RV		1.32.10, 1.33.11, 1.69.4, 1.105.10, 1.105.11, 1.108.12, 1.158.3, 1.164.30, 1.182.7, 3.14.2, 4.58.5, 5.1.6, 5.44.3, 5.47.3, 6.12.1, 7.41.4, 7.49.3, 7.68.7, 7.89.4, 8.27.20, 8.34.18, 8.70.10, 9.65.23, 10.5.1, 10.11.2, 10.15.14, 10.102.5, 10.102.9, 10.138.3, 10.139.2
	Yajurveda	YV		13.38, 13.42, 13.49, 15.52, 17.59, 17.60, 17.87, 17.93 <sup>R</sup> , 19.60, 21.54, 32.2, 34.37 <sup>R</sup>
	Atharva-veda	AV		1.33.2 <sup>R</sup> , 3.16.4 <sup>R</sup> , 4.15.14, 4.15.15, 5.19.3, 6.106.2, 9.10.8 <sup>R</sup> , 10.7.38, 10.8.15, 13.2.36, 18.1.19 <sup>R</sup> , 18.2.35 <sup>R</sup> , 19.13.9
	Śāṅkhāyana-brāhmaṇa	ŚāṅB		7.4a, 7.4b, 7.4c <sup>R</sup> , 7.4d <sup>R</sup> , 11.3, 11.5, 12.3, 14.1a, 14.1b, 14.3, 15.3a, 15.3b, 15.4a <sup>R</sup> , 15.4b <sup>R</sup> , 16.7a, 16.7b, 18.9, 21.5, 22.8 <sup>R</sup> , 23.8, 25.5, 25.8
	Aitareya-brāhmaṇa	AiB		2.19.1, 2.33.2, 2.33.3, 3.10.2, 3.19.5, 3.35.6a, 3.35.6b <sup>R</sup> , 3.35.6c, 3.41.4, 4.15.6, 4.15.15, 4.18.1, 4.18.4, 4.18.5, 4.19.7, 4.22.1, 4.23.5, 6.27.13, 8.21.10
	Śatapatha-brāhmaṇa	ŚaB		1.1.1.11 <sup>C</sup> , 1.2.5.16a, 1.2.5.16b, 1.7.4.20, 1.8.1.9 <sup>C</sup> , 2.5.3.6, 2.6.1.11a <sup>C</sup> , 2.6.1.11b, 3.2.1.9, 3.3.2.20, 3.5.1.34, 3.5.2.10, 3.5.2.13, 3.6.1.2a <sup>C</sup> , 3.6.1.2b, 3.7.1.12a, 3.7.1.12b, 4.2.2.16, 4.3.1.12 <sup>C</sup> , 5.1.2.19, 5.2.4.5, 5.4.3.9, 5.5.1.1, 6.1.1.2, 6.2.1.24a, 6.2.1.24b, 6.7.4.11, 7.1.1.44 <sup>C</sup> , 7.5.1.15, 7.5.1.23a, 7.5.1.23b <sup>R</sup> , 7.5.2.11, 7.5.2.14, 7.5.2.18, 7.5.2.34, 8.1.2.7, 8.1.3.6, 8.1.3.10, 8.1.4.1 <sup>R</sup> , 8.1.4.3 <sup>R</sup> , 8.2.4.19, 8.4.4.8 <sup>R</sup> , 8.5.1.12, 8.6.1.9, 8.6.1.20, 8.6.1.23, 8.6.2.8, 8.6.2.16 <sup>R</sup> , 8.6.3.4a, 8.6.3.4b, 8.6.3.4c, 8.6.3.11, 8.6.3.29, 9.2.3.17a, 9.2.3.17b <sup>R</sup> , 9.2.3.18a, 9.2.3.18b <sup>R</sup> , 9.3.1.10a, 9.3.1.10b, 9.3.4.18, 9.4.3.6, 10.2.3.1, 10.5.4.15, 11.4.1.15, 11.7.3.3a, 11.7.3.3b, 12.5.1.17, 13.6.1.9, 13.7.1.15, 13.8.3.9, 14.1.2.17a <sup>R</sup> , 14.1.2.17b <sup>R</sup>
Pali	Dighanikāya	Digh	I	62 <sup>C</sup> , <b>83a</b> , <b>83b</b> , 83c <sup>R</sup> , 87 <sup>C</sup> , 88 <sup>C</sup> , 100 <sup>C</sup> , 111 <sup>C</sup> , 128 <sup>C</sup> , 150 <sup>C</sup> , 206 <sup>C</sup> , 209a <sup>C</sup> , 209b <sup>C</sup> , 209c <sup>C</sup> , 225 <sup>C</sup> , 250 <sup>C</sup>
			II	46a <sup>C</sup> , 46b <sup>C</sup> , 47a <sup>C</sup> , 47b <sup>C</sup> , 48 <sup>C</sup> , 135, 176, 235, 274, 337
			III	76a <sup>C</sup> , 76b <sup>C</sup> , 267 <sup>C</sup> , 285 <sup>C</sup>
	Majjhima-nikāya	Majjh	I	27a <sup>C</sup> , 27b <sup>C</sup> , 27c <sup>C</sup> , 179 <sup>C</sup> , 213 <sup>C</sup> , 216 <sup>C</sup> , 225, 276 <sup>C</sup> , 279, 285 <sup>C</sup> , 290 <sup>C</sup> , 337, 344 <sup>C</sup> , 356 <sup>C</sup> , 401 <sup>C</sup> , 469 <sup>C</sup>
			II	8 <sup>C</sup> , 9 <sup>C</sup> , <b>21</b> , 55 <sup>C</sup> , <b>64</b> , 133 <sup>C</sup> , 141 <sup>C</sup> , 146a <sup>C</sup> , <b>146b</b> , 162 <sup>C</sup> , 164 <sup>C</sup> , <b>196</b> , 226 <sup>C</sup> , 250a <sup>C</sup> , 250b <sup>C</sup>
			III	11 <sup>C</sup> , <b>128a</b> , 128b <sup>R</sup> , 129 <sup>C</sup> , 130 <sup>C</sup> , 131 <sup>C</sup> , 134 <sup>C</sup> , <b>166</b> , 175, 178, 183 <sup>R</sup> , 280 <sup>C</sup> , 291 <sup>C</sup>
	Saṃyutta-nikāya	Saṃ	I	17a <sup>C</sup> , 17b <sup>C</sup> , 56 <sup>C</sup> , 57 <sup>C</sup> , 105 <sup>C</sup> , <b>156</b> , 169 <sup>C</sup>
			II	<b>229</b>
			III	<b>93</b>
			IV	122 <sup>C</sup> , <b>179a</b> , 179b <sup>R</sup> , <b>180a</b> , <b>180b</b> , <b>181</b> , <b>194a</b> , 194b <sup>R</sup> , <b>195</b> , <b>199</b> , 306a <sup>C</sup> , 306b <sup>C</sup> , 306c <sup>C</sup> , 306d <sup>C</sup> , 306e <sup>C</sup> , 306f <sup>C</sup> , 307a <sup>C</sup> , 307b <sup>C</sup> , 307c <sup>C</sup> , 307d <sup>C</sup> , 307e <sup>C</sup> , 307f <sup>C</sup> , 307g <sup>C</sup> , 307h <sup>C</sup> , 308a <sup>C</sup> , 308b <sup>C</sup> , 315a <sup>C</sup> , 315b <sup>C</sup> , 316a <sup>C</sup> , 316b <sup>C</sup> , 317a <sup>C</sup> , 317b <sup>C</sup>
			V	<b>219</b> , 315 <sup>C</sup> , 352 <sup>C</sup>

	Aṅguttara-nikāya	Aṅg	I	68 <sup>C</sup> , 69 <sup>C</sup> , 130a <sup>C</sup> , 130b <sup>C</sup> , 130c <sup>C</sup> , 130d <sup>C</sup> , 131a <sup>C</sup> , 131b <sup>C</sup> , <b>141</b> , 180 <sup>C</sup> , 247 <sup>C</sup> , 248 <sup>C</sup>
			II	<b>95</b> , 208 <sup>C</sup> , 239 <sup>C</sup>
			III	30 <sup>C</sup> , <b>64a</b> , 64b <sup>R</sup> , 113 <sup>C</sup> , 114 <sup>C</sup> , 115 <sup>C</sup> , 120 <sup>C</sup> , 135 <sup>C</sup> , 152 <sup>C</sup> , 155 <sup>C</sup> , 262 <sup>C</sup> , 381a <sup>C</sup> , 381b <sup>C</sup> , 382 <sup>C</sup> , <b>399a</b> , <b>399b</b> , <b>399–400</b> , <b>400a</b> , <b>400b</b> , <b>400c</b> , 400d <sup>R</sup> , <b>401a</b> , 401b <sup>R</sup> , 401–2 <sup>R</sup>
			IV	6 <sup>C</sup> , 152 <sup>C</sup> , 154 <sup>C</sup> , 193 <sup>C</sup> , <b>201</b> , 202 <sup>R</sup> , <b>205</b> , 361a <sup>C</sup> , 361b <sup>C</sup> , 361c <sup>C</sup> , 362a <sup>C</sup> , 362b <sup>C</sup> , 362c <sup>C</sup>
			V	23 <sup>C</sup> , 26 <sup>C</sup> , 72 <sup>C</sup> , 73 <sup>C</sup> , 80 <sup>C</sup> , 89 <sup>C</sup> , 163 <sup>C</sup> , 198 <sup>C</sup> , 204 <sup>C</sup> , 338 <sup>C</sup> , 341 <sup>C</sup>
	Khuddaka-nikāya (in the order of the Pali Text Society)	Dhamm	Dhamma-pada	36a <sup>C</sup> , 36b <sup>C</sup> , <b>98</b> , <b>118</b>
		Ud	Udāna	46 <sup>C</sup> , <b>52</b> , <b>55</b> , <b>84</b>
		Iti	Itivuttaka	79a <sup>C</sup> , 79b <sup>C</sup> , 79c <sup>C</sup> , <b>90</b> , 111 <sup>C</sup>
		Sutta	Sutta-nipāta	7a <sup>C</sup> , 7b <sup>C</sup> , <b>60</b> , <b>99</b> , 103 <sup>C</sup> , <b>108</b> , <b>122</b> , <b>161</b> , <b>162</b> , <b>180</b> , <b>184</b> , <b>199a</b> , 199b <sup>R</sup> , <b>202</b> , 205 <sup>R</sup> , <b>211</b> , 212 <sup>R</sup> , 213a <sup>R</sup> , <b>213b</b>
		Vim	Vimāna-vatthu	82 <sup>C</sup> , <b>122</b> , 123 <sup>C</sup> , <b>126</b> , 127 <sup>R</sup>
		Peta	Peta-vatthu	<b>11a</b> , <b>11b</b> , 25 <sup>C</sup> , <b>67</b> , 79a <sup>R</sup> , 79b <sup>R</sup>
		Thera	Thera-gāthā	<b>33</b> , <b>68</b> , <b>78</b> , <b>107</b> , <b>113</b>
		Therī	Therīgāthā	144 <sup>C</sup> , 162a <sup>C</sup> , 162b <sup>C</sup>
		Jā	Jātaka I	<b>412</b> , <b>472</b>
			Jātaka II	<b>180</b> , <b>228</b>
			Jātaka III	108 <sup>C</sup> , <b>158</b> , 313 <sup>C</sup> , <b>348</b> , <b>380</b> , <b>396</b> , <b>413</b> , 425 <sup>C</sup> , <b>508</b> , 523 <sup>C</sup> , 541 <sup>C</sup>
			Jātaka IV	19a <sup>C</sup> , 19b <sup>C</sup> , 46 <sup>C</sup> , 60 <sup>C</sup> , 77 <sup>C</sup> , <b>156</b> , 165a <sup>C</sup> , 165b <sup>C</sup> , 310 <sup>C</sup> , 312 <sup>C</sup> , <b>395</b> , 399 <sup>C</sup> , <b>400</b> , 412a <sup>C</sup> , <b>412b</b> , 412c <sup>C</sup> , <b>412d</b> , <b>464</b>
			Jātaka V	<b>55</b> , 89 <sup>C</sup> , <b>169</b> , <b>223</b> , <b>233</b> , <b>302</b> , 360 <sup>C</sup> , <b>417</b>
			Jātaka VI	<b>23</b> , <b>35a</b> , 35b <sup>R</sup> , <b>117</b> , 272 <sup>C</sup> , <b>282</b> , <b>325a</b> , <b>325b</b>
		Nidd	Niddesa I	<b>164a</b> , 164b <sup>R</sup> , <b>164c</b> , 164d <sup>R</sup> , 165 <sup>R</sup> , <b>169a</b> , 169b <sup>R</sup> , 169c <sup>R</sup> , 169d <sup>R</sup> , 170 <sup>R</sup> , <b>353a</b> , 353b <sup>R</sup> , <b>353c</b> , 353d <sup>R</sup> , <b>433</b> , 434a <sup>R</sup> , 434b <sup>R</sup> , 435 <sup>R</sup> , <b>445</b> , <b>449</b>
			Niddesa II	<b>9</b> , 10 <sup>R</sup> , <b>17</b> , 22 <sup>R</sup> , <b>32</b> , 33 <sup>R</sup> , <b>35</b> , 37 <sup>R</sup> , 58 <sup>C</sup> , 59 <sup>C</sup>
		Paṭi	Paṭisambhidā-magga	<b>167a</b> , <b>167b</b> , 168a <sup>R</sup> , 168b <sup>R</sup> , 168c <sup>C</sup> , 168d <sup>R</sup> , 168e <sup>R</sup> , 169a <sup>R</sup> , 169b <sup>R</sup> , 169c <sup>R</sup> , 169d <sup>R</sup> , 169e <sup>R</sup> , 170a <sup>R</sup> , 170b <sup>R</sup> , 170c <sup>C</sup>
			Apadāna I	8a <sup>C</sup> , 8b <sup>C</sup> , 22 <sup>C</sup> , 27 <sup>C</sup> , 51 <sup>C</sup> , <b>326</b>

(continued)

TABLE A1 (continued)

Language	Text	Abbr.	Vol./ Subsection	All occurrences (either verse no. or page no. given); examples in bold were included in the quantitative analysis
Pali (cont.)	Khuddaka-nikāya (cont.)	Apa	Apadāna II	348 <sup>C</sup> , 349 <sup>C</sup> , 387, 390 <sup>C</sup> , 460 <sup>C</sup> , 461 <sup>C</sup> , 465 <sup>C</sup> , 467 <sup>C</sup> , 474 <sup>C</sup> , 489 <sup>C</sup> , 491 <sup>C</sup> , 495 <sup>C</sup> , 540 <sup>C</sup> , 570 <sup>C</sup> , 579, 587 <sup>C</sup> , 588 <sup>C</sup>
		Buddha	Buddha-vaṃsa	44, 77a <sup>C</sup> , 77b <sup>C</sup>
		Cari	Cariyā-pitaka	7, 12, 18
Apa-bhramsha	Bhavisattakaha	BH		38.2 <sup>C</sup> , 42.6, 46.5, 46.8, 82.8, 130.1, 193.8, 207.16, 207.17, 250.14, 284.8, 284.10, 294.8
	Karakaṇḍacariu	KA		2.18.2, 2.20.8, 6.16.1, 7.12.4, 8.5.11, 8.15.3, 9.8.4, 9.9.7, 9.14.13, 10.2.6
	Paumacariu	PA		1.11.3, 5.11.9, 6.3.6, 15.13.8
	Sanatkmāra-caritam	SA		484.3, 504.8, 514.6, 557.1, 566.5, 595.7 <sup>C</sup> , 689.5, 707.4, 713.7, 732.3 <sup>C</sup> , 755.5, 756.5 <sup>C</sup>
Old Awadhi	Rāmacarita-mānasa <sup>4</sup>	RA		<p><b>mahā</b> 2.42, 48.3, 2.111.2, 2.113.3, 2.129.1, 2.136.2, 2.203.5, 2.232.4, 2.234, 3.15, 3.18.1, 3.19, 3.28.6, 3.42.2, 3.54, 4.3.5, 5.3.1, 5.56.4, 5.57.2, 6.5, 6.12.1, 6.16.2, 6.16.3, 6.23.4, 6.28, 6.43, 6.52.1, 6.56.2, 6.66.6, 6.71.4, 6.81.1, 6.90, 6.90, 6.91.1, 6.108, 6.114.5, 6.115.6, 6.118.2, 6.126.5, 6.129.4, 6.133.2, 6.135.5, 7.85.2, 7.110, 7.120, 7.123.4, 7.125.3, 7.131.3, 7.132.4, 7.134, 7.155, 7.156.4, 7.165, 7.172.4, 7.222.2</p> <p><b>mahū</b> 1.19.1, 1.108.5, 1.109, 1.141, 1.157, 2.36.2, 2.118.1, 2.228.3, 3.24.3, 3.24.9, 3.24, 3.30.1, 3.33.8, 3.34.8, 3.35.13, 3.43.3, 3.46, 4.16.4, 4.18, 4.21.1, 4.22.3, 4.25.4, 4.32, 5.5.4, 5.6.1, 5.6.4, 5.7.1, 5.8.4, 5.9.1, 5.10.5, 5.20.4, 5.23.1, 5.27.3, 5.37.1, 5.45.4, 6.45.2, 6.63.4, 6.83, 6.91.3, 6.110.5, 6.116.2, 6.122.1, 6.124, 6.125, 6.133.5, 6.134.7, 6.153.1, 7.5, 7.7.1, 7.9.1, 7.15.4, 7.68.4, 7.78.1, 7.78.2, 7.78.3, 7.97, 7.102</p> <p><b>mājha</b> 1.201.4, 2.48.1, 2. 2.165.2, 206.3, 3.13.8, 3.24.10, 3.27, 4.10, 6.52.4, 6.56.2, 6.91.3, 6.99.8, 6.119.4, 6.119.5, 6.120.1, 6.123.8, 7.119.2, 7.172.1</p> <p><b>māhi</b> 1.11.3, 1.19.4, 1.22.6, 1.41.2, 1.54.3, 1.70.1, 1.78.3, 1.80.1, 1.84.4, 1.91.3, 1. 94.3, 1.95.4, 1.96.3, 1.105.2, 1.107.3, 1.118.2, 1.122.3, 1.126.3, 1.133.4, 1.143.3, 1.153.4, 1.155.3, 1.159.3, 1.164.1, 1.174.2, 1.176.1, 1.177.2, 1.179.1, 1.184.3, 1.192.1, 1.193.1, 1.194.3, 1.196.2, 1.197.1, 1.217.3, 1.220.2, 1.221.1, 1.240.2, 1.243.6, 1.253.3, 1.258.3, 1. 264.4, 1.270.4, 1.278.1, 1.282.1, 1.283.4, 1.287.2, 1.291.4, 1.300.3, 1.307.2, 1.310.4, 1.327.3, 1.390.2, 2.3.2, 2.5.2, 2.8.4, 2.31.2, 2.34.1, 2.43.2, 2.45.3, 2.72.1, 2.82.3, 2.85.3, 2.93.4, 2.111.2, 2.116.2, 2.120.1, 2.121.1, 2.122.3, 2.130.3, 2.131.3, 2.139.2, 2.151.4, 2.163.3, 2.173.1, 2.206.4, 2.209.2, 2.213.1, 2.215.4, 2.228.2, 2.234.4, 2.253.4, 2.266.1, 2.323.1, 3.6.6, 3.13.3, 3.18.4, 3.19.4, 3.20.5, 3.29.1, 3.31.3, 3.37.2, 3.37.7, 3.38.5, 3.42.4, 3.44.5, 4.8.7, 4.9.9, 4.14.1, 4.16.1, 4.23.4, 4.25.2, 4.27.3, 4.29.1, 4.30.4, 4.33.3, 5.5.3, 5.7.2, 5.15. 4, 5.17.5, 5.26.3, 5.32.4, 5.37.5, 5.42.2, 5.43.2, 5.48.2, 5.49.3, 5.50.5, 5.58.2, 6.4.3, 6.16.4, 6.44.3, 6.50.5, 6.67.4, 6.83.6, 6.108.4, 6.113.3, 6.115.2, 6.135.4, 7.7.7, 7.34.2, 7.35.1, 7.44.2, 7.47.4, 7.70.3, 7.77.3, 7.81.2, 7.82.4, 7.83.2, 7.102.4, 7.119.1, 7.127.1, 7.150.4, 7.164.4, 7.166.3, 7.178.8, 7.192.2, 7.194.3, 7.206.3, 7.206.5, 7.208.7, 7.223.2</p>

<sup>4</sup> The passages in the *Rāmacarita-mānasa* are not categorized regarding repetitions or occurrence in compounds, as they figure only peripherally in the quantitative analyses in this study.

TABLE A2. UPARI (all attested examples)

Language	Text	Abbr.	Vol./ Subsection	All occurrences (either verse no. or page no. given); examples in bold were included in the quantitative analysis
Vedic	Rigveda	RV		1.24.7, 1.34.8, 4.31.15, 5.61.12, 8.33.19, 8.47.2, 8.101.9, 9.54.3, 10.34.9, 10.75.3 <sup>U</sup> , 10.129.5
	Yajurveda	YV		13.58, 15.19, 18.44, 33.74 <sup>R</sup> , 33.85 <sup>R</sup>
	Śatapatha-brāhmaṇa	ŚB		1.3.4.15, 5.4.4.1, 6.2.2.39a, 6.2.2.39b, 6.2.2.39c, 6.2.2.39d, 6.4.1.6 <sup>R</sup> , 6.4.3.10 <sup>R</sup> , 6.4.4.18, 6.6.2.14 <sup>R</sup> , 6.6.3.17 <sup>R</sup> , 7.1.1.33a <sup>R</sup> , 7.1.1.33b <sup>R</sup> , 7.1.1.36 <sup>R</sup> , 7.3.2.13 <sup>R</sup> , 7.3.2.19 <sup>R</sup> , 7.4.2.8 <sup>R</sup> , 8.1.2.7a, 8.1.2.7b, 8.1.2.7c, 8.1.4.2 <sup>R</sup> , 8.1.4.10a <sup>R</sup> , 8.1.4.10b <sup>R</sup> , 8.2.4.20a <sup>R</sup> , 8.2.4.20b <sup>R</sup> , 8.3.1.10 <sup>R</sup> , 8.3.4.15a <sup>R</sup> , 8.3.4.15b <sup>R</sup> , 8.4.4.12a <sup>R</sup> , 8.4.4.12b <sup>R</sup> , 8.6.1.20a, 8.6.1.20b, 8.6.1.20c, 8.6.1.20d, 9.2.3.30a, 9.2.3.30b, 9.2.3.35 <sup>R</sup> , 9.3.4.10 <sup>R</sup> , 9.4.1.14a <sup>R</sup> , 9.4.1.14b <sup>R</sup> , 9.4.1.15a <sup>R</sup> , 9.4.1.15b <sup>R</sup> , 9.4.1.16a, 9.4.1.16b, 9.4.2.11a, 9.4.2.11b
Pali	Dīgha-nikāya	Dīgh	I	47 <sup>C</sup> , 95, 112 <sup>C</sup> , 128 <sup>C</sup>
			II	172a <sup>C</sup> , 172b <sup>C</sup> , 317 <sup>C</sup> , 325
			III	60 <sup>C</sup> , 61a <sup>C</sup> , 61b <sup>C</sup> , 85, 155
	Majjhima-nikāya	Majjh	I	44a <sup>C</sup> , 44b <sup>C</sup> , 44c <sup>C</sup> , 45 <sup>C</sup> , 46 <sup>C</sup> , 213 <sup>C</sup> , 231, 306a, 306b <sup>R</sup>
			II	79, 93, 117, 164, 214 <sup>C</sup>
			III	130a <sup>C</sup> , 130b <sup>C</sup> , 130c <sup>C</sup> , 131a <sup>C</sup> , 131b <sup>C</sup> , 131c <sup>C</sup> , 131d <sup>C</sup> , 131e <sup>C</sup> , 131f <sup>C</sup> , 131g <sup>C</sup> , 133, 172 <sup>C</sup> , 184, 302
	Saṃyutta-nikāya	Saṃ	I	75 <sup>C</sup> , 144a, 144b <sup>R</sup> , 144c <sup>R</sup> , 145a <sup>R</sup> , 145b <sup>R</sup> , 145c <sup>R</sup>
			II	32 <sup>C</sup> , 184, 231
			III	188 <sup>C</sup>
			IV	196a, 196b <sup>R</sup> , 196c <sup>R</sup> , 196d <sup>R</sup>
			V	168a <sup>C</sup> , 168b <sup>C</sup> , 396, 437, 438 <sup>R</sup> , 439 <sup>C</sup> , 445
	Aṅguttara-nikāya	Aṅg	I	243
			II	140
			III	243a, 243b
			IV	404
			V	114 <sup>C</sup> , 115 <sup>C</sup> , 117 <sup>C</sup> , 119 <sup>C</sup>
	Khuddaka-nikāya	Ud	Udāna	10, 47a <sup>C</sup> , 47b <sup>C</sup>
			Jā	Jātaka III 207 <sup>C</sup> , 433
		Jā	Jātaka IV	4 <sup>C</sup> , 365
			Jātaka VI	269 <sup>C</sup> , 528, 534 <sup>R</sup> , 550a, 550b, 553a <sup>R</sup> , 553b <sup>R</sup>
		Nidd	Niddesa I	353, 405
			Niddesa II	57
		Apa	Apadāna I	273, 276, 306 <sup>C</sup> , 334 <sup>C</sup>
			Apadāna II	380 <sup>C</sup> , 384 <sup>R</sup> , 442 <sup>C</sup> , 608
		Cari	Cariyā-pitaka	18 <sup>C</sup> , 19

(continued)

TABLE A2 (continued)

<i>Language</i>	<i>Text</i>	<i>Abbr.</i>	<i>Vol./ Subsection</i>	<i>All occurrences (either verse no. or page no. given); examples in bold were included in the quantitative analysis</i>
Apa- bhramsha	Bhavisatta- kaha	BH		17.8 <sup>C</sup> , <b>110.1</b> , <b>167.2</b> , 182.7 <sup>C</sup> , <b>222.3</b> , <b>228.8</b> , 277.6
	Karakaṇḍa- cariu	KA		<b>2.9.1</b> , <b>2.19.10</b> , <b>3.8.14</b> , <b>3.12.10</b> , <b>3.13.8</b> , <b>4.1.13</b> , <b>4.3.8</b> , <b>4.4.4</b> , <b>4.6.1</b> , <b>4.12.1</b> , <b>4.15.2</b> , <b>5.4.7</b> , <b>5.7.9</b> , <b>5.13.6</b> , <b>6.7.6</b> , <b>7.6.7</b> , <b>7.11.13</b> , <b>9.4.6</b> , <b>9.15.5</b> , <b>10.23.2</b>
	Pauma- cariu	PA		<b>1.3.3</b> , 2.2.9 <sup>C</sup> , <b>2.3.4</b> , 2.3.8 <sup>C</sup> , <b>3.1.13</b> , <b>6.6.2</b> , 11.5.2 <sup>C</sup> , 11.14.1 <sup>C</sup> , 11.14.8 <sup>C</sup> , <b>12.11.9</b> , <b>13.1.10</b> , <b>13.2.6</b> , 14.5.6 <sup>C</sup> , 14.7.9 <sup>C</sup> , <b>17.8.10</b> , <b>19.3.2</b> , <b>19.11.4</b>
	Sanatkumā- racaritam	SA		<b>489.5</b> , 510.6 <sup>C</sup> , <b>522.9</b> , <b>589.8</b> , <b>594.1</b> , <b>598.1</b> , <b>650.5</b> , <b>674.5</b> , <b>693.2</b> , 694.8 <sup>C</sup> , 727.6 <sup>C</sup>

# References

## Primary sources

For clarity's sake, I list the editions of the primary texts in the alphabetical order of their titles as they are normally referred to today (with standardized titles in brackets if necessary). Translations drawn upon are listed directly below the primary source.

*Das Aitareya Brāhmaṇa. Mit Auszügen aus dem Commentare von Sāyaṇācārya und anderen Beilagen.* Th. Aufrecht (ed.). Bonn: A. Marcus, 1879.

[Aitareyabrāhmaṇa] *Rigveda Brahmanas: The Aitareya and Kauṣītaki Brāhmaṇas of the Rigveda.* A. B. Keith (transl.). Cambridge (MA): Harvard University Press, 1920.

*The Aṅguttaranikāya* (Vols I–V). Richard Morris (ed.). (Vol I in 2nd revised edn by A. K. Warder). London: Luzac & Company, 1955–1961.

[Aṅguttaranikāya] *The Book of the Gradual Sayings: (Anguttara-nikāya) or More Numbered Suttas* (Vols I–V). F. L. Woodward (ed. and transl.). London: Luzac & Company, 1960–1965.

*The Apadāna of the Khuddaka Nikāya* (Parts I & II). Mark E. Lilley (ed.). London: Oxford University Press, 1925, 1927.

*Apadānapāli.* ed. M. E. Lilley, 1925, 1927; 2 vols. reprinted as one.

*Atharva Veda Sanhita.* R. Roth and W. D. Whitney (eds.), (2nd edn by Max Lindenau). Berlin: Dümmlers, 1924.

[Atharvaveda] *Atharva-Veda Sanhita* (Vols I & II). W. D. Whitney (transl.). Delhi, Patna, Varanasi: Motilal Banarsidass, 1962.

*Bhavisatta Kaha von Dhaṇavāla. Eine Jaina Legende in Apabhramsha.* Hermann Jacobi (ed.). München: Akademie der Wissenschaften, 1918.

*Buddhavaṃsa and Cariyāpiṭaka.* N. A. Jayawickrama (ed.). London, Boston: The Pāli Text Society, 1974.

[Buddhavaṃsa] *The Minor Anthologies of the Pāli Canon (Vol. III): Chronicle of Buddhas (Buddhavaṃsa) and Basket of Conduct (Cariyāpiṭaka).* I. B. Horner (transl.). London, Boston: The Pāli Text Society, 1975.

[Cariyāpiṭaka] *Buddhavaṃsa and Cariyāpiṭaka.* N. A. Jayawickrama (ed.). London, Boston: The Pāli Text Society, 1974.

[Cariyāpiṭaka] *The Minor Anthologies of the Pāli Canon (Vol. III): Chronicle of Buddhas (Buddhavaṃsa) and Basket of Conduct (Cariyāpiṭaka).* I. B. Horner (transl.). London, Boston: The Pāli Text Society, 1975.

*Dhammapada.* O. von Hinüber and K. R. Norman (eds). Oxford: The Pāli Text Society, 1995.

[Dhammapada] *The Dhammapada: A Collection of Verses.* F. Max Müller (transl.). Delhi, Varanasi, Patna: Motilal Banarsidass, 1965 [1881].

*The Dīgha Nikāya* (Vols I–III). T. W. Rhys Davids and J. Estlin Carpenter (eds). London: Luzac & Company, 1947–1949.

- [Dīghanikāya] *Dialogues of the Buddha*. T. W. Rhys Davids (transl.). London: Luzac & Company, 1956, 1965, 1966.
- Itivuttaka*. Ernst Windisch (ed.). London: Geoffrey Cumberlege, Oxford University Press, 1948.
- [Itivuttaka] *The Minor Anthologies of the Pāli Canon (Vol II). Udāna: Verses of Uplift and Itivuttaka: As it was said*. F. L. Woodward and Geoffrey Cumberlege (transl.). Oxford: Oxford University Press, 1948.
- The Jātaka together with its Commentary: being Tales of the Anterior Births of Gotama Buddha (Vols I–VI)*. V. Fausbøll and T. W. Rhys Davids (eds). London: Luzac & Company, 1963–1964.
- [Jātaka] *The Jātaka or Stories of the Buddha's Former Births: translated from the Pāli by Various Hands (Vols I–VI)*. E. B. Cowell (ed.). London: Luzac & Company, 1957.
- Karakaṇḍa Cariu of Muni Kanakāmara*. Hiralal Jain (ed.). Karanja, Berar: Karanja Jaina Publication Society, 1934.
- Majjhima-nikāya (Vols I–III)*. Robert Chalmers (ed.). London: Oxford University Press, 1948.
- [Majjhimanikāya] *The Collection of the Middle Length Sayings (Majjhima-Nikāya) (Vols I–III)*. I. B. Horner (transl.). London: Luzac & Company, 1957, 1967.
- The Milindapañho*. V. Trenckner (ed.). London: Williams and Norgate, 1880.
- [Milindapañha] *The Questions of Kind Milinda*. T. W. Rhys Davids (transl.). Oxford: Clarendon, 1890.
- Niddesa (I, Mahāniddesa & II, Cullanidessa)*. L. de la Vallée Poussin and E. J. Thomas (eds). London: Humphrey Milford (Oxford University Press), 1916, 1918.
- Paṭisambhidāmagga (Vols I & II)*. Arnold C. Taylor (ed.). London, Henley, Boston: The Pāli Text Society, 1979.
- [Paṭisambhidāmagga] *The Path of Discrimination*. Bhikkhu Nāṇamoli (transl.). London, Boston: The Pāli Text Society, 1982.
- Paumacariu of Kavirāja Svayambhūdeva (A Pre-Tenth Century Purāṇic Epic in Apabhramsha)*. Harivallabh C. Bhayani (ed.). Bombay: Bharatiya Vidya Bhavan, 1953.
- [Paumacariu] *Een kritische studie van Svayambhūdeva's Pāumacariu*. Eva de Clercq (ed., transl. and comment.). Ghent: Universiteit Gent, Faculteit Letteren en Wijsbegeerte (dissertation), 2002–2003.
- [Paumacariu] *Jain Rāmāyaṇa Paumacaryu. Rendering into English from Apabhramśa*. Shantilal Nagar (transl.). Delhi: P. R. Publishing Corporation, 2002.
- Rāmacaritamānasa (रामचरतिमानस)*. Tulsīdās. K. Mitra (ed.). Allahabad: Indian Press.
- [Rāmacaritamānasa] *Tulasidasa's Shri Ramacharitamānasa (2nd edn)*. R. C. Prasad (ed. and transl.). Delhi: Motilal Banarsidass, 1990.
- [Rigveda] *Die Hymnen des Rigveda (Vols I & II)*. Theodor Aufrecht (ed.). Wiesbaden: Harrassowitz, 1955.
- [Rigveda] *Rig-Veda: Das heilige Wissen: Erster und zweiter Liederkreis*. Michael Witzel and Toshifumi Gotō (ed. and transl.). Frankfurt am Main: Verlag der Weltreligionen, 2007.
- [Rigveda] *Der Rig-Veda. Aus dem Sanskrit ins Deutsche übersetzt und mit einem laufenden Kommentar versehen von Karl Friedrich Geldner (Harvard Oriental Series 63)*. Michael Witzel (ed.). Cambridge (MA): The Department of Sanskrit and Indian Studies, Harvard University, 2003.



- Sāmaveda Saṁhitā: Sanskrit Text with English Translation*. R. T. H. Griffith (transl.). Ravi Prakash Arya (ed., intro. and rev.). Delhi: Parimal, 1996.
- [Sāmaveda] *Die Hymnen des Sāma-Veda: Einleitung—Glossar—Übersetzung*. Theodor Benfey (ed. and transl.). Darmstadt: Wissenschaftliche Buchgesellschaft, 1969.
- Śāṁkhāyanabrāhmaṇa* [= Kauṣītaki]. *Das Kaushitaki Brāhmaṇa*. B. Lindner (ed.). Jena: Costenoble, 1887.
- [Śāṁkhāyanabrāhmaṇa] *Rigveda Brahmanas: The Aitareya and Kauṣītaki Brāhmaṇas of the Rigveda*. A. B. Keith (transl.). Cambridge, MA: Harvard University Press, 1920.
- The Saṃyuttanikāya of the Suttaṭṭhaka* (Vols I–V). M. Léon Feer (ed.). London: Luzac & Company, 1960.
- [Saṃyuttanikāya] *The Book of the Kindred Sayings: (Saṃutta-nikāya) or grouped suttas* (Vols I–V). T. W. Rhys Davids and Sūriyagoḍa Sumangala Thera (transl.). London: Luzac & Company, 1950–1965.
- Sanatkumāracaritam—ein Abschnitt aus Haribhadrās Nemināthacaritam*. Hermann Jacobi (ed.). München: Akademie der Wissenschaften, 1921.
- [Śatapathabrāhmaṇa] *The Śatapatha-Brāhmaṇa in the Mādhyandina-Çākhā with extracts from the commentaries of Sāyaṇa, Harisvāmin and Dvivedāṅga*. A. Weber (ed.). Berlin, 1849.
- [Śatapathabrāhmaṇa] *The Śatapatha-Brāhmaṇa according to the text of the Mādhyandina School*. Julius Eggeling (transl.). Oxford: Clarendon Press, 1882.
- The Sutta Nipāta*. Dines Andersen and Helmer Smith (eds.). London: Frowde, 1913.
- [Suttanipāta] *The Group of Discourses (Sutta-Nipāta)* (Vols I & II). K. R. Norman, I. B. Horner, and Walpola Rahula (transl.). 1984, 1992.
- The Thera- and Therīgāthā (Stanzas ascribed to Elders of the Buddhist Order of Recluses)* (2nd edn with appendix). Hermann Oldenburg and Richard Pischel (eds); K. R. Norman and L. Alsdorf (eds 2nd edn). London: Luzac & Company, 1966.
- [Theragāthā] *The Elders Verses I Theragāthā*. K. R. Norman (transl.). London: Luzac & Company, 1969.
- [Therīgāthā] *The Elders Verses II. Therīgāthā*. K. R. Norman (transl.). London: Luzac & Company, 1971.
- Udāna*. Paul Steinthal (ed.). London: Oxford University Press, 1948.
- [Udāna] *The Minor Anthologies of the Pāli Canon* (Vol. II). *Udāna: Verses of Uplift and Itivuttaka: As it was said*. F. L. Woodward and Geoffrey Cumberlege (transl.). Oxford: Oxford University Press, 1948.
- Vimānavatthu and Petavatthu*. N. A. Jayawickrama (ed.). London, Henley, Boston: The Pāli Text Society, 1977.
- [Vimānavatthu and Petavatthu] *The Minor Anthologies of the Pāli Canon* (Vol IV). *Vimānavatthu: Stories of the Mansions, Petavatthu: Stories of the Departed*. I. B. Horner and H. S. Gehman (ed. and transl.). London/Boston: The Pāli Text Society, 1974.
- [Vinayaṭṭhaka] *The Vinaya Piṭakam* (Vols II & IV). H. Oldenberg (ed.). London/Edinburgh: Williams and Norgate, 1880, 1882.
- [Yajurveda] *The Vajasaneyi-Saṁhitā in the Mādhyandina and the Kāṇva-Śākhā with the Commentary of Mahidhara*. Albrecht Weber (ed.). Varanasi: Chowkhamba Sanskrit Series Office, 1972 [1852].
- [Yajurveda] *The Texts of the White Yajurveda*. R. T. H. Griffith (transl.). Benares: Lazarus and Company, 1899.

**Secondary sources**

- Agnihotri, Rama Kant. 2007. *Hindi. An Essential Grammar*. London: Routledge.
- Andersen, Paul Kent. 1979. 'Word order typology and prepositions in Old Indic', in Bela Brogyanyi (ed.), *Studies in Diachronic, Synchronic, and Typological Linguistics: Festschrift for Oswald Szemérenyi on the Occasion of his 65th Birthday*. Amsterdam: Benjamins, 23–34.
- Anderson, Stephen R. 1993. 'Wackernagel's revenge: Clitics, morphology, and the syntax of second position', *Language* 69 (1), 68–98.
- Asbury, Anna, Berit Gehrke, Henk van Riemsdijk, and Joost Zwarts. 2008. 'Introduction. Syntax and semantics of Spatial P', in Anna Asbury, Jakub Dotlacil, Berit Gehrke, and Rick Nouwen (eds), *Syntax and Semantics of Spatial P*. Amsterdam/Philadelphia: Benjamins, 1–32.
- Austin, Peter. 2001. 'Word order in a free word order language: the case of Jiwari', in Jane Simpson, David Nash, Mary Laughren, Peter Austin, and Barry Alpher (eds), *Forty Years On: Ken Hale and Australian languages*. Canberra: Pacific Linguistics, 205–323.
- Austin, Peter and Joan Bresnan. 1996. 'Non-configurationality in Australian Aboriginal languages', *Natural Language & Linguistic Theory* 14 (2), 215–68.
- Baker, Mark C. 2001. 'Configurationality and polysynthesis', in Martin Haspelmath and Ekkehard König (eds), *Language Typology and Language Universals (Handbücher zur Sprach- und Kommunikationswissenschaft 20.2)*. Berlin/New York: de Gruyter, 1433–41.
- Baker, Mark C. 2006. *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Bashir, Elena. 1999. 'The Urdu and Hindi ergative postposition *ne*: Its changing role in the grammar', in Rajendra Singh (ed.), *The Yearbook of South Asian Languages and Linguistics*. New Delhi: Sage Publications, 11–36.
- Beames, John. 1970 [1875]. *A Comparative Grammar of the Modern Aryan Languages of India—to wit Hindi, Punjabi, Sindhi, Gujarati, Marathi, Oriya and Bengali (Vol. II. The Noun and Pronoun)*. London: Trübner.
- Benveniste, E. 1952. 'La construction passive du parfait transitif', *Bulletin de la Société de Linguistique de Paris* 48, 52–62.
- Bhat, D. N. S. 1994. *The Adjectival Category. Criteria for Differentiation and Identification*. Amsterdam/Philadelphia: Benjamins.
- Bisang, Walter. 2004. 'Grammaticalization without coevolution of form and meaning: The case of tense-aspect-modality in East and mainland Southeast Asia', in Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer (eds), *What Makes Grammaticalization. A Look from its Fringes and its Components*. Berlin: Mouton de Gruyter, 109–38.
- Blake, Barry J. 1983. 'Structure and word order in Kalkatungu: The anatomy of a flat language', *Australian Journal of Linguistics* 3, 143–75.
- Bloch, Jules. 1934. *L'Indo-Aryen—du Veda aux temps modernes*. Paris: Librairie d'Amérique et d'Orient Adrien—Maisonneuve.
- Bloch, Jules. 1965. *Indo-Aryan. From the Vedas to Modern Times* (revised and translated by Alfred Master). Paris: Librairie d'Amérique et d'Orient Adrien—Maisonneuve.

- Bloomfield, Leonard. 1933. *Language*. London: Allen & Unwin.
- Böhtlingk, Otto and Rudolph Roth. 1855–75. *Grosses Petersburger Wörterbuch*. Online at <http://www.sanskrit-lexicon.uni-koeln.de/scans/PWGSscan/2013/web/index.php>, last accessed 26 October 2015.
- Bortone, Pietro. 2010. *Greek Prepositions: From Antiquity to the Present*. Oxford: Oxford University Press.
- Bouillon, Henri. 1978. 'Kontrastive Syntagmatik der deutschen Präpositionen *auf* und *über*', in J. Lerot and R. Kern (eds), *Mélanges de linguistique et de littérature offerts au professeur Henri Draye à l'occasion de son éméritat*. Louvain: Bibliothèque de l'Université, Bureau de recueil, 3–13.
- Bowern, Claire. 2005. 'Correlates of non-configurationality', *Proceedings of NELS* 36. Amherst, MA: University of Massachusetts, 53–62.
- Bresnan, J. 2001. *Lexical-Functional Syntax*. Malden, MA: Blackwell.
- Brugman, Claudia. 1983. *The Story of Over*. Bloomington: Indiana University Linguistics Club.
- Brugmann, Karl. 1911. *Vergleichende Laut-, Stammbildungs- und Flexionslehre nebst Lehre vom Gebrauch der Wortformen der Indogermanischen Sprachen* (Vol. II, 2). Strassburg: Trübner.
- Bubeník, Vít. 1998. *A Historical Syntax of Late Middle Indo-Aryan (Apabhramsha)*. Amsterdam/Philadelphia: Benjamins.
- Bubeník, Vít. 2006a. 'Cases and postpositions in Indo-Aryan', in John Hewson and Vít Bubeník (eds), 2006. *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: Benjamins, 102–30.
- Bubeník, Vít. 2006b. 'Cases and prepositions in Iranian', in John Hewson and Vít Bubeník (eds), 2006. *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: Benjamins, 131–59.
- Bubeník, Vít and Chitra Paranajpe. 1996. 'Development of pronominal systems from Apabhramśa to New-Indo-Aryan', *Indo-Iranian Journal* 39, 111–32.
- Burston, Monique Andrée. 1977. *A semantic analysis of the Pāli case system*. Cornell University dissertation.
- Butt, Miriam. 2001. 'A re-examination of the accusative to ergative shift in Indo-Aryan', in Miriam Butt and Tracy Holloway King (eds), *Time over Matter. Diachronic Perspectives on Morphosyntax*. Stanford, CA: CSLI, 105–41.
- Butt, Miriam. 2006. 'The dative-ergative connection', in O. Bonami and P. Cabredo Hofherr (eds), *Empirical Issues in Syntax and Semantics* 6, 69–92.
- Butt, Miriam and Tafseer Ahmed. 2011. 'The redevelopment of Indo-Aryan case systems from a lexical semantic perspective', *Morphology* 21 (3–4), 545–72.
- Butt, Miriam and Tracy Holloway King. 2004. 'The status of case', in V. Dayal and A. Mahajan (eds), *Clause Structure in South Asian Languages*. Berlin: Springer, 153–98.
- Bybee, Joan. 1988. 'Semantic substance vs. contrast in the development of grammatical meaning', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 14, 247–79.
- Bybee, Joan. 2003. 'Mechanisms of change in grammaticization: The role of frequency', in B. D. Joseph and J. Janda (eds), *The Handbook of Historical Linguistics*. Oxford: Blackwell, 602–23.

- Bybee, Joan and Östen Dahl. 1989. 'The creation of tense and aspect systems in the languages of the world', *Studies in Language* 13 (1), 51–103.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago/London: University of Chicago Press.
- Caldwell, Robert. 1875 [1856]. *A Comparative Grammar of the Dravidian or South-Indian Family of Languages*. London: Trübner & Company.
- Canedo, José. 1937. *Zur Wort- und Satzstellung in der alt- und mittelindischen Prosa*. Göttingen: Vandenhoeck & Ruprecht.
- Cardona, George. 1970. 'The Indo-Iranian construction *mana (mama) kṛtam*', *Language* 46 (1), 1–12.
- Casaretto, Antje. 2010a. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. IX: *ánu*', *Münchener Studien zur Sprachwissenschaft* 65, 7–64.
- Casaretto, Antje. 2010b. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. VI: *práti*', *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 7, 1–53.
- Casaretto, Antje. 2010 [2011]. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. VIII: *abhi*', *Historische Sprachforschung* 123, 97–156.
- Casaretto, Antje. 2011 [2012]a. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. XVI: *áti* und *tiráś*', *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 8, 173–216.
- Casaretto, Antje. 2011 [2012]b. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. XII: *ví*', *Historische Sprachforschung* 124, 134–77.
- Casaretto, Antje. 2011 [2012]c. 'Lokalpartikel im Ṛgveda. XIV: *áccā*', *Zeitschrift für Indologie und Südasiastudien* 28, 1–27.
- Casaretto, Antje. 2012. 'Syntax und Wortarten der Lokalpartikeln im Ṛgveda. XXI: *purás*, *purástāt* und *purá*', *Münchener Studien zur Sprachwissenschaft* 66, 11–53.
- Chatterji, Sunti Kumar. 1970 [1926]. *The Origin and Development of the Bengali Language*. London: Allen & Unwin.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Cinque, Guglielmo and Luigi Rizzi (eds). 2010. *Mapping Spatial PPs. The Cartography of Syntactic Structures*. Oxford: Oxford University Press.
- Coleman, Robert. 1991. 'Latin prepositional syntax in Indo-European perspective', in Robert Coleman (ed.), *New Studies in Latin Linguistics* (Selected papers from the 4th Colloquium on Latin Linguistics, Cambridge, April 1987). Amsterdam/Philadelphia: John Benjamins, 323–38.
- Compes, Isabel, Silvia Kutscher, and Carmen Rudolf. 1993. *Pfade der Grammatikalisierung: Ein systematisierter Überblick*. Arbeitspapier No. 17 (Neue Folge). Köln: Institut für Allgemeine Sprachwissenschaft.
- Cuny, Albert. 1907. *Les préverbes dans le Çatapatabrahmaṇa*. Paris, University dissertation.
- Daivids, T. W. and William Stede. 2007 [1921–25]. *Pali-English Dictionary*. Delhi: Motilal Banarsidass.
- Davis, Henry and Lisa Matthewson. 2009. 'Issues in Salish syntax and semantics', *Language and Linguistics Compass* 3 (4), 1097–166.

- de Hoop, Helen and Bhuvana Narasimhan. 2005. 'Differential case marking in Hindi', in M. Amberber and H. de Hoop (eds), *Competition and Variation in Natural Languages: The Case for Case*. Oxford: Elsevier.
- Delbrück, Berthold. 1878. *Die altindische Wortfolge aus dem Çatapathabrâhmaṇa dargestellt*. Halle: Verlag der Buchhandlung des Waisenhauses.
- Delbrück, Berthold. 2009 [1888]. *Altindische Syntax*. Cambridge: Cambridge University Press.
- Delbrück, Berthold. 1900. *Vergleichende Syntax der Indogermanischen Sprachen*. Strassburg: Trübner.
- De Mulder, Walter and Anne Vanderheyden. 2002. 'The evolution of French *sur*: Toward a diachronic Cognitive Semantics', in Hubert Cuyckens and Günter Radden (eds), *Perspectives on Prepositions*. Tübingen: Niemeyer, 177–93.
- Detges, Ulrich. 2009. 'How useful is case morphology? The loss of the Old French two-case system within a theory of Preferred Argument Structure', in Jóhanna Barðdal and Shobhana L. Chelliah (eds), *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*. Amsterdam/Philadelphia: Benjamins, 93–120.
- Diewald, Gabriele and Elena Smirnova. 2010. 'Introduction: Paradigmaticity and obligatoriness of grammatical categories', *Acta Linguistica Hafniensia* 42 (1), (Special issue on *Paradigmaticity and Obligatoriness*), 1–10.
- Diewald, Gabriele and Elena Smirnova. 2012. '"Paradigmatic integration". The fourth stage in an expanded grammaticalization scenario', in Kristin Davidse, Bert Cornillie, H. Cuyckens, and Torsten Leuschner (eds), *Grammaticalization and Language Change. New Reflections*. Amsterdam/Philadelphia: Benjamins, 111–33.
- Dryer, Matthew. 1992. 'The Greenbergian word order correlations', *Language* 68 (1), 81–138.
- Du Bois, John W. 1987. 'The discourse basis of ergativity', *Language* 63, 805–55.
- Dunkel, George E. 1992. 'Die Grammatik der Partikeln', in Robert Beekes, Alexander Lubotsky, and Jos Weitenberg (eds), *Rekonstruktion und relative Chronologie*. Innsbruck: Indogermanische Gesellschaft, 153–77.
- Ebert, Karen. 2001. 'Südasiens als Sprachbund', in Martin Haspelmath and Ekkehard König (eds), *Language Typology and Language Universals. An International Handbook* (HSK vol. 20.2). Berlin: de Gruyter, 1529–39.
- Eckardt, Regine. 2006. *Meaning Change in Grammaticalization. An Enquiry into Semantic Reanalysis*. Oxford: Oxford University Press.
- É. Kiss, Katalin 2008. 'Free word order, (non)configurationality, and phases', *Linguistic Inquiry*, 39 (3), 441–75.
- Emeneau, Murray B. 1956. 'India as a Linguistic Area', *Language* 32 (1), 3–16.
- Emeneau, Murray B. 1980. *Language and Linguistic Area* (Essays selected and introduced by Anwar S. Dil). Stanford, CA: Stanford University Press.
- Fahs, Achim. 1989. *Grammatik des Pali*. Leipzig: VEB Verlag.
- Fontana, Josep. 1993. *Phrase structure and the syntax of clitics in the history of Spanish*. University of Pennsylvania dissertation (UMI).
- Gambhir, Vijay. 1981. *Syntactic restrictions and discourse functions of word order in Standard Hindi*. University of Pennsylvania dissertation (UMI).
- Geiger, Wilhelm. 1916. *Pāli. Literatur und Sprache*. Strassburg: Trübner.
- Gillon, Brendan. 1996. 'Word order in classical Sanskrit', *Indian Linguistics* 57 (1), 1–35.

- Gillon, Brendan and Benjamin Shaer. 2005. 'Classical Sanskrit, "wild trees", and the properties of free word order languages', in *Classical Languages and Generative Linguistics*. Katalin É. Kiss (ed.), Amsterdam/Philadelphia: Benjamins, 281–317.
- Giusti, Giuliana and Rossella Iovino. 2011. 'Evidence for a split DP in Latin', *University of Venice Working Papers in Linguistics* 21, 111–29.
- Giusti, Giuliana and Renato Oniga. 2007. 'Core and periphery in the Latin noun phrase', in Gérald Purnelle and Joseph Denooz (eds), *Ordre et cohérence en latin. Communications présentées au 13e colloque international de linguistique latine*. Geneva: Droz, 81–95.
- Golumbia, David. 2004. 'The interpretation of nonconfigurality', *Language & Communication* 24, 1–22.
- Gonda, Jan. 1959. 'On amplified sentences and similar structures in the Veda', in Jan Gonda (ed.), *Four Studies in the Language of the Veda*. 's-Gravenhage: Mouton, 7–70.
- Grassmann, Hermann. 1996 [1875]. *Wörterbuch zum Rig-Veda* (6th revised edn). Wiesbaden: Harrassowitz.
- Grierson, George A. 1903. 'On certain suffixes in the Modern Indo-Aryan vernaculars', *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 38, 473–91.
- Grierson, George A. 1922. 'Spontaneous nasalization in the Indo-Aryan languages', *Journal of the Royal Asiatic Society of Great Britain and Ireland* 3, 381–8.
- Gumperz, John J. and Robert Wilson. 1971. 'Convergence and creolization: A case from the Indo-Aryan/Dravidian border', in Dell Hymes (ed.), *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press, 15–167.
- Gusain, Lakhani. 2004. *Marwari*. München: Lincom.
- Hagège, Claude. 2010. *Adpositions*. Oxford: Oxford University Press.
- Haiman, John. 1994. 'Ritualization and the development of language', in William Pagliuca (ed.), *Perspectives on Grammaticalization*. Amsterdam/Philadelphia: Benjamins, 3–28.
- Hale, Kenneth. 1983. 'Warlpiri and the grammar of non-configurational languages', *Natural Language & Linguistic Theory* 1 (1), 5–47.
- Hale, Kenneth. 1989. 'On nonconfigurational structures', in L. Marácz and P. Muysken (eds), *Configurationality: The Typology of Asymmetries*. Dordrecht: Foris, 293–300.
- Hale, Mark. 1995. Wackernagel's law. Phonology and syntax in the Rigveda. Unpublished manuscript.
- Hale, Mark. 1996. 'Deriving Wackernagel's law: Prosodic and syntactic factors determining clitic placement in the language of the Rigveda', in Aaron L. Halpern and Arnold M. Zwicky (eds), *Approaching Second. Second Position Clitics and Related Phenomena*. Stanford, CA: CSLI, 165–97.
- Hall, T. A. 1997. 'The historical development of retroflex consonants in Indo-Aryan', *Lingua* 102, 203–21.
- Halpern, Aaron L. and Arnold M. Zwicky (eds). 1996. *Approaching Second. Second Position Clitics and Related Phenomena*. Stanford, CA: CSLI.
- Hamp, Eric P. 1996. 'On the Indo-European origins of the retroflexes in Sanskrit', *Journal of the American Oriental Society* 116 (4), 719–23.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press.

- Harris, Martin B. 1978. 'The inter-relationship between phonological and grammatical change', in Jacek Fisiak (ed.), *Recent Developments in Historical Phonology*. The Hague/Paris/New York: de Gruyter, 159–72.
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. München/Newcastle: Lincom Europa.
- Heath, Jeffrey. 1986. 'Syntactic and lexical aspects of nonconfigurationality in Nunggubuyu (Australia)', *Natural Language & Linguistic Theory* 4 (3), 375–408.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hettrich, Heinrich. 1988. *Untersuchungen zur Hypotaxe im Vedischen*. Berlin/New York: de Gruyter.
- Hettrich, Heinrich. 1991. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. I: *ádhi*', *Münchener Studien zur Sprachwissenschaft* 52, 27–76.
- Hettrich, Heinrich. 1993. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. II: *antár*', *Münchener Studien zur Sprachwissenschaft* 54, 147–76.
- Hettrich, Heinrich. 2002. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. III: *pári*', in Matthias Fritz and Susanne Zeilfelder (eds), *Novalis Indogermanica. Festschrift für Günter Neumann zum 80. Geburtstag*. Leykam: Graz, 215–42.
- Hettrich, Heinrich. 2007. *Materialien zu einer Kasussyntax des Ṛgveda*. Würzburg: Universität zu Würzburg, Institut für Altertumswissenschaften. Online at <http://www.vergl-sprachwissenschaft.phil.uni-wuerzburg.de/fileadmin/04080400/Materialien.pdf>, last accessed 26 October 2015.
- Hettrich, Heinrich, Antje Casaretto, and Carolin Schneider. 2010. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. IV: Allgemeines (H. Hettrich); II. *úpa* (A. Casaretto); III. *áva* (C. Schneider)', *Münchener Studien zur Sprachwissenschaft* 64, 17–130.
- Hewson, John. 2006. 'Typological evolution in Indo-European: From paradigm to phrase', in John Hewson and Vít Bubeník. *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: Benjamins, 1–27.
- Hewson, John and Vít Bubeník. 2006. *From Case to Adposition. The Development of Configurational Syntax in Indo-European Languages*. Amsterdam/Philadelphia: Benjamins.
- Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer.
- Himmelmann, Nikolaus. 1998. 'Regularity in irregularity: Article use in adpositional phrases', *Linguistic Typology* 2, 315–53.
- Himmelmann, Nikolaus. 2005. 'Gram, construction, and class formation', in Clemens Knobloch and Burkhard Schaefer (eds), *Wortarten und Grammatikalisierung: Perspektiven in System und Erwerb*. Berlin/New York: de Gruyter, 79–93.
- Hinds, J. 1974. *On the Status of the VP Node in Japanese*. Bloomington: Indiana University Linguistics Club.
- Hock, Hans Henrich. 1975. 'Substratum influence on (Rig-Vedic) Sanskrit?', *Studies in Linguistic Sciences* 5, 76–125.
- Hock, Hans Henrich. 1982. 'Clitic verbs in PIE or discourse-based verb fronting? Sanskrit *sá hovāca gárgyaḥ* and congeners in Avestan and Homeric Greek', *Studies in the Linguistic Sciences* 12 (2), 1–38.

- Hock, Hans Henrich. 1984. '(Pre-)Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence', *Studies in the Linguistic Sciences* 14 (1), 89–108.
- Hock, Hans Henrich. 1986. "'P-oriented" constructions in Sanskrit', in Bhadriraju Krishnamurti, Colin P. Masica, and Anjani K. Sinha (eds), *South Asian Languages: Structure, Convergence and Diglossia*. Delhi: Motilal Banarsidass, 15–26.
- Hock, Hans Henrich. 1991 (ed.), *Studies in Sanskrit Syntax. A Volume in Honor of the Centennial of Speijer's Sanskrit Syntax (1886–1986)*. Delhi: Motilal Banarsidass.
- Hock, Hans Henrich. 1992. 'What's a nice word like you doing in a place like this? Syntax vs. phonological form', *Studies in the Linguistic Sciences* 22, 39–87.
- Hock, Hans Henrich. 1994. 'Discourse linkage in Sanskrit narratives with special emphasis on the story of Nala', *SALA* XV, 117–35.
- Hock, Hans Henrich. 1996a. '(Pre-)Rgvedic convergence between Indo-Aryan (Sanskrit) and Dravidian? A survey of the issues and controversies', in Jan E. M. Houben (ed.), *Ideology and Status of Sanskrit. Contributions to the History of the Sanskrit Language*. Leiden/New York/Köln: Brill.
- Hock, Hans Henrich. 1996b. 'Who's on first? Toward a prosodic account of P2 clitics', in Aaron Halpern and Arnold Zwicky (eds), 1996. *Approaching Second. Second Position Clitics and Related Phenomena*. Stanford, CA: CSLI, 199–270.
- Hock, Hans Henrich. 1997. 'Chronology or genre? Problems in Vedic syntax', in Michael Witzel (ed.), *Inside the Texts—Beyond the Texts: New Approaches to the Study of the Vedas*. Cambridge, MA: Harvard University Press, 103–26.
- Hock, Hans Henrich. 2006. 'Reflexivization in the Rig-Veda (and beyond)', in H. Hettrich and B. Tikkanen (eds), *Themes and Tasks in Old and Middle Indo-Aryan Linguistics*. Delhi: Motilal Banarsidass, 19–44.
- Hoernle, Rudolf. A. F. 1880. *A Comparative Grammar of the Gaudian Languages with Special Reference to the Eastern Hindi Accompanied by a Language-map and a Table of Alphabets*. London: Trübner.
- Hofmann, J. B. 1965. *Lateinische Syntax und Stilistik* (neubearbeitet von Anton Szantyr). München: C. H. Beck
- Hook, Peter Edwin. 1988. 'Paradigmaticization: A case study from South Asia', in *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 293–303.
- Hopper, Paul J. 1991. 'On some principles of grammaticalization', in Elizabeth C. Traugott and Bernd Heine (ed.), *Approaches to Grammaticalization* (Vol. I). Amsterdam/Philadelphia: Benjamins, 17–35.
- Hopper, Paul J. 1996. 'Some recent trends in grammaticalization', *Annual Review of Anthropology* 25, 217–36.
- Hopper, Paul J. and Elizabeth C. Traugott. 2003. *Grammaticalization* (2nd edn). Cambridge: Cambridge University Press.
- Iovino, Rossella. 2011. 'Word order in Latin nominal expressions. The syntax of demonstratives', in Renato Oniga, Iovino Rossella, and Giuliana Giusti (eds), *Formal Linguistics and the Teaching of Latin. Theoretical and Applied Perspectives in Comparative Grammars*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 51–64.



- Jacobi, Hermann. 1967 [1886]. *Ausgewählte Erzählungen in Māhārāṣṭrī*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Jelinek, Eloise. 1984. 'Empty categories, case, and configurationality', *Natural Language & Linguistic Theory* 2 (1), 39–76.
- Kachru, Yamuna. 2006. *Hindi*. Amsterdam/Philadelphia: Benjamins.
- Kahr, Joan Casper. 1975. 'Adpositions and locationals: Typology and diachronic development', *Working Papers on Language Universals, Stanford* 19, 21–54.
- Kahr, Joan Casper. 1976. 'The renewal of case morphology: Sources and constraints', *Working Papers on Language Universals, Stanford* 20, 107–51.
- Kānta, Sūrya. 1937. *Index verborum to Tulasī Dāsa's Rāmāyaṇa*. Lahore: University of Panjab.
- Kaufman, Daniel Aaron. 2010. *The morphosyntax of Tagalog clitics: A typologically-driven approach*. Cornell University dissertation. Online at <https://ecommons.cornell.edu/handle/1813/17165>, last accessed 26 October 2015.
- Kellogg, S. H. 1972 [1875]. *A Grammar of the Hindī Language, in which are treated the High Hindī, Braj, and the Eastern Hindī of the Ramāyan of Tulsī Dās, etc.* New Delhi: Oriental Books Reprint Corporation.
- Keydana, Götz. 2013. *Infinitive im Rigveda: Formen, Funktion, Diachronie*. Leiden: Brill.
- Keydana, Götz and Silvia Luraghi. 2012. 'Definite referential null objects in Vedic Sanskrit and Ancient Greek', *Acta Linguistica Hafniensia* 44 (2), 116–28.
- Khan, Tafseer Ahmed. 2009. *Spatial Expressions and Case in South Asian Languages*. Konstanz: Fachbereich Sprachwissenschaft (dissertation). Online at <http://kops.ub.uni-konstanz.de/handle/urn:nbn:de:bsz:352-125089>, last accessed 29/08/2013.
- Klein, Jared S. 1991. 'Syntactic and discourse correlates of verb-initial sentences in the Rigveda', in Hans Henrich Hock (ed.), *Studies in Sanskrit Syntax. A Volume in Honor of the Centennial of Speijer's Sanskrit Syntax (1886–1986)*. Delhi: Motilal Banarsidass, 123–43.
- Krisch, Thomas. 2005. 'Preliminaries to the study of adjectival syntax in Proto-Indo-European', in Günther Schweiger (ed.), *Indogermanica. Festschrift Gert Klingenschmitt*. Tübingen: Schweiger VWT-Verlag, 303–19.
- Kuiper, F. B. J. 1967. 'The genesis of a linguistic area', *Indo-Iranian Journal* 10 (2–3), 81–102.
- Kulikov, Leonid. 2007. 'The reflexive pronouns in Vedic: A diachronic and typological perspective', *Lingua* 117, 1412–33.
- Kuryłowicz, Jerzy. 1958. *L'Accentuation des langues Indo-Européennes* (2nd edn). Wrocław/Kraków: Zakład Narodowy Wydawnictwo.
- Kuryłowicz, Jerzy. 1964. *Inflectional Categories of Indo-European*. Heidelberg: Winter.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
- Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Ledgeway, Adam. 2012. *From Latin to Romance*. Oxford: Oxford University Press.
- Lehmann, Christian. 1991. 'The Latin nominal group in typological perspective', in Robert Coleman (ed.), *New Studies in Latin Linguistics*. Amsterdam/Philadelphia: Benjamins.
- Lehmann, Christian. 2002 [1982]. *Thoughts on Grammaticalization* (2nd revised edn). (Arbeitspapiere Seminar für Sprachwissenschaft 9). Erfurt: Seminar für Sprachwissenschaft der Universität Erfurt.

- LeSourd, Phillip S. 2006. 'Problems for the pronominal argument hypothesis in Maliseet-Passamaquoddy', *Language* 82, 486–514.
- Levin, Beth and Malka Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Levinson, Stephen C. and S. Meira, and The Language & Cognition Group. 2003. "Natural concepts" in the spatial topological domain', *Language* 79 (3), 485–516.
- Lundskær-Nielsen, Tom. 1993. *Prepositions in Old and Middle English. A Study of Prepositional Syntax and the Semantics of At, In and On in some Old and Middle English Texts*. Odense: Odense University Press.
- Luraghi, Silvia. 2010. 'The rise (and possible downfall) of configurationality', in Silvia Luraghi and Vít Bubeník (eds), *The Continuum Companion to Historical Linguistics*. London/New York: Continuum, 212–29.
- Marácz, László and Pieter Muysken (eds). 1989. *Configurationality. The Typology of Asymmetries*. Dordrecht: Foris.
- Marouzeau, J. 1922. *L'ordre des mots dans la phrase Latine*. Paris: Champion.
- Masica, Colin P. 1976. *Defining a Linguistic Area. South Asia*. Chicago/London: University of Chicago Press.
- Masica, Colin P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Matthews, P. H. 1981. *Syntax*. Cambridge: Cambridge University Press.
- Matthews, P. H. 2007. *Syntactic Relations. A Critical Survey*. Cambridge: Cambridge University Press.
- Mayrhofer, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen* (Vol. I). Heidelberg: Winter.
- Mayrhofer, Manfred. 1996. *Etymologisches Wörterbuch des Altindoarischen* (Vol. II). Heidelberg: Winter.
- McGregor, R. S. 1967. 'The rise of standard Hindi and Early Hindi prose fiction', *Journal of the Royal Asiatic Society* 99 (2), 114–32.
- McGregor, R. S. 2001. 'On the evolution of Hindi as a language of literature', *South Asia Research* 21 (2), 203–17.
- McGregor, William B. 1997. 'Functions of noun phrase discontinuity in Gooniyandi', *Functions of Language* 4, 83–114.
- Meillet, Antoine. 1975 [1912]. 'L'évolution des formes grammaticales', in Antoine Meillet, *Linguistique historique et linguistique générale*. Paris: Champion, 130–48.
- Meillet, Antoine. 1975 [1915/1916]. 'Le renouvellement des conjonctions', in Antoine Meillet, *Linguistique historique et linguistique générale*. Paris: Champion, 159–74.
- Meillet, Antoine and J. Vendryes. 1927. *Traité de grammaire comparée des langues classiques*. Paris: Champion.
- Mithun, Marianne. 1987. 'Is basic word order universal?', in Russel S. Tomlin (ed.), *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins, 281–328.
- Mohanan, Tara. 1994. *Argument Structure in Hindi*. Stanford, CA: CSLI.
- Monier-Williams, Monier. 1899. *Sanskrit-English Dictionary* (revised version from 2008). Online at <http://www.sanskrit-lexicon.uni-koeln.de>, last accessed 25 October 2015.

- Montaut, Annie. 2006. 'The evolution of the tense–aspect system in Hindi/Urdu: The status of the ergative alignment', in Miriam Butt and Tracy Holloway King (eds), *Proceedings of the LFG06 Conference*. Online at [http://ceias.ehess.fr/docannexe/file/559/membres\\_montaut\\_diachrony\\_ergativelfg06.pdf](http://ceias.ehess.fr/docannexe/file/559/membres_montaut_diachrony_ergativelfg06.pdf), last accessed 26 October 2015.
- Montaut, Annie. 2009. 'Ergative and pre-ergative patterns in Indo-Aryan as predications of localization: A diachronic view of past and future systems', in A. Faithi (ed.), *Language Vitality in South Asia*. Aligarh: Aligarh Muslim University, 295–325.
- Müller, Antje. 2013. *Spatiale Bedeutungen deutscher Präpositionen. Bedeutungsdifferenzierung und Annotation*. Bochum: Sprachwissenschaftliches Institut, University of Bochum (dissertation).
- Murasugi, Kumiko. 2014. 'Noun incorporation, nonconfigurationality, and polysynthesis', in Andrew Carnie, Yosuke Sato, and Daniel Siddiqi (eds), *The Routledge Handbook of Syntax*. London: Routledge, 283–304.
- Mylius, Klaus. 2003. *Geschichte der altindischen Literatur. Die 3000jährige Entwicklung der religiös-philosophischen, belletristischen und wissenschaftlichen Literatur Indiens von den Veden bis zur Etablierung des Islam* (2nd revised and extended edn). Wiesbaden: Harrassowitz.
- Nevis, Joel Ashmore. 1988. *Finnish Particle Clitics and General Clitic Theory*. New York/London: Garland.
- Noël, Dirk. 2007. 'Diachronic construction grammar and grammaticalization theory', *Functions of Language* 14 (2), 177–202.
- Nordlinger, Rachel. 1998. *Constructive Case. Evidence from Australian Languages*. Stanford, CA: CSLI.
- Nordlinger, Rachel. 2014. 'Constituency and grammatical relations in Australian languages', in Harold Koch and Rachel Nordlinger (eds), *The Languages and Linguistics of Australia*. Berlin/New York: de Gruyter, 215–62.
- Oberlies, Thomas. 1998. *Historische Grammatik des Hindi: Die Genese seines morphologischen Systems aus dem Mittel- und Altindischen*. Freiburg: Reinbek.
- Oberlies, Thomas. 2001. *Pāli. A Grammar of the Language of the Theravāda Tipiṭaka*. Berlin/New York: de Gruyter.
- Oldenberg, H. 1917. *Zur Geschichte der altindischen Prosa*. Berlin: Weidmann.
- Osthoff, Hermann. 1879. 'Kleine beiträge zur declinationslehre der indogermanischen sprachen II', in Hermann Osthoff and Karl Brugmann, *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen II*. Leipzig: Hirzel, 1–147.
- Pokorny, Julius. 1959. *Indogermanisches Etymologisches Wörterbuch* (Vol. I). Bern: A. Francke AG.
- Raza, Ghulam and Tafseer Ahmed. 2011. 'Argument scrambling within Urdu NPs', in Miriam Butt and Tracy Holloway King (eds), *Proceedings of the LFG11 Conference*. Online at <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/16/papers/lfg11razaahmed.pdf>, last accessed 26 October 2015.
- Reichelt, Hans. 1978 [1909]. *Awestisches Elementarbuch*. Heidelberg: Carl Winter Universitätsverlag.

- Reinöhl, Uta. 2009. *Zum Begriff der Renovation im Rahmen der Grammatikalisierungstheorie*. Münster: Institut für Allgemeine Sprachwissenschaft. (Magisterarbeit (MA thesis), unpublished).
- Renou, Louis. 1933. 'La séparation du préverbe et du verbe en Védique', *Bulletin de la Société de Linguistique de Paris* 34, 49–96.
- Renou, Louis. 1956. 'Préposition et préverbe dans le Ṛgveda', in Louis Renou, *Études védiques et pāṇinéennes* (Vol. II). Paris: E. de Boccard, 113–27.
- Rijkhoff, J. 2002. *The Noun Phrase*. Oxford: Oxford University Press.
- Robbeets, Martine and Walter Bisang (eds). 2014. *Paradigm Change. In the Transeurasian Languages and beyond*. Amsterdam/Philadelphia: Benjamins.
- Rögnvaldsson, Eiríkur. 1995. 'Old Icelandic: A non-configurational language?', *NOWELE* 26, 3–29.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.
- Rubba, Jo. 1994. 'Grammaticization as semantic change. A case study of preposition development', in William Pagliuca (ed.), *Perspectives on Grammaticalization*. Amsterdam/Philadelphia: Benjamins, 81–101.
- Saint-Dizier, Patrick (ed.). 2006. *Syntax and Semantics of Prepositions*. Dordrecht: Springer.
- Saksena, B. 1971. *Evolution of Avadhi (a Branch of Hindi)* (2nd edn). Delhi: Motilal Banarsidass.
- Sandahl, Stella. 2000. *A Hindi Reference Grammar*. Leuven: Peeters.
- Schäufele, Steven William. 1991a. 'A note on the term "scrambling"', *Natural Language & Linguistic Theory* 9 (2), 365–8.
- Schäufele, Steven William. 1991b. *Free word-order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory*. Urbana-Champaign: University of Illinois dissertation.
- Schäufele, Steven William. 1996. 'Now that we're all here, where do we sit? Phonological Ordering in the Vedic clause-initial string', in Aaron L. Halpern and Arnold M. Zwicky (eds), *Approaching Second. Second Position Clitics and Related Phenomena*. Stanford, CA: CSLI, 447–75.
- Schneider, Carolin. 2009 [2010]. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. V: *ní*', *Historische Sprachforschung* 122, 118–69.
- Schneider, Carolin. 2010a. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. VII: *nís*', *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 7, 149–93.
- Schneider, Carolin. 2010b. 'Lokalpartikel im Ṛgveda (Folge 13): *ápa*', *Studien zur Indologie und Iranistik* 27, 1–35.
- Schneider, Carolin. 2011. 'Syntax und Wortarten der Lokalpartikeln des Ṛgveda. X: *úd*', *Münchener Studien zur Sprachwissenschaft* 65, 185–241.
- Schneider, Carolin. 2012. 'Local particles in the Ṛgveda. XVII: *pára*', *Münchener Studien zur Sprachwissenschaft* 66, 221–46.
- Schultze-Berndt, Eva and Candide Simard. 2012. 'Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure', *Linguistics* 50 (5), 1015–58.

- Schwarzschild, L. A. 1954. 'The possessive adjectives of Late Prakrit', *Journal of the Royal Asiatic Society of Great Britain and Ireland* 3-4, 127-36.
- Shapiro, Michael. 2007. 'Hindi', in George Cardona and Dhanesh Jain (eds), *The Indo-Aryan Languages*. London/New York: Routledge, 250-85.
- Sharma, Devyani. 1999. 'Nominal clitics and constructive morphology in Hindi', in Miriam Butt and Tracy Holloway King (eds), *Proceedings of the LFG99 Conference*. Stanford, CA: CSLI. (Online at <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/4/lfg99sharma.pdf>, last accessed 26 October 2015.
- Silverstein, Michael. 1976. 'Hierarchy of features and ergativity', in R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. New Jersey: Humanities Press, 112-71.
- Simpson, Andrew, Arunima Choudhary, and Mythili Menon. 2013. 'Argument ellipsis and the licensing of covert nominals in Bangla, Hindi and Malayalam', *Lingua* 134, 103-28.
- Simpson, Jane. 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Dordrecht: Kluwer.
- Simpson, Jane and Ilana Mushin. 2008. 'Clause-initial position in four Australian languages', in Ilana Mushin and Brett Baker (eds), *Discourse and Grammar in Australian Languages*. Amsterdam/Philadelphia: Benjamins, 25-57.
- Singh, Amar Bahadur. 1977. 'On time and place postpositions in Hindi', *Indian Linguistics* 38, 12-20.
- Southworth, Franklin C. and M. L. Apte (eds). 1974. *Contact and Convergence in South Asian Languages*. (Special publication of the *International Journal of Dravidian Linguistics*). Trivandrum.
- Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Speijer, J. S. 1968 [1886]. *Sanskrit Syntax*. Leyden: Brill.
- Spencer, Andrew. 2005. 'Case in Hindi', in Miriam Butt and Tracy Holloway King (eds), *Proceedings of the LFG05 Conference*. Online at <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/10/lfg05spencer.pdf>, last accessed 26 October 2015.
- Staal, J. F. 1967. *Word Order in Sanskrit and Universal Grammar*. Dordrecht: D. Reidel.
- Starosta, Stanley. 1985. 'Relator nouns as a source of case inflection', *Oceanic Linguistics Special Publications* 20 (For Gordon H. Fairbanks), 111-33.
- Strnad, Jaroslav. 2013. *Morphology and Syntax of Old Hindi*. Leiden/Boston: Brill.
- Stucky, Susan. 1981. 'Free word order languages, free constituent order languages, and the gray area in between', *Proceedings of NELS* 11, 364-76.
- Svorou, Soteria. 1988. *The experiential basis of the grammar of space. Evidence from the languages of the world*. Buffalo: SUNY dissertation.
- Sweet, Henry (ed.) 1871. *King Alfred's West-Saxon Version of Gregory's Pastoral Care*. London: Trübner.
- Sweetser, Eve E. 1988. 'Grammaticalization and semantic bleaching', *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 389-405.
- Talmy, Leonard. 1972. *Semantic structures in English and Atsugewi*. Berkeley: University of California PhD dissertation.
- Taylor, Ann. 1990. *Clitics and configurationality in Ancient Greek*. University of Pennsylvania dissertation.
- Tessitori, L. P. 1913. 'On the origin of the dative and genitive postpositions in Gujarati and Marwari', *Journal of the Royal Asiatic Society of Great Britain and Ireland*, 553-67.

- Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley/Los Angeles/Oxford: University of California Press.
- Thommen, Eduard. 1905. 'Die Wortstellung im nachvedischen Altindischen und im Mittelindischen', *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 38 (4), 504–63.
- Trumpp, E. 1872. *Grammar of the Sindhi Language*. London: Trübner; Leipzig: Brockhaus.
- Turner, R. L. 1916. 'The Indo-Germanic accent in Marathi', *Journal of the Royal Asiatic Society of Great Britain and Ireland*, 203–51.
- Turner, R. L. 1936. 'Ā-kṣeti and Pali acchati in Modern Indo-Aryan', *Bulletin of the School of Oriental Studies* 8 (2–3), 795–812.
- Turner, R. L. 1960. *Some Problems of Sound Change in Indo-Aryan*. Poona [Pune]: University of Poona [Pune].
- Turner, R. L. 1966. *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.
- Ultan, Russell. 1978. 'Towards a typology of substantival possession', in Joseph Greenberg (ed.), *Universals of Human Language 4: Syntax*. Stanford, CA: Stanford University Press, 12–49.
- van de Velde, Freek. 2009. 'The emergence of modification patterns in the Dutch noun phrase', *Linguistics* 47 (4), 1021–49.
- van de Velde, Freek. 2010. 'The emergence of the determiner in the Dutch NP', *Linguistics* 48 (2), 263–99.
- Verbeke, Saartje and Ludovic De Cuypere. 2009. 'The rise of ergativity in Hindi. Assessing the role of Grammaticalization', *Folia Linguistica Historica* 30, 1–24.
- Vincent, Nigel. 1999. 'The evolution of C-Structure: prepositions and PPs from Indo-European to Romance', *Linguistics* 37 (6), 1111–53.
- Viśva-Bandhu Śāstrī. 1935–65. *Vaidika-Padānukrama-Koṣa: A Vedic Word-Concordance*. Hoshiarpur: Vishveshvaranand Vedic Research Institute.
- Viti, Carlotta. 2009. 'A quantitative analysis of the OSV word order in Vedic', in Elisabeth Rieken and Paul Widmer (eds), *Pragmatische Kategorien. Form, Funktion und Diachronie*. Wiesbaden: Reichert, 307–22.
- von der Gabelentz, Georg. 1901. *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse* (2nd edn). Leipzig: Tauchnitz.
- von Hinüber, Oskar. 2001. *Das ältere Mittelindisch im Überblick* (Vol. II). Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Wackernagel, Jacob. 1892. 'Über ein Gesetz der indogermanischen Wortstellung', *Indogermanische Forschungen* 1, 333–436.
- Webelhuth, Gert. 1984–85. 'German is configurational', *The Linguistic Review* 4, 203–46.
- Weber, Albrecht. 1877. 'Review of S. H. Kellogg, *Grammar of the Hindi Language*', *Jenaer Literatur-Zeitung* 33, 517–19.
- Weil, Henri. 1844. *De l'ordre des mots dans les langues anciennes comparées aux langues modernes*. Paris: Franck.
- West, Martin L. 2011. *Old Avestan Syntax and Stylistics*. Berlin/Boston: de Gruyter.
- Whitney, William Dwight. 1855–56. 'Contributions from the Atharvaveda to the theory of Sanskrit verbal accent', *Journal of the American Oriental Society* 5, 385–419.

- Whitney, William Dwight. 1869–70. ‘On the nature and designation of the accent in Sanskrit’, *Transactions of the American Philological Association* 1, 20–45.
- Whitney, William Dwight. 1983 [1879]. *Sanskrit Grammar* (5th edn). Delhi: Motilal Banarsidass.
- Wilhelm, Christopher. 2001. *The origin and development of adpositions and apositional phrases in the Indo-European languages*. Los Angeles: University of California dissertation.
- Wodtko, Dagmar S., Britta Irslinger, and Carolin Schneider. 2008. *Nomina im Indogermanischen Lexikon*. Heidelberg: Universitätsverlag Winter.
- Yamaguchi, Kazuyuki. 2004. *A Typological, Historical, and Functional Study of Adpositions in the Languages of the World*. Albuquerque, New Mexico (UMI Microform).
- Zwicky, Arnold. 1985. ‘On clitics and particles’, *Language* 61 (2), 283–305.





# Language index

Ancient Greek 38, 45–6, 56, 66–7, 132–3  
Apabhramsha 2, 10, 11–12, 14, 15, 16–17,  
18, 54, 57, 60, 61, 64, 85, 87, 92–3, 111,  
112, 122–3, 127–30, 133, 135–9, 142,  
145, 149–70, 172, 183, 186, 188–9, 203

Ashokan inscription 133

Classical Sanskrit, *see* Sanskrit

Dravidian 68, 121–2, 131

English 2, 3, 4, 7, 21–2, 25, 27, 32, 36, 54, 56,  
67–8, 92–3, 107, 110–13, 121, 144, 147,  
155, 161, 165, 166, 171, 177, 181, 182, 196

French 8, 20, 23, 49, 67–8, 90, 110, 111, 114,  
144, 196, *see also* Romance

German 22, 31, 56, 90, 93, 100, 106, 110–13,  
144, 152, 195–6

Gujarati 60

Haryani 60–3, 193

Hindi 1–7, 10–11, 14–15, 17, 18, 19, 21,  
32–3, 37, 39, 40, 42–5, 47, 51, 53–64, 65,  
68, 69, 83, 85, 94, 101, 107, 111–13,  
114–15, 116, 121–4, 128, 129, 131,  
163–70, 172, 175–9, 181, 182, 183–4,  
185–7, 190–9, 201–4  
early Hindi 2, 14, 62, 131, 182, 194,  
*see also* Old Awadhi

Indo-Aryan 1–12, 21, 23, 28, 47–8, 50, 51,  
53, 65–6, 67–70, 80–4, 85, 87–8, 90, 94,  
110, 113–17, 120–31, 132, 164, 170, 176,  
184, 187, 196, 201–4

Old Indo-Aryan 2, 6, 9, 11, 12, 13, 14, 17,  
18, 36, 40, 54, 55, 56, 57, 64, 70, 71, 78,  
83, 85, 86, 96, 97, 100, 103, 104, 111, 122,

124, 133, 134, 143, 175, 186, 187, 188,  
201, 202, *see also* Sanskrit

Middle Indo-Aryan 1, 9–18, 53, 55, 56,  
57, 64, 66, 70, 71, 76, 79, 83, 85, 86,  
87–8, 90, 94, 100, 103–4, 116, 120–5,  
131–3, 143, 144, 148, 151, 160, 167,  
170, 172, 181, 183, 185, 186, 201–3,  
*see also* Apabhramsha; Ashokan  
inscription; Pali; Prakrit

New Indo-Aryan 1–19, 50–1, 53–64, 65,  
68, 69, 70, 83, 86, 92, 100, 111, 119,  
120–31, 132, 144, 147, 151, 161, 163,  
164, 169, 181, 193, 196, 201, *see also*  
Gujarati; Haryani; Hindi; Kashmiri; Old  
Awadhi, Rajasthani; Urdu

Indo-European 2, 3, 5, 8, 9, 18, 46–8, 54,  
56, 65–83, 86, 87, 89–90, 93–4, 110–11,  
113–17, 119, 132, 171, 185, 196

Iranian 80, 82–3, 94, 121

Jaminjung 30–2, 38, 39

Jiwarli 3, 26

Kashmiri 121, 133

Latin 5, 8, 20, 31, 32, 45–9, 54, 56, 67, 68,  
87, 89, 90, 92, 111–12, 114, 116, 132,  
158–60, 171, 203

Mahārāṣṭrī 76, 133

Nepali 62

Nunggubuyu 25, 31–2, 41, 87, 139

Old Awadhi 2, 10, 11, 15, 16, 54, 56, 59,  
62, 86, 131, 161, 166, 172, 189, 194,  
*see also* Hindi, early Hindi

- Pali 2, 10, 11, 12, 13, 14, 15, 16, 17, 37, 54,  
55, 58, 62, 70, 76, 78, 79, 85, 86, 89,  
90–4, 95–111, 112, 121, 122–30, 133,  
136, 137, 142, 148, 150, 152, 153,  
155–8, 159–62, 205, *see also* Indo-Aryan,  
Middle Indo-Aryan
- Persian, *see* Iranian
- Prakrits 10–11, 54, 60, 67, 100, 121, 131, 133,  
*see also* Mahārāṣṭrī
- Proto-Indo-European 18, 46, 54, 65, 67, 68,  
85, 110, 202
- Rajasthani 60–3, 193
- Romance 5, 47, 171, 196, *see also* French
- Sanskrit 10, 15, 16, 17, 32, 53, 54, 56, 58, 60,  
68, 94, 104, 111, 112, 121, 131, 133,  
*see also* Indo-Aryan, Old Indo-Aryan
- Classical Sanskrit 10, 11, 13, 74,  
76, 91, 97, 100, 101, 104, 112,  
120, 133
- Vedic Sanskrit xv, 1–6, 10–19, 24, 28,  
29, 31–42, 43, 44, 46, 51, 66, 67, 68, 69,  
70–6, 79, 80, 81, 82, 83, 86, 88–90, 91,  
92, 93, 94, 95, 96, 97, 98, 99–102, 103,  
104, 110, 113, 114, 116, 119, 120, 122,  
123, 125, 126, 127, 129, 130, 131, 135,  
136, 137, 138, 139, 142, 143, 145, 147,  
148, 149–55, 156, 158, 161, 163, 165,  
202, 205
- Urdu 1, 58–60, 64, 129, *see also* Hindi
- Vedic Sanskrit, *see* Sanskrit
- Warlpiri 23–9, 32, 42, 51

# Subject index

- accentuation 28, 30, 71, 72, 80–3, 125–6,  
133, 134, 135, 137
- adjunct 26, 28, 43, 179, 194, 198,  
*see also* apposition
- adposition 2, 3, 6, 7, 8, 20, 32, 40, 46, 47, 49,  
54, 65–83, 90, 93, 120, 131, 147, 163,  
164, 166, 172, 187, 190, 194, 196, 198,  
202, 203, *see also* postposition
- preposition 18, 46, 48, 49, 68, 73, 74, 80,  
83, 94, 95, 96, 107, 110, 115, 116, 119,  
121, 132, 137, 144, 145, 158, 171,  
196, 202
- adpositional phrase 5, 8, 9, 23, 48, 65, 70,  
83, 87, 89, 93, 110, 113, 114, 115, 116,  
117, 120, 171, 177, 187, 202,  
*see also* postpositional phrase
- prepositional phrase 45, 46, 47, 196
- adverb 2, 3, 7, 8, 11, 16, 40, 44, 49, 56, 57, 64,  
65, 66–83, 86–117, 175, 181, 185, 186,  
194, 199, 202, *see also* local particle
- afterthought 30, 38–40
- alignment 36, 193
- analogical change 100–10, 175, 185
- analytic expression 8, 47, 96, 114
- animacy 58, 148, 151, 158–60, 162, 194–5
- apposition 3, 69, 89, 153, 162, 185, 202
- article 1, 2, 5, 31–2, 40, 45–6, 48, 113,  
144, 145, 171–2, 177, 182–4, 187, 190,  
196, 198
- asymmetrical group, *see* group
- autosemantic 20–2, 46, 50, 147, 148,  
166–73, 181
- auxiliary complex 26, *see also* Wackernagel  
position
- binding 25, 27
- body part term 21, 22, 30, 54, 58, 64, 86,  
94, 171
- case, *see* direct case; ergativity; genitive;  
local case; locative; oblique case
- functional 1, 3, 8, 11, 27, 46, 58, 78, 83, 85,  
87, 95, 96, 99, 114, 159, 160, 162, 175,  
177, 178, 181, 183, 184, 194–6, 198,  
199, 204
- morphological 3, 6, 8, 9, 12, 14, 26, 41, 42,  
44, 46, 47, 48, 54, 56, 58, 63, 66–70, 71,  
73, 80, 86, 87–90, 93, 95, 96, 99–110,  
112, 113, 114, 115, 148, 153, 159, 160,  
161, 162, 167, 175, 178, 181, 184, 185,  
186, 187, 191, 194, 195, 196, 198, 202
- clitics 11, 20, 23, 28, 38, 81, 82, 123–8, 132–7,  
141, 145, 178, 189, 203
- clitic insertion 37, 123, 127
- cliticization 11, 15, 23, 80–2, 137, 178, 189
- P2 clitic 28–9, 37, 38, 123, 124, 126–7,  
132–4, 136, 141
- configurationality 1–7, 23–51, 70, 86, 90,  
117, 123, 132, 201, *see also*  
non-configurationality; phrase structure
- low-level configurationality 1, 2, 4, 14, 19,  
27, 28, 33, 46, 51, 65, 201
- strong configurationality 27
- context expansion 20, *see also* host class
- continuity 28, 119, 135, *see also* juxtaposition
- contrastive focus 30, 34, 39
- control 25, 36–7
- convergence 121, 122, 175, 176, 182, 198
- co-referentiality 3, 4, 29, 30, 37, 40–1, 46, 50,  
113, 140, 202
- demonstrative 3, 40–1, 44, 46, 47, 113, 119,  
126, 133, 136, 137, 139–45, 182, 183,  
190, 191, 203
- dependency relation 29, 94, 114, 115, 117,  
124, 153, 171, 203
- dependent-marking 3, 27, 47

- direct case 8, 44, 191, 193, 194–6, 197, 204
- discontinuity 19, 23–7, 28, 29–32, 37–40, 43,  
51, 119, 123–31, 132–5, 136, 138, 139,  
141, 143, 144, 145, 169, 188, *see also*  
continuity; word order
- apparent discontinuity 28, 29–31, 37, 38,  
39, 40, 51, 124
- true discontinuity 19, 28, 29–32, 37–40, 51
- discrete change 4, 5, 20, 21, 50, *see also*  
gradual change
- distributional similarity 89, 175, 177–9, 185,  
190–3, 196–7
- dominance 24
- dummy pronoun 32, 36, *see also* expletive;  
null anaphora
- emphasis 151, 201–2, *see also* contrastive  
focus
- epiphenomenality 22, 23, 29, 176, 183,  
196–7, 198, 204
- epiphenomenality scenario 176, 180–1,  
184
- ergativity 8, 14, 25, 36, 59–63, 86, 186, 191,  
193–4, 195, 204
- erosion, *see* sound change
- expletive 36–7, *see also* dummy pronoun
- flat syntax 25, 28, 190
- focus, *see* contrastive focus
- free constituent order, *see* word order
- free phrase order, *see* word order
- free word order, *see* word order
- frequency 7, 15, 21, 34, 36, 37, 38, 40, 41, 46,  
53, 55, 63, 70, 71, 74, 75, 76, 78, 82, 92,  
94, 96, 97, 99, 100, 101, 102, 103, 104,  
105, 106, 107, 115, 121, 122, 130, 132,  
133, 135, 139, 141, 144, 145, 151, 155,  
158, 159, 161, 162, 172–3, 176, 183, 185,  
196–9, 202–4
- genitive 3, 6, 7, 8, 39, 40, 43, 79, 85, 86, 87,  
90–3, 94–110, 111–13, 114–17, 119, 120,  
124, 126, 129, 130, 133, 138, 139, 141,  
144, 145, 147–73, 175, 185, 186, 187,  
190–3, 195, 197, 199, 201, 203, 204,  
*see also* post-Vedic genitive shift
- gradual change 5, 21, 203, *see also* discrete  
change
- gram 20, 22–3, *see also* grammaticalization
- grammaticalization 1–9, 15, 19–23, 31, 32,  
47, 48–51, 55, 58, 59, 62, 64, 65, 69, 79, 83,  
86, 87, 90, 101, 108, 109, 113, 114, 115,  
116, 117, 119, 124, 148, 164, 166, 167,  
170–3, 175–6, 179–81, 182–3, 185, 187,  
188, 190, 191, 193, 196–7, 198, 201–4
- ground 6, 86, 106–10, 115, 147, 148, 149–58,  
159–60, 161, 162, 163, 169–70, 171–2,  
201, 203, *see also* possession
- basic 150, 155, 169
- non-basic 147, 148–58, 159, 160, 161, 162,  
163, 169, 170, 171, 172
- group 3, 5, 18, 46, 47, 86, 87, 89, 90, 92, 94,  
103, 106, 109, 110–11, 112, 113–16, 117,  
123, 124, 133, 135, 137, 141, 142, 143,  
145, 147–73
- asymmetrical 3, 5, 18, 47, 48, 50–1, 87, 93,  
113–16, 117, 124, 170, 172, 187, 196,  
202–3
- symmetrical 3, 5, 46, 47, 48, 50–1, 87–90,  
113–16, 117, 124, 171, 173, 187, 202–3
- habituation 49
- headedness 3, 9, 12, 39, 40, 43, 44, 47, 51, 68,  
86, 87, 92, 94, 95, 99, 106, 116, 119, 124,  
126, 131, 132, 141, 145, 171, 172, 175,  
178, 186, 187, 193, 202, *see also* group,  
asymmetrical; group, symmetrical
- head-marking 3, 23, 26–7, 42, 47
- host class 21, 170
- image schema 107–11, 150, 164, 169
- information structure 4, 30, 42, 120, 137,  
139–44, 154, *see also* contrastive focus
- intonation, *see* accentuation
- juxtaposition 26, 41, 42, 49, 50, 69, 73, 86, 87,  
89, 90, 94, 101, 117, 119, 135, 145,  
*see also* continuity

- language contact 121, *see also* Dravidian;  
linguistic area
- lexical structure 25–7
- linguistic area 121–2, 131
- local particle 2, 3, 8, 11, 15, 18, 40, 46, 47, 56,  
65–83, 85–90, 95, 96, 103, 110, 113–16,  
122, 132, 171, 196, 202, 203, *see also*  
group, symmetrical; preposition;  
preverb
- local case, *see* case, morphological; locative
- locative 6, 40, 54–5, 56, 57, 58, 60, 61, 62, 79,  
82, 85, 86, 92, 93, 94, 95, 96, 97, 100, 103,  
104, 107, 129, 159, 168, 193, 197
- metaphor 7, 21, 22, 50, 54, 63, 109, 116,  
147–58, 166, 170, 171, 172, 201, 203
- metonymy 21, 22, 50, 109, 116, 147–8, 151,  
163–72, 201–3
- nominalization 43–4, 86–7, 92–5, 97–110,  
112, 202
- non-configurationality 3–7, 23–32, 33–42,  
43, 46, 48–51, 87, 90, 93, 113–14, 132,  
135, 148, 170–3, 201–4, *see also*  
configurationality
- noun-adjective distinction 5, 7, 28, 29, 31–2,  
33, 40–2, 44–5, 46, 96, 204
- noun phrase 1, 45, 47, 69, 70, 124, 171, 187,  
190
- null anaphora 23–7, 34–6, 42
- obligatorification, *see* obligatoriness
- obligatoriness 1, 3–7, 18, 28, 31–3, 36, 40–5,  
48–50, 68, 86, 94, 130, 139, 147–73, 183,  
201–4, *see also* null anaphora
- oblique case 1, 6, 11, 43, 44, 55, 62–3, 68, 70,  
85–6, 121, 161, 163, 167, 175, 177, 185,  
187–8, 189, 191, 193–7
- overt expression, *see* obligatoriness
- overtiness, *see* obligatoriness
- P2 26, 28, 29, 31, 37, 38, 39, 40, 47, 51,  
123–31, 132–5, 136–45, 203, *see also*  
Wackernagel position
- paradigm 2, 104, 144, 145, 161–2, 203
- paradigmatic complementarity 175–9,  
180, 183, 184, 193–6, 197–8
- paradigmaticity 176–9, 184, 193, 194, 198
- paradigmatization 2, 8, 18, 48, 50, 51, 64,  
175–99, 204
- participle 3, 7, 9, 11, 12, 57, 58, 60, 62, 63–4,  
79, 87, 95, 96, 97, 99, 102, 114, 117, 120,  
153, 181, 185, 186, 188, 191, 195, 197, 202
- periphrasis 11, 12, 58, 71, 79, 83, 94, 95, 96,  
99, 102, 148, 158–60, 161, 162, 181
- phrase structure 1, 5, 9, 17, 25, 27, 28, 32, 42,  
48, 49, 50, 65–6, 69, 80, 83, 114, 148,  
170–3, 190, 201–4, *see also*  
configurationality; noun phrase;  
postpositional phrase
- possession 39, 40, 43, 44, 63–4, 86, 94, 113,  
114, 119, 123, 124, 135–44, 147, 148–58,  
159, 160–2, 163–70, 171–2, 186, 191,  
201, 203, *see also* genitive; pronominal  
possessor
- postposition 1–18, 32, 40, 43–4, 47, 48, 49,  
50, 51, 53–64, 65–6, 68–71, 72, 76, 77,  
79, 80, 83, 85–7, 94, 100, 114, 116,  
119–31, 132, 144–5, 147, 161, 169, 172,  
175–99, 201–4, *see also* adposition;  
preposition; postpositional phrase
- simple 2, 7, 11, 15, 43, 44, 47, 53–64,  
175–99, 201, 204
- complex 1, 14, 43, 53, 57, 58, 100, 175,  
178, 185, 193, 194
- compound 43, 175, 178
- postpositional phrase 1–11, 18, 21, 33, 37,  
44, 47, 50–1, 65, 69, 85–6, 92–3, 94, 109,  
111, 112, 113, 114, 122, 124, 128, 132,  
135, 137, 139, 143, 147, 148, 162, 175,  
187, 193, 196, 201–4, *see also*  
postposition
- post-Vedic genitive shift 94–110, 116, 119,  
169, 175, 185, 197, 199
- precedence 24
- preposition, *see* adposition, preposition
- prepositional phrase, *see* adpositional phrase,  
prepositional phrase

- preverb 56, 65, 67, 71–6  
pro-drop, *see* null anaphora  
projection principle 25–7  
Pronominal Argument Hypothesis 26, 34  
pronominal possessor 16, 18, 40, 50, 119,  
126, 133, 135–45, 147, 148, 152, 154–5,  
157, 160–2, 167, 170, 172, 203  
pronoun 16, 18, 26, 28, 29, 32, 35, 36, 37, 38,  
39, 40, 43, 50, 51, 81, 85, 96, 119, 123,  
125, 126, 132, 133, 135–44, 145, 147,  
148, 152, 154, 155, 157, 158, 159,  
160–2, 167, 170, 172, 183, 185, 188,  
189, 190–3, 196, 197, 198, 203, 204,  
205, *see also* pronominal argument  
hypothesis (PAH)  
prosody, *see* accentuation  
  
reanalysis 3, 8, 46, 47, 49, 65–70, 80–3, 86,  
87, 89, 90, 93, 94, 95, 97, 100, 102,  
103–10, 113–16, 117, 132, 148, 167, 169,  
171, 172–3, 185, 196, 199, 203  
relational noun 7, 11, 20, 40, 43, 86–94, 95–9,  
100–10, 111, 113, 114, 116–17, 119–20,  
123, 129, 130, 145, 147, 168, 175, 185,  
197, 199, 202  
ritualization 49  
  
scrambling, *see* word order  
secondary predication 30, 38, 39  
semantic change, *see* metaphor;  
metonymy  
semantic orientation 66, 67, 70, 72–83, 89,  
95, 122, 202  
  
slot-formation scenario 176, 182–4,  
197–8, 204  
sound change 8, 9, 53, 59, 62, 63, 167, 171,  
179, 196  
source construction 2, 9, 65, 85–117, 175,  
184, 191, 196, 202, 203  
split topicalization 31  
Sprachbund, *see* linguistic area  
subject-object asymmetry 24, 25, 27, 36,  
37, 42  
symmetrical group, *see* group  
syncretism 8, 86, 87, 93, 103–6, 129, 167,  
187, 197, 202  
synsemantic 20–2, 46, 49, 50, 71, 147–8, 166,  
167, 171–3, 181  
synthetic expression 8, 47, 96, 99, 114,  
161, 162  
  
tmesis 72, 74, 76  
  
Wackernagel position 28–9, 119, 123–31,  
132–5, 143, 144–5, *see also* P2  
word order 1, 3, 4, 7, 8, 9, 11, 14, 23–7, 28, 29,  
33, 39, 44, 49, 68, 69, 71, 77, 80, 86, 88,  
90, 119–45, 148, 166, 167, 171, 187,  
196, 203  
free constituent order 23, 27, 32, 42, 51  
free phrase order 4, 27  
free word order 4, 23–7, 33, 49, 119, 120,  
139, 144, 171  
juxtaposition, *see* continuity  
permutation 24, 43  
scrambling 4, 27