Adjective attribution

Michael Rießler

DRAFT
of Friday 20th March, 2015, 16:11

Studies in Diversity Linguistics 2

language science press

Adjective attribution

This book is the first typological study of adjective attribution marking. Its focus lies on Northern Eurasia, although it covers many more languages and presents an ontology of morphosyntactic categories relevant to noun phrase structure in general. Beside treating synchronic data, the study contributes to historical linguistics by reconstructing the origin of new types specifically in the language contact area between the Indo-European and Uralic families.

DRAFT
of Friday 20th March, 2015, 16:11

ISBN 978-3-944675-49-7

Michael Rießler

Adjective attribution

Studies in Diversity Linguistics

Chief Editor: Martin Haspelmath
Consulting Editors: Fernando Zúñiga, Peter Arkadiev, Ruth Singer, Pilar Valen-zuela

In this series:

1. Handschuh, Corinna. A typology of marked-S languages
2. Rießler, Michael. Adjective attribution
3. Klamer, Marian (ed.). The Alor-Pantar languages: History and typology
4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea)
5. Wilbur, Joshua. A grammar of Pite Saami

ISSN: 2363-5568

ISSN: 2363-5568

Adjective attribution

Michael Rießler

DRAFT
of Friday 20th March, 2015, 16:11

language
science
press

Language Science Press
Berlin

Language Science Press
Habelschwerdter Allee 45
14195 Berlin, Germany

langsci-press.org

This title can be downloaded at:
http://langsci-press.org/catalog/book/19
© 2015, Michael Rießler
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/
ISBN: 978-3-944675-49-7
ISSN: 2363-5568

Cover and concept of design: Ulrike Harbort
Typesetting: Michael Rießler, ??
Proofreading: Alexandra Kellner, Joshua Wilbur, ??

Storage and cataloguing done by FU Berlin

Language Science Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables and other factual information given in this work are correct at the time of first publication but Language Science Press does not guarantee the accuracy of such information thereafter.

за най-любимите ми Алма, Ива и Кристина

Preface

Acknowledgements

List of abbreviations and notational conventions

Morphological glosses

The following list includes only abbreviations for glossing of linguistic examples not defined by the Leipzig Glossing Rules.[1]

ABESS	abessive
ADJZ	adjectivizer, adjectivization
AGR	(any kind of) agreement
ATTR	or (attr.); attribution, attributive
ANR	action nominal(izer)
COMPAR	comparative (adjective derivation)
CONTR	contrastive focus
CRS	currently relevant state
DERIV	derivative, derivation (unspecified)
DIM	diminutive
ESS	essive
HUM	human (gender)
ILL	illative
INFL	(any) inflection
MOD	modification
NAR	narrative (case)
NONFUT	non-future
NONHUM	non-human (gender)
PFCT	perfective (verb derivation)
PRED	or (pred.); predication, predicative
PREPOS	prepositional
REAL	realis
STAT	stative (verb derivation)
SUPER	superlative
UTR	utrum, common (gender)

Syntactic classes and phrase constituents

A	adjective
AdP	adpositional phrase
AP	adjective phrase

[1] http://www.eva.mpg.de/lingua/resources/glossing-rules/ 16.02.2014

ART	(attributive) article
CASE	case (clitic)
DEF	definite article
Deg	degree word
HEAD	phrase head
INDEF	indefinite article
N	noun
NP	noun phrase
PSD	possessed (head in possessive noun phrase)
PSR	possessor (dependent in possessive noun phrase)
Rel	relative clause
V	verb

Abbreviations for cardinal directions

C	Central	S	South(ern)
E	East(ern)	SE	South-East(ern)
N	North(ern)	SW	South-West(ern)
NE	North-East(ern)	W	West(ern)
NW	North-West(ern)		

Other symbols

The following symbols are used for the illustration of linguistic changes.

~	variant
<	borrowing
←	derivation or other synchronic process
⇐	grammaticalization or other diachronic process[2]

[2] Note that the term *grammaticalization* is used for different types of linguistic changes leading to re-analysis of a given construction's grammatical meaning. A prototypical instance in this rather broad sense of grammaticalization is the morphologization of a formerly lexical morpheme to a grammatical morpheme, as the development of definite markers from anaphoric pronouns in Germanic languages, e.g. English *the (house)* (*the* ⇐ Old English *þæt*) and Swedish *(hus)-et* (*-et* ⇐ Old Norse *hið*).

Contents

Part I

Preliminaries

1 Introduction

Aim

The aim of this investigation is to typologize adjective attribution marking devices in the languages of northern Eurasia. Agreement and construct state marking are commonly known morphological devices for the licensing of adjectival modifiers; an example of a purely syntactic device is juxtaposition.

The main parts of this thesis include an ontological classification of all attested devices in the geographic area of investigation and a survey of adjective attribution marking devices occurring across the northern Eurasian language families. Finally, several attested scenarios for the evolution of adjective attribution marking devices in languages of northern Eurasia are discussed.

Question

The most central questions dealt with in this investigation regard the formal licensing of the syntactic relation between a head noun and its adjectival dependent inside a noun phrase:

- What syntactic, morphological or other adjective attribution marking devices are available in languages?
- How can these devices be systematically described and typologized?
- How is the occurrence of the different types distributed geographically?
- How does attribution marking arise and diffuse across languages?

Method

The present study is the result of empirical research based on data from grammatical descriptions on the investigated languages. It follows a data-driven, bottom-up and framework-neutral approach (**haspelmath2010**) and also the method of "Autotypology" following **bickel-etal2002** and **bickel2007**

The method of sampling and mapping of data is inspired by the *Autotyp*[1] and *Eurotyp*[2] research programs and the *WALS* project.[3] The approach presented here is closer to *Eurotyp* than to *WALS* or *Autotyp* in coding as many different genera from the geographic area of investigation as possible.

Content

The book is divided into four main parts. In Part I "Theoretical preliminaries", a few basic comparative concepts relevant to a framework-neutral description of a noun phrase and its constituents are introduced. This part also discusses the syntax-morphology interface in noun-phrase structure which is of central importance for the present study.

Part II "Typology" presents a general ontology of adjective attribution marking devices based on data from northern Eurasian and other languages.

In Part III "Synchrony", a synchronic-typological survey of noun phrase structure with attributive adjectives in northern Eurasia is presented and exemplified with data from all genera of the area.

The book's last Part **??** "Diachrony" is devoted to the evolution of adjective attribution marking devices. It describes several different paths of evolving and abolishing adjective attribution marking devices in northern Eurasian languages.

[1] Cf. http://www.spw.uzh.ch/autotyp/ 16.02.2014

[2] Cf. http://www.degruyter.com/view/serial/16329 16.02.2014

[3] Cf. http://wals.info 16.02.2014

2 Noun phrases and adjectival modifiers

2.1 Noun phrases

A noun phrase is a referential syntactic unit which can serve as subject, object or oblique argument of a verb or as predicative complement of a nominal sentence. Furthermore, a noun phrase can be used in adverbial and adnominal functions. According to common syntactic models, the head determines the category of the phrase and governs the dependent constituent(s) in the phrase (**nichols1986**). Consequently, the head of a noun phrase is a noun (or a pronoun). Dependent constituents in noun phrases, also called "attributes", narrow the denotation, i.e. modify the head noun descriptively. Typical modifiers in noun phrases are nouns (or noun phrases), adjectives (or adjective phrases), adpositions (or adpositional phrases) and relative clauses, as in the following example.[1]

(1) $[_{NP}[_{PSR}$ *her*$][_{AP}$ *brand new*$]$ *house* $[_{AdP}$ *over there*$][_{Rel}$ *which is big*$]]$

Noun phrases can thus contain simple modifiers, like nouns or adjectives, or more complex types of modifiers which are complex phrases themselves: for instance (possessor) noun phrases (*my*), adjective phrases (*brand new*), adpositional phrases (*in the village*) or relative clauses (*which was expensive*).

2.2 Adjectival modifiers

This book presents a cross-linguistic comparison of attributive adjectives. It investigates the syntactic and morphosyntactic behavior of adjectives inside noun phrases, in particular how they are formally licensed as dependent constituents in noun phrases.

The notion "adjective" needs some clarification because adjectives do not constitute a universal syntactic category. Whereas in some languages adjectives form a distinct word class, in other languages adjectives may not be clearly distinguishable from other parts of speech and constitute a flexible category to-

[1] Possible hierarchies of constituents inside this noun phrase are ignored in this illustrating example.

Hixkaryana

gether with nouns or with verbs. In a third group of languages, adjectives do not exist et all.

For the survey of languages considered in this investigation, the term *adjective* had thus to be defined in a purely semantic sense, as words with a lexical meaning referring to properties or qualities such as 'high', 'beautiful', 'red', etc. "Qualifying modifiers" (**rijkhoff2002**) in this broad sense are all lexical elements specifying properties of their referents. This definition excludes possessive pronouns, demonstratives, numerals, and words meaning 'other', all of which may behave syntactically like adjectival modifiers in several languages. On the other hand, the semantic definition of adjectives includes adjectival nouns and adjectival verbs (cf. "nouny" and "verby" adjectives in **wetzer1996**) and even qualifying modifiers which are true verbs or true nouns in some languages.

Even though adjectives do not constitute a universal syntactic category, almost all languages seem to exhibit some type of modifier construction in the noun phrase to specify qualitative properties. Hixkaryana, a Carib language spoken in Brazil, however, has been mentioned as a counterexample because qualitative properties are only expressed in predicative constructions (**derbyshire1979 rijkhoff2002**). If a language does not exhibit a distinct class of adjectives, inher-

Type 1	(Flexible)	V / N / A		
Type 2	(Flexible)	V	N / A	
Type 3	(Differentiated)	V	N	A
Type 4	(Rigid)	V	N	
Type 5	(Rigid)	V		

Figure 2.1: Parts-of-speech systems (**hengeveld-etal2004**)

ent properties of the referent are most often expressed by other lexical means, for example by a relative clause (headed by a finite stative or descriptive verb) used as an adnominal modifier or by a qualifying noun phrase (headed by an abstract, property marking noun) as adnominal modifiers (**rijkhoff2002**).

Similar to **hengeveld-etal2004** the present study is based on the characterization of adjectives as semantic predicates which can be used as modifiers of nouns without further (derivational) operations. A typology of parts-of-speech systems is illustrated in Table 2.1.

In the "flexible" language types 1–2 in Table 2.1, certain classes of lexemes can occur in more than one function (as verbs/nouns/adjectives in Type 1 or as nouns/adjectives in Type 2). In the "differentiated" type of languages, on the

Indo-European languages
Basque
Uralic languages
Caucasian languages
Mongolic
Tungusic languages
Turkic languages
Ainu (Shizunai)
Korean
Nivkh

other hand, the various classes of lexemes are strictly divided according to their function and constitute a tripartite system of lexeme classes with verbs/nouns/-adjectives (Type 3). The "rigid" types of languages exhibit either a bipartite system with verbs/nouns (Type 4) or a system exhibiting only one class of lexemes: verbs (Type 5).[2]

Most northern Eurasian languages belong to a type of language which exhibits a distinct class of adjectives, whether flexible or rigid (and whether this class is open or closed and counts only very few lexemes). Languages spoken on the European subcontinent predominantly belong to Type 3 and exhibit adjectives as a distinct major class. Most Indo-European languages of northern Eurasia belong to this type, but also Basque, the Uralic languages of Europe and most languages belonging to one of the three Caucasian language families.

Type 2 languages with a flexible class of "noun-adjectives" are also well represented in northern Eurasia. In practically all Mongolic, Tungusic and Turkic languages, for example, there is usually no sharp distinction between adjectives and nouns (**rijkhoff2002 poppe1964**).

Type 4 languages lacking a flexible or distinct class of adjectives are represented, for example, by Ainu, Korean and Nivkh. In these languages, verbs are normally employed as qualifying adnominal modifiers.

Languages of Type 1 (with a flexible class of "verb-adjectives") or 5 (exhibiting exclusively verbs) are not represented in the northern Eurasian area.

2.3 Syntax of adjectival modification

The present book deals with noun phrases in which adjectives occur as attributes. Predicative adjectives are not dealt with systematically,[3] although in some cases attributive and predicative adjectives will be contrasted to each other, especially if the languages in question code them differently. The main question to answer with my investigation is how different languages license the syntactic position of adjectival modifiers inside noun phrases, i.e. what grammatical devices are used for the encoding of the syntactic relationship between an adjectival dependent and its head noun.

[2] The classification of **hengeveld-etal2004** has seven types because the authors also include manner adverbs as a distinct class. According to the original classification, Type 3 in Table 2.1 should thus be divided further yielding the three subtypes V–N–A/Adv (flexible), V–N–A–Adv (rigid) and V–N–A (rigid).

[3] A typology of adjective predication is **wetzer1996**

English
English
English
English
English
English

2.3.1 Noun phrase internal syntax

The syntactic relationship between noun phrase constituents can be encoded by means of purely syntactic structures, i.e. simply stringing together constituents, or by adding syntactic or morphological devices.

The adjective can take up the modifier slot in the noun phrase without further syntactic or morphological marking taking place inside the noun phrase. Such syntactic licensing means that the relationship between dependent and head is encoded purely structurally in terms of designated positions. An instance of purely syntactic licensing are noun phrases with adjectival modifiers in English. The adjective obligatorily precedes the noun but is not marked otherwise.

(2) English (Indo-European; own knowledge)
large houses

An example of a syntactic device is the dummy head *one* in English which occurs obligatorily in noun phrases without lexical heads.

(3) English (Indo-European; own knowledge)

a. *a large one*
INDEF large HEAD:SG
'a large one'

b. *large ones*
large HEAD:PL
'large ones'

The dummy head *one* is a noun phrase constituent itself, hence a true syntactic attribution marking device, even though morphology is also involved in this syntactic structure because *one* is inflected for number. The difference between covert and overt syntactic attribution marking devices can also be illustrated with different relative clauses in English.

(4) English (Indo-European; own knowledge)

a. [$_{NP}$ *the house* [$_{REL}$ *I built*]]

b. [$_{NP}$ *the house* [$_{REL}$ *that I built*]]

i. [$_{NP}$ *the man* [$_{REL}$ *who*$_{nom}$ *built a house*]]

ii. [$_{NP}$ *the man* [$_{REL}$ *whose*$_{gen}$ *house was built*]]

Whereas (4a) exemplifies a covert syntactic device because the relative clause is simply juxtaposed, (4b) is an overt syntactic device because the the relative clause

is marked by an invariable formative. In (4b-i+4b-ii), the relativizer *who* is also an overt syntactic device. But in the marking of this relative clause construction, morphology is involved too because the relativizer inflects for case according to the semantic role of the relativized noun.

German
German
German
German
German
German

Morphological attribution marking devices are either overt (linear or else) morphemes bound to constituents or covert morphological processes, like incorporation.[4] A prototypical instance of a morphological adjective attribution marking device is agreement inflection, as in German.

(5) German (Indo-European; own knowledge)
groß-e Häus-er
big-PL house-PL
'large houses'

Agreement inflection of attributive adjectives in German is a morphological device, it exists only because syntax requires it, hence a morphosyntactic device. Other morphological marking in German occurs on syntactic units or on constituents of syntactic units without belonging to morphosyntax. For instance, the plural inflection on the head noun (*Häus-er*) or the inflectional circumfix yielding a participle (*ge-bau-t*) in (6) belongs exclusively to the level of (inflectional and derivational) morphology but not to syntax.

(6) German (Indo-European; own knowledge)
ge-bau-t-e *Häus-er*
PTCP-build-PTCP-PL house-PL
'built houses'

Note that adjectives have been characterized as predicates which can be used as modifiers of nouns without further (derivational) operations. Consequently, the German participle stem *gebaut* (← *bauen* + *ge- … -t*) is an adjective in this broad sense. Syntactically, the participle behaves like a true adjective and takes similar attribution marking. The attribution marking device (i.e. the agreement inflection) attaches to the participle stem as such (marked with parentheses in 6). The participle inflection of the verb root *bau-* yielding this new stem does not belong to the sphere of syntax. Similarly, category-changing derivational morphology in other languages yielding, for example, a stative verb or a participle function, is not considered to be morphological licensing of adjectival modification.

[4] Morphological attribution marking devices can also attach to complex constituents, as the possessor marking clitics in English or Swedish which attach to noun phrases: Swedish [NP[NP kungen]=s rike] the_king=POSS empire 'the empire of the king', [NP[NP kungen av Sverige]=s rike] the_king of Sweden=POSS empire 'the empire of the King of Sweden'.

German
German
German

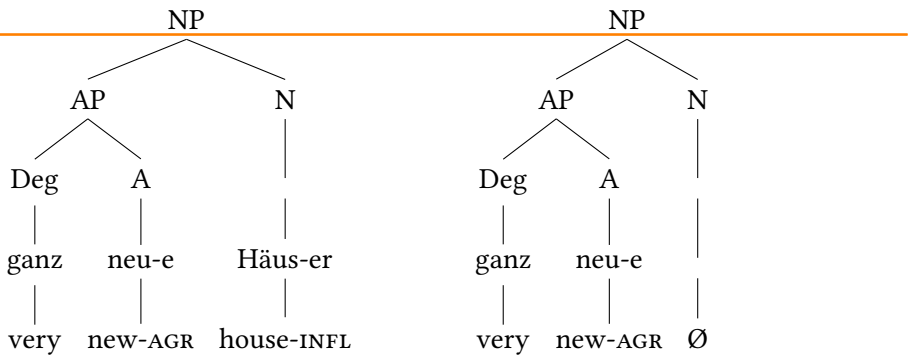
2.3.2 Headless noun phrases

Adjectives as well as various other modifiers can also occur in noun phrases without a noun. Normally, this is the case with adjectives in elliptical constructions or adjectives which are otherwise substantivized. In many languages, noun phrases with and without an overtly expressed head noun exhibit a similar phrase structure, as in the following examples from German.

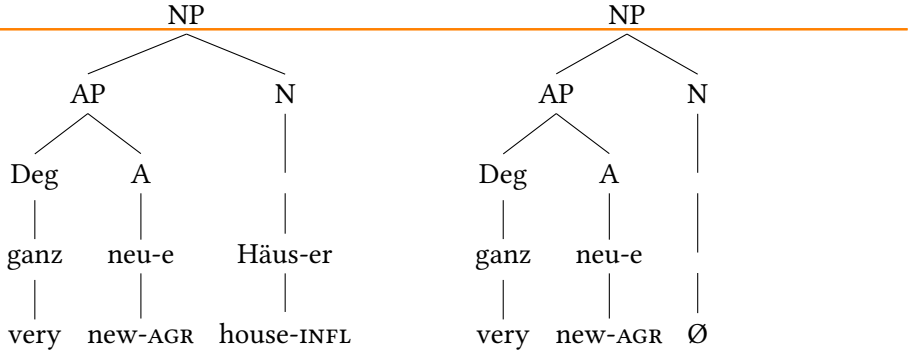
(7) German (Indo-European; own knowledge)

a. *ganz neue Häuser* b. *ganz neue* (viz. *Häuser*)

check tree (solid lines)
check spacing after trees



check tree (solid lines)
check spacing after trees



The syntactic structure of the two examples in (7) is principally identical except for the missing head noun 'house' with its morphological plural marking in the second structure. The attributive adjective 'new' is marked for the same morphosyntactic agreement features in both examples. Even though the adjective in the headless phrase is semantically substantivized and used referentially, it is still syntactically the modifier of the (ellipted) noun 'house'. The syntactic status of the modifier as head of an adjective phrase is indicated by its ability to take dependents, such as the degree word 'very'. German thus allows the syntactic

head position to remain empty in elliptical constructions.

In other languages, accepting an empty head position in the (elliptical) noun phrase seems less straightforward. In Kildin Saami, for example, nouns and adjectives share identical inflection paradigms. As modifiers of nouns, however, adjectives are not inflected but are simply juxtaposed,[5] as in (8a) and (8b). Only when attributive adjectives occur in elliptical noun phrases are they inflected identically to nouns, as in (8c) and (8d).[6]

Saamic, Kildin
Saamic, Kildin
Saamic, Kildin
German
German
Saamic, Kildin
Saamic, Kildin
Saamic, Kildin

(8) Saamic, Kildin (Uralic; own knowledge)

a. *čofta odt pērrht*
very new house(NOM:SG)
'a very new house'

b. *čofta odt pērht-es't*
very new house-LOC:SG
'in a very new house'

c. *čofta odt (viz. pērrht)*
very new(NOM:SG)
'a very new one'

d. *čofta od-es't (viz. pērht-es't)*
very new-LOC:SG
'in a very new one'

If the elliptical construction in Kildin Saami is analyzed as having an empty syntactic head position, as in German, an explanation for the different behavior of the (nominal) case inflection is needed. Unlike in German, where (nominal) inflection is always bound to the noun, inflection in Kildin Saami can occur bound to nouns or adjectives. Case marking in Kildin Saami could thus be analyzed as clitic and bound to the whole noun phrase and hence showing up on the rightmost phrase constituent.

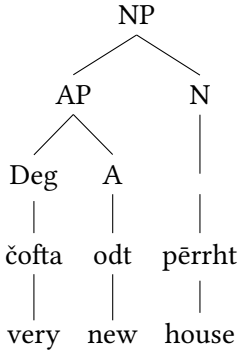
(9) Saamic, Kildin (Uralic; own knowledge)

[5] This is true only for one class of adjectives. Other adjective classes show different morphosyntactic behavior, see §?? below.

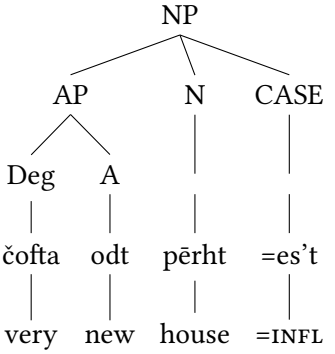
[6] The stem alternation in the adjective *odt : od-* is due to a regular morpho-phonological process.

English
English
English

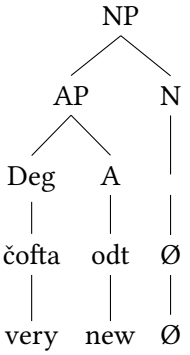
a.



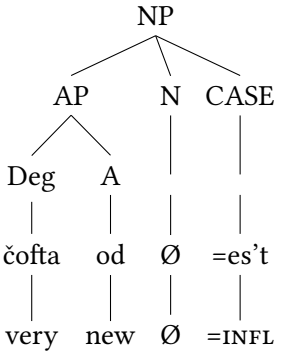
b.



c.



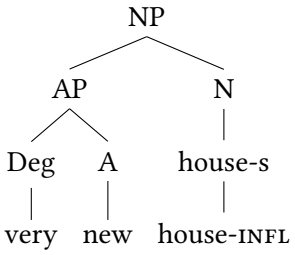
d.



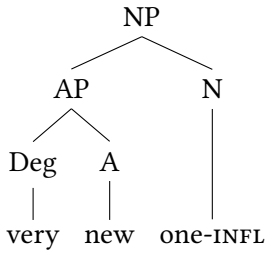
Another type of language in which elliptical noun phrases behave differently is exemplified by English. In elliptic constructions, attributive adjectives are obligatorily marked with the marker *one*. This marker is exclusively used in headless noun phrases with adjectival (and some other) modifiers. It never occurs if the head noun is overtly expressed.

(10) English (Indo-European; own knowledge)

a. very new houses



b. very new ones (viz. houses)



check spacing after trees

Being a grammatical word, hence a constituent in the phrase structure, the marker *one* in English is sometimes described as "dummy head" (**rijkhoff2002**) replacing

the noun at the syntactic head position. Consequently, it could be argued that the syntactic head position is never empty in English.

English
Georgian
Georgian

2.3.3 Appositional modification

Apposition[7] is commonly described as a sequence of two (or more) co-referential constituents on the same syntactic level and hence with the same syntactic function, as in the following expression.

(11) $(_{np}[_{NP}$ *Alma and Iva*$][_{NP}$ *my daughters*$])$ *are in this picture.*

Syntactically, the two independent noun phrases *Alma and Iva, my daughters* together serve as one argument phrase in (11).[8] In other words, apposition can be defined as a single semantic phrase which consists of several independent syntactic phrases which together serve one syntactic function.

Appositional modification differs from true apposition in that the apposed constituent phrase is semantically and syntactically dependent on the other constituent phrase. Similar to the definition presented in **rijkhoff2002** appositional (noun) modification is here understood as a construction in which the dependent constituent is not part of the (integral) phrase headed by the modified noun. Semantically, the appositional modifier is headed by the modified noun. Syntactically, however, the appositional modifier has an empty head which is co-referential with the head noun of the apposed noun phrase.

Appositional modification seems to occur as a secondary marked type of adjective attribution marking in several languages, for instance in Georgian. Attributive adjectives are normally preposed and show only limited agreement (see 12a). In postposition (marking emphasis), however, the adjective inflects for the full set of cases and numbers (12b). This construction thus resembles an independent (headless) noun phrase in apposition to the semantic head (**testelec1998**); cf. also §**??** below.

(12) Georgian (Kartvelian; **testelec1998**)

a. *am* *or* ***lamaz*** *kal-s*
that:OBL two nice:OBL woman-DAT
'to those two nice women'

[7] Note the different meaning of "juxtaposition", which is defined as a distinct functional type in §4.2.

[8] The notation of the appositional unit in round brackets is borrowed from **rijkhoff2002**

Bulgarian
Bulgarian
Bulgarian
Bulgarian
Georgian
Bulgarian
Georgian
German
German

b. *kal-eb-s* ***lamaz-eb-s***
woman-PL-DAT nice-PL-DAT
'to the NICE women'

Even without differentiated attribution marking, constituent order change between attribute and head can indicate apposition, as in Bulgarian. Note that the constituent order in noun phrases of Bulgarian is strictly head-final. In poetic language, however, it is possible to move the adjective after the noun.

(13) Bulgarian (Indo-European; own knowledge)

a. *tezi* ***golem-i*** *gradove*
these big-PL towns
'these big towns'

b. *tezi* *gradove* ***golem-i***
these towns big-PL
'these big towns'

It seems impossible to prove that Bulgarian presents an example of appositional modification. The emphasized noun phrase in (13b) could simply be analyzed as integral noun phrase differentiated from other non-emphasized noun phrases by word order. Georgian, however, is different from Bulgarian. The emphasized noun phrase in (12b) exhibits different morphosyntactic marking due to the additional agreement features (Georgian) and is very likely to be analyzed as an attributive appositional construction.

Evidence for appositional modification as a syntactically distinguished noun phrase type is also found in constructions were the apposed headless noun phrase is overtly marked by means of attributive nominalization (cf. §4.5.2.3). Attributive nominalization can be illustrated with the epithet construction in German.

(14) German (Indo-European; own knowledge)
$[_{NP}$ *Friedrich* $[_{NP}$ *der Große*$]]$ *'Frederick the Great'*

3 The morphology-syntax interface

3.1 Morphosyntax

An inventory of grammatical features relevant to morphology and its interfaces with semantics and syntax has recently been systematized and presented in a volume edited by **kibort-etal2010** specifically in the chapter by **kibort2010a** Kibort and Corbett's typology of morphosyntactic features, which is grounded in other work, for instance by **aronoff1994**; **corbett1987**; **carstairs-mccarthy1999**; **corbett2006**; **corbett-etal2006**; **bickel-etal2007**; **kibort2008a** will be evaluated in the following sections. It will be shown that true morphosyntactic features (i.e. features not interfacing with semantics) relevant to noun phrase structure are missing but have to be added to such an inventory.

Note that the term "morphosyntax" is sometimes inaccurately used for any type of syntactic construction in which morphological processes take place. It is also commonly used as a homonym for "grammar" or "morphology and/or syntax" thus subsuming all kinds of morphological and syntactic structure of a language. For the present study, however the scopes of syntactic and morphological processes are differentiated from each other. Consequently morphosyntax is here understood as the interface between syntax and morphology, i.e. syntactic structure assigning morphology on one or more of its constituents.

3.1.0.0.1 Morphological features Strictly morphological features have only inherent values, i.e. the assignment of these values is not sensitive to syntax. Morphological features include values which are either fixed, i.e. supplied on the lexical level, or selected from a range of values. The selection of these values is based only on formal criteria. A prototypical example of a purely morphological feature is inflection class.

3.1.0.0.2 Morpho-semantic features Morpho-semantic features also only have inherent values whose assignment is not sensitive to syntax. The values of morpho-semantic features are selected from a range of values. However, unlike purely

Albanian
Bulgarian
Rumanian
Balkan languages

morphological features, the selection is based on semantic criteria. A prototypical example of the assignment of a morpho-semantic feature is definite marking.

Albanian
Rumanian

3.1.0.0.3 Morphosyntactic features Morphosyntactic features are sensitive to syntax because either agreement or government is involved in the assignment of their values. In the case of agreement, however, a morphosyntactic feature belongs per definition both to morphosyntax – due to the feature's contextual assignment to the agreement target – and simultaneously to pure morphology (or morpho-semantics) – due to the feature's status inherent in the agreement trigger.

The difference between morphosyntactic and purely morphological (or morpho-semantic) features can be illustrated by definiteness marking in Albanian, Bulgarian and Rumanian. The definite markers in these three Balkan languages are bound morphemes in postposition, cf. (1a) (2a) (3a). The syntactic behavior of the definite marker in all three languages is also similar: In noun phrases with modifying adjectives the marker attaches enclitically to the first constituent.

(1) Albanian (Indo-European; **buchholz-etal1987**)

a. *djal=i*
boy(M)=DEF:M.SG
'the boy'

b. *djal=i* *i* *mire*
boy(M)=DEF:M.SG ATTR:DEF.M.SG good.M.SG
'the good boy'

c. *i* *mir=i* *djalë*
ATTR:DEF.M.SG good=DEF:M.SG boy(M)
'the GOOD boy'

(2) Rumanian (**beyer-etal1987**)

a. *băiat=ul*
boy(M)=DEF.M.SG
'the boy'

b. *băiat=ul* *bun*
boy(M)=DEF.M.SG good.M.SG
'the good boy'

c. *bun=ul* *băiat*
good=(M)-DEF.M.SG boy(M)
'the GOOD boy'

(3) Bulgarian (own knowledge)

a. *momče=to*
boy(N)=DEF.N.SG
'the boy'

b. *dobro=to momče*
good=DEF.M.SG boy(N)
'the good boy'

Bulgarian
Albanian
Bulgarian
Rumanian
Albanian
Albanian
Bulgarian
Rumanian

The feature SPECIES,[1] however, does not belong to morphosyntax in all of these three languages. Even though the definite marker shows the same syntactic behavior (i.e. attaching in second-position), the morphological feature SPECIES is sensitive to syntax only in Albanian. Whereas definiteness is a purely morpho-semantic feature not involved in any syntactic triggering in Bulgarian and Rumanian, in Albanian a second marker of definiteness occurs on the adjective. This marker is required by syntax through the mechanism of agreement. Hence, definiteness is morphosyntactic only in Albanian. In Bulgarian and Rumanian definiteness is purely morphological.

3.2 Morphosyntactic features

As shown in the previous section, *morphosyntactic marking* can basically be defined as *morphological marking relevant to syntax*. According to **kibort2010a** the syntactic relevance of a certain morphological marker is determined by the involvement of this marker in either agreement or government. Kibort's view of morphosyntax, however, is based on definitions of agreement and government which imply obligatory interfacing of the respective grammatical features with all three components: morphology, syntax and semantics. Hence, a "more accurate term [...] would be 'morpho-semantico-syntactic' features" (**kibort2010a**).

Kibort pages??

Both agreement and government require a syntactic constituent as trigger and another constituent as target of morphosyntactic marking. Kibort's terms *trigger* and *target* are used in the case of agreement marking, whereas *governor* and *governee* are the respective labels in the cases of government. Consequently,

[1] Typical values of SPECIES are, for instance, DEFINITE, INDEFINITE or SPECIFIC. The use of the term SPECIES (from Latin 'appearance, form') is borrowed from Swedish and Finnish grammatical terminology, **holm-etal1970**; **itkonen-t1980a** It will be used throughout this investigation instead of the commonly known "definiteness" because it seems terminologically odd to have a feature DEFINITENESS exhibiting a value with the similar label DEFINITE.

Persian
Persian
Persian
Persian
Persian
Persian
Bulgarian
Bulgarian

Kibort's *government* covers only morphosyntactic marking assigned by triggers (governors) which are constituents – like a head noun marked for certain gender and number values triggering gender and number *agreement* on the modifier.

Instances of morphological marking triggered not by constituents but by the syntactic structure as such seem to fall outside the range of Kibort's typology of morphosyntactic features. A prototypical example of morphosyntactic marking without a trigger inside the noun phrase is attributive state marking in Persian.

(4) Persian (**mahootian1997**)

a. *"Construct state" (i.e. attributive state)*
xâne-ye *bozorg*
house-CONSTRUCT big
'large house'

b. *"Absolute state" (i.e. predicative state)*
in *xâne* *bozorg* *ast*
DEM house(ABSOLUTE) big is
'the house is large'

In Persian, a nominal head is obligatorily inflected in the construct state if an adjective is present in the noun phrase. The trigger of the head-marking attributive suffix *-ye* in Persian is the syntactic structure alone. Since no other value than [+construct] is assigned, semantics cannot be involved. It could be argued that semantics is relevant to the choice of whether to use the adjective as attribute or as predicate and that the attributive inflection on the head noun is inherent (i.e. morpho-semantically assigned). Semantics (or pragmatics) is of course relevant to the speaker's decision to utter a noun phrase instead of a predication. Semantics is, however, irrelevant to the argumentation about the syntactic structure requiring certain morphological marking: Once the speaker has made her or his decision, it is the syntactic structure alone which is involved in the assignment of the relevant morphological marking. Consequently, attributive construct state in Persian is an example of true morphosyntactic marking.

Attributive construct state marking morphosyntactically similar to the Persian construct state marking occurs in many other languages. In Bulgarian, for instance, some nouns require a special inflection after numerals.

(5) Bulgarian (own knowledge)
dva stol-a
two chair(M)-CONSTRUCT
'two chairs'

Unlike attributive construct state marking in Persian, which occurs obligatorily in noun phrases with different types of modifiers (adjectives, nouns, and some other), attributive construct state marking in Bulgarian is restricted with regard to both dependent and head. Thus, it occurs only in noun phrases in which the modifier is a numeral higher than 'one' and in which the head noun belongs to the class of non-human masculines. In the Bulgarian grammatical tradition this inflectional marking is called the "counting form".[2] The marker originates historically from the genitive singular inflection of masculines. The diachrony, however, does not affect the analysis of this marker as belonging to the morphosyntactic feature STATE from a synchronic-typological point of view. Even though attributive construct state marking in Bulgarian is much more restricted than in Persian, it clearly belongs to the same type of syntactically assigned inflection on the head noun.

Persian
Bulgarian
Bulgarian
Bulgarian
Persian
Persian
Bulgarian
Persian
Saamic, Kildin
Saamic, Kildin
Saamic, Kildin
Saamic, Kildin

The term *state* here is adapted from **melcuk2006** who defines it as an inflectional category of nouns heading a noun phrase. According to Mel'čuk, the function of morphological state marking is licensing the syntactic relationship between the phrase constituents. In the case of head-marking state, as in Persian and Bulgarian (4+5), the head noun is inflected and shows the morphological value [+construct] if it is the governing member in the present syntactic relation (i.e. the noun phrase).

Even though *state* in Mel'čuk's (and others') terms is usually associated with head-marking constructions of the Persian type (cf. example 4), a similar morphosyntactic mechanism applies to dependent marking construct states in other languages. Consider, for example, Kildin Saami in which the dependent noun phrase of a postposition is obligatorily inflected in the genitive case.

(6) Saamic, Kildin (Uralic; own knowledge)
tuel' *al'n*
chair\GEN on
'on the chair'

It could be argued that the genitive inflection of 'chair' in example 6 is a morphological value of the feature CASE assigned to the dependent noun phase by the mechanism of *government*. But since genitive is the obligatory and only possible marker in postpositional phrases in Kildin Saami, there is no motivation for assuming that any case value is marked here. There is no semantic connection to a genitive case which marks a possessor noun in Kildin Saami either.[3] Since this

[2] Bulgarian *brojna forma*

[3] This is true from a synchronic point of view. Historically, the origin of the genitive marking

Saamic, Kildin
Saamic, Kildin
Saamic, Kildin
Saamic, Kildin
Saamic, Kildin
Russian
Russian

modification marker is assigned by the syntax of the specific construction alone, and since the only function of this marker is licensing the given syntactic relation (i.e. an adpositional phrase), a more appropriate gloss could be CONSTRUCT.

Several languages also exhibit dependent marking construct state in noun phrases. The matching value is usually glossed as ATTRIBUTIVE. In Kildin Saami, for example, members of one (lexically defined) subclass of adjectives are obligatorily inflected for attributive state if they are used as modifiers in a noun phrase.

(7) Saamic, Kildin (Uralic; own knowledge)

a. *Attributive adjective (cf. "attributive state")*
vīl'k-es' *puaz*
white-ATTR reindeer
'white reindeer'

b. *Predicative adjective (cf. "absolute state")*
puaz *lī* *vīll'k-e*
reindeer is white-PRED
'the reindeer is white'

The assignment of attributive inflection on (adjectival) modifiers of nouns as well as the assignment of genitive inflection on (nominal) modifiers of adpositions thus follow a similar syntactic mechanism in Kildin Saami: A certain syntactic relationship (i.e. an adpositional phrase or a noun phrase, respectively) is licensed by marking the dependent phrase constituent with the feature STATE.

Finally, the feature STATE may not only be dependent-marked, as in Kildin Saami, but can even interfere with other features. Whereas attributive state marking is invariable in Kildin Saami, in other languages it shows interference with semantic values assigned through the mechanism of agreement. The agreement inflection of attributive adjectives in Russian, for instance, marks the syntactically governed feature STATE simultaneously with the morphosyntactically governed features NUMBER/GENDER/CASE.

(8) Russian (Indo-European; own knowledge)

a. *Attributive adjective inflection (cf. "attributive state")*
belyj *olen'*
white:ATTR:M.SG reindeer
'the white reindeer'

in adpositional phrases is easily accounted for and goes back to possessor marking in noun phrases with relational head nouns. But again, the diachrony of a certain marker is not relevant to its synchronic-typological categorization.

b. *Predicative adjective inflection (cf.“absolute state”)*
olen’ bel
reindeer white:PRED:M.SG
‘the reindeer is white’

3.3 An ontology of morphosyntactic features

Besides introducing a few very basic notions connected to noun phrase structure and adjectival modification, the syntax-morphology interface has been discussed in the theoretical sections above. In particular, Kibort’s (**kibort2010a**) inventory of grammatical features relevant to morphology and its interfaces with semantics and syntax have been critically evaluated. True morphosyntactic features (i.e. features not interfacing semantics) are not yet included in her inventory of grammatical features. The argumentation in the present chapter aims at establishing a new feature STATE, which according to Kibort’s own definitions must be regarded as a true morphosyntactic feature and which should definitely be added to Kibort’s list.

Figure 3.1 shows the morphosyntactic features relevant to the present inventory of noun phrase types. Note that only the rightmost feature (6) in that figure can be characterized as being of true *morphosyntactic* nature. The group of features under (5) must be characterized as *morpho-semantico-syntactic* because the syntactic assignment of these features on the agreement target requires their semantically based assignment on the agreement trigger as well. The group of features under (2–4) are *morpho-semantic* features. The group (1) is purely *morphological*. Note also that the feature CASE shows up in several leaves because it can be assigned both in morphosyntax (through agreement on adjectives) or in morphology (through the assignment of either grammatical or semantic cases on head nouns). In the following Part **??** of this book, dependent marking *state* will be dealt with in more detail since this type occurs in several languages of the geographical area under investigation.

Tree formatting

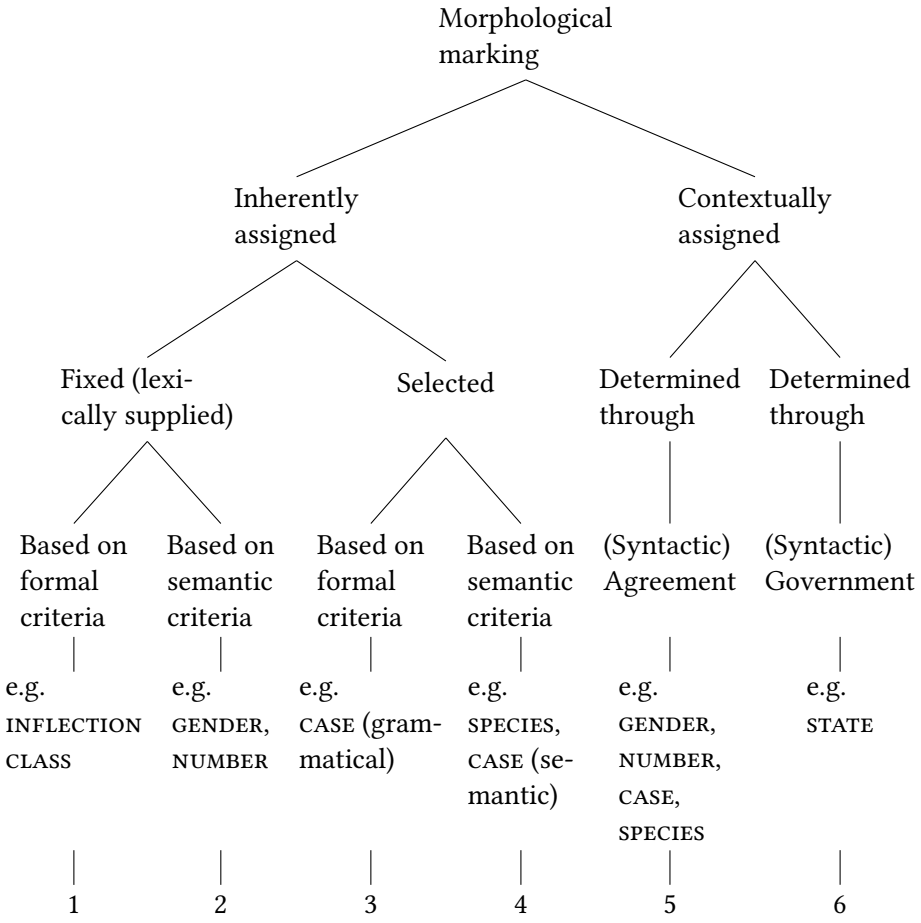


Figure 3.1: An ontology of morphosyntactic features relevant to the present inventory of noun phrase types (**kibort2010a kibort2008a**) and extended with the feature STATE)

Part II

Typology

4 A typology of adjective attribution marking devices

In the present chapter, different types of adjective attribution marking devices attested in natural languages will be described and systematized with a special focus on their typologization according to the morphology of attributive adjectives.

4.1 Typologizing noun phrase structure

The goal of the following chapters is to typologize noun phrases and to present a comprehensive ontology of different syntactic, morphosyntactic, and morphosemantico-syntactic attribution marking devices attested in the languages of northern Eurasia and beyond.

In order to illustrate the different noun phrase types to which these devices belong, data from several languages both within and outside the geographic area of investigation are taken into consideration. The focus, however, will be on constructions and features especially relevant to adjective attribution in the northern Eurasian area.

The term *adjective attribution marking* will be used to refer to a grammatical operation relating an adjectival modifier to its noun head. *Attribution marking device* will be used to subsume both overt and covert grammatical operations which license the syntactic relation of attribution.

The term *noun phrase type* used here denotes the specific syntactic or morphosyntactic structure type of a noun phrase. This term is thus superordinate and belongs to noun phrase structure in general. Since the present study is restricted to a rather small subset of noun phrases, namely noun phrases with adjectival modifiers, the subordinated term *adjective attribution marking device* (instead of *adjective attribution marking type*) will be used to cover all grammatical operations which license the syntactic relation of adjective attribution.

4.1.0.0.4 Attribution marking Minimally, an attribution marking device will simply license the syntactic structure without ranking single constituents, i.e. without licensing any of the constituents as head or dependent. This is the case for the pure syntactic devices *juxtaposition* and *incorporation*.

The syntactic relation of attribution can also be licensed by a device linking the modifying and the modified constituents morphologically to each other, namely in the case of agreement marking. The morphological device of *agreement marking* is characterized by the assignment of an inherent (i.e. true morphological) feature from one constituent to another through morphosyntactic government.

A different instance of "indirect" licensing of attribution is the marking of a semantic relation between the modifier and the modified, as with possessor case (genitive) marking.

It is not at all unusual that the syntactic, morphological, and/or semantic relations between noun phrase constituents are marked simultaneously. If, for instance, an attributiion marker is attached to a modifier which additionally inflects for agreement features, both the syntactic and the morphological relation between the noun phrase constituents are marked. Another example for simultaneously marked syntactic and semantic relations is a noun phrase with a case marked possessor noun (e.g. in genitive case) and a head noun which is additionally marked for dependent-driven agreement (e.g. with a cross-referencing possessive affix).

4.1.0.0.5 Typological parameters Noun phrase types with formally distinct characteristics can be defined according to several parameters. Such parameters are, for example, the order of constituents inside the noun phrase (e.g., attribute-head order, head-attribute order, free order), the attribution marker's locus (e.g., on-head, on-dependent), the marker's behavior relative to the whole phrase (e.g. clitic), its phonological fusion (e.g., free, bound, non-linear), or its position relative to the word host (e.g., pre, post, circum).[1]

Examples for a variety of phonologically, morphologically, syntactically, and semantically distinct types of attribution marking devices will be given in the following chapter. The focus of the ontology presented here is on morphological and morphosyntactic parameters, especially with regard to the absence or presence of additional attribution marking morphemes, as well as to their kind and behavior. An overall picture of the ontology of attribution devices relevant to

[1] These parameters, adapted from Croft's (**croft1995**) typological classification of genitive constructions, are applied for a general typology of noun phrase structure in the noun phrase structure module of AUTOTYP (cf. **AUTOTYP-NP**).

this study is given in Figure 4.2 at the end of §4.

Noun phrase types can also be defined on a polyfunctionality scale with regard to the class of modifying elements: Attributive adjectives and other adnominal modifiers (demonstratives, bare nouns, noun phrases, adpositional phrases, clauses, etc.) may or may not occur in similar noun phrase types. The polyfunctionality parameter even takes the content of certain devices beyond attribution marking into consideration. Since the present study investigates adjective attribution marking, the polyfunctionality of attribution marking devices will be dealt with in less detail (see Chapter 5).

4.1.0.0.6 How many noun phrase types does a language exhibit? Most languages exhibit more than one distinct noun phrase type because different attribute classes may occur as modifiers in noun phrase structures which behave differently in their syntax or morphosyntax. In English, for instance, adjectives and clauses behave syntactically different as modifiers in noun phrases: Whereas attributive clauses are marked by relative pronouns (or particles) (*the dog* ***which is nice***), adjectives are juxtaposed (*the* ***nice*** *dog*). However, since the present book is devoted to the morphosyntax of one single class of adnominal modifiers, namely adjectives, variation in attribution marking devices across different classes of attributed elements is of minor importance.

Nonetheless, attributed elements belonging to one and the same class may also occur in noun phrases which are marked differently: Possessive pronouns in English, for example, can be attributed either by means of juxtaposition (***her*** *dog*) or by using them in a prepositional construction (*the dog* ***of hers***). Even attributive adjectives may occur in two formally distinct noun phrase types. In Turkish, for instance, attributive adjectives are unmarked (***kara*** *kalem* 'black pencil'); in headless noun phrases marked as direct objects, however, adjectives must be nominalized by means of the 3rd person singular possessive suffix (***kara-sını*** 'the black one (=pencil)' [POSS:3SG.ACC]; see also §**??** below).

Prototypically, the use of different devices for licensing one and the same class of attributed elements is not arbitrary but governed by constraints. Nominalization of adjectives in Turkish, for instance, is due to a syntactic subset constraint affecting those phrases in direct object position and without a lexical head noun. In other languages, the occurrence of a given noun phrase type may also be constrained lexically and/or semantically by subsets of either attributes or heads. A well-known example beyond adjective attribution comes from languages in which the choice of possession marking devices is determined semantically by the alienable or inalienable subset of the head noun (i.e. the possessed). Even

other subsets of head nouns are known to constrain the choice of possession marking in some languages, such as kinship terms, (non-) referential nouns, etc.

Similarly, languages may exhibit subset constraints on the semantic class of heads modified by adjectives. The epithet-construction marked with an attributive article in English (or other Germanic languages, cf. *Frederick **the Great**, Friedrich **der Große***; see also §4.5.2.3 below) may serve as an example. In English, this special noun phrase type only occurs if the head noun belongs to the semantic subclass of proper nouns.

Examples of a semantic subset of attributes governing a special attribution marking device are commonly found in languages with contrastive focus marking of adjectives. In Rumanian, for instance, adjective attribution marking is usually characterized by a noun phrase type with head-initial constituent order. A different noun phrase type, formally distinguished by the reversed order of constituents, occurs if the adjective bears contrastive focus (see the Rumanian example (2c) on page 16 above).

Finally, many languages exhibit lexically defined subclasses of adjectives (or other adnominal modifiers) which are sensitive with regard to the required attributive marking. In Albanian, for instance, the members of one adjective class are regularly marked by head-driven agreement whereas the members of another adjective class require an additional agreement marker (see the Albanian example (**??**) on page **??**).

In many languages these lexical subclasses seem marginal and are thus often mentioned merely *en passant* (if at all) in grammatical descriptions. The adjective *pikku* 'little' in Finnish is an example for such a marginal subclass: *pikku* is juxtaposed to the modified noun while other adjectives in Finnish show number and case agreement as a rule (**karlsson1999**). Similarly in German a few adjectives like the colors *lila* 'lavender' and *rosa* 'pink' behave morphosyntactically differently and do not agree with the modified noun. Another example for a marginal subclass of adjectives comes from Itelmen, where attributive adjectives are regularly marked with a special attributive suffix (see the Itelmen example (**??**) on page **??**). Only a few loan adjectives from Russian occur in juxtaposition (**volodin1997**).

These marginal adjective classes are often hard to come across in a rather broad typological survey. It seems to be one limitation of the typological method (i.e. sampling and coding a huge amount of different languages on the basis of qualitatively highly diverse grammatical descriptions) that interesting cases are often missed due to limited knowledge or understanding of the structure of all particular languages. From a diachronic perspective, however, "irregular"

linguistic structures are very important because they often reflect innovative tendencies or archaic features, i.e. features which are due to language change. Marginal noun phrase types should thus be included in typological surveys if they are discovered. Komi-Zyrian

4.2 Syntactic attribution marking: juxtaposition

Juxtaposition can be defined as an unmarked sequence of phrase constituents in which one constituent is syntactically subordinated to the other. It has to be distinguished from *apposition*. The latter term is usually used to denote an appositive construction of two noun phrases, as in *Alma, meine Tochter* 'Alma, my daughter' or *Iva, die jüngere Tochter* 'Iva, the younger daughter' where neither constituent is syntactically subordinated. See also the short discussion in §2.3.3. Juxtaposition is thus characterized by adjacency of noun phrase constituents alone. There is no construction marker present. Consider the following Komi-Zyrian examples where neither agreement markers or any other additional morphemes are present. The attributive adjective in (**??**) is represented by its pure stem form. It does not inflect for any of the categories marked on the head noun.[2]

(1) Komi-Zyrian (Uralic; (**lytkin1966a**))

a. *bur mort*
good person
'good person'

b. *bur mort-jas*
good person-PL
'good people'

Juxtaposition constitutes a very widespread attribution marking device cross-linguistically. Among the northern Eurasian languages, juxtaposition occurs as the default attribution marking device in several families, among others in Mongolic, Turkic and Uralic. Whereas juxtaposition constitutes the default type even in the proto-stages in these language groups, the occurrence of juxtaposition in several other languages results from a relatively recent linguistic change in which the original agreement marking on adjectives was lost.

Defining juxtaposition as a "device" for marking attribution might, however, be questionable. Given the definition that attribution is licensed by the sequence of constituents alone, i.e. that an adnominal modifier and a head noun occur next

[2] Beside NUMBER, these categories include CASE and POSSESSION in Komi-Zyrian.

Ainu (Shizunai)
Kalmyk

to each other in the syntactic structure, juxtaposition resembles a "non-marking" rather than a marking device. In English, for instance, one could also argue that the non-occurrence of the copula *is/are* is relevant to the marking of attribution. In order to use an adjective as predicate in English (*the man **is** good, the men **are** good*), the copula is obligatory. However, word order may be relevant, too. In English, again, juxtaposed attributive adjectives precede the noun as a rule, whereas predicative adjectives follow it.

Word order can in fact be crucial in languages were both adjective attribution and predication are marked simply through adjacency of noun and adjective but with reversed word order, as for example, in Ainu or Kalmyk.

(2) Ainu (Shizunai) (isolate; **refsing1986**)

a. *Attribution: adjective-noun order*
pirka cep
be_good fish
'a fine fish'

b. *Predication: noun-adjective order*
cep pirka
fish be_good
'the fish is fine'

(3) Kalmyk (Mongolic; **jachontova1997**)

a. *Attribution: adjective-noun order*
čyγan časun
white snow
'white snow'

b. *Predication: noun-adjective order*
časun čyγan
snow white
'the snow is white'

The only difference between attribution and predication of adjectives in Ainu[3] and Kalmyk is in word order.

[3] Note that there are no true adjectives in Ainu. Property words are stative verbs in this language (see also §??).

4.3 Covert morphosyntactic construct marking: adjective incorporation

Västerbotten-Swedish

Similarly to juxtaposition, *adjective incorporation* is characterized by adjacency of phrase constituents. There is no additional morpheme present in this type of noun phrase either. The syntactic relation of attribution is, however, marked by a syntactic composition of modifier and head noun. This type can thus be characterized as covertly marked operation.

(4) Västerbotten-Swedish (Indo-European; **larsson1929**)

a. *stor-båt-en*
big-boat-DEF:M.SG
'the big boat'

b. *stor-hus-et*
big-house-DEF:N.SG
'the big house'

Since adjective incorporation in northern North-Germanic dialects is syntactically and semantically distinguishable from derivational compounding it is often referred to as *Adjective-Noun-Incorporation* (for instance by **sandstrom-etal2003** **dahl2007** or **julien2005**).

4.3.0.0.7 Phonological vs. syntactic compounds In Västerbotten-Swedish (as well as in other North-Germanic varieties where adjective-noun compounds occur), accent patterns clearly indicate that adjectives are morpho-phonologically compounded (**dahl2003**). Non-compounded monosyllabic stems, such as *tré*, 'tree', *bǻt* 'boat', *bǻt-er* 'boats', *bǻt-er-na* 'the boats', have an acute accent (marked with ´ in the examples) as a rule and whether or not they are equipped with inflectional affixes. Bisylabic stems, including compounds, by contrast have pitch accent on the stem (marked with ´ ‘ in the examples). Compare *tré-bå̀t-en* 'the wooden boat' or *stór-bå̀t-en* with the noun phrase *bǻt-en mín* 'my boat', where both the noun and the (non-compounded) possessive pronoun have acute accent.

Phonological composition, however, cannot be sufficient evidence for syntactic compounding (i.e. incorporation). Phrase internal phonological or prosodic processes at the juncture of adjectives and nouns (as, e.g., the accent pattern described above) seem to be very common in languages. Such processes can perhaps prove morpho-phonological composition. For the present typology, however, adjective incorporation is defined purely syntactically as a noun phrase

Västerbotten-Swedish

where the attributive adjective occurs obligatorily as a (syntactically) bound morpheme. To prove syntactic boundedness one has to show that the adjective cannot occur unbound. In Västerbotten-Swedish (and other North-Swedish dialects), for instance, the adjective stem cannot occur unbound unless alternative morphosyntactic marking is applied. Using the adjective 'big' in Västerbotten-Swedish in a headless noun phrase results in a construction in which the adjective is marked for agreement and is obligatorily followed by en article serving as a dummy head.[4]

(5) Västerbotten-Swedish (Indo-European; **larsson1929**)

a. *en* *stor* *en*
INDEF:M big(M) ART:INDEF:M.SG

b. *ett* *stor-t ett*
INDEF:N big:N ART:INDEF:N.SG
'a big one'

If evidence for syntactic incorporation cannot be found compounded adjectives can only by described as a special case of juxtaposition. But interestingly, if the described test of syntactic boundedness is applied, then English falls in the category of incorporating languages as a result. In English too, attributive adjectives can only occur bound to a head. This head is either lexical or, similar to Västerbotten-Swedish indefinite noun phrases, an obligatory article as dummy head.[5]

Whether or not English is coded as an incorporating language, adjective incorporation seems to constitute a minor type of attribution marking. Among languages of the northern Eurasian area, however, this type is attested in geographically quite distinct languages: besides the peripheral North-Germanic dialects, it is also found in Adyge and in Chukchi, Kamchatkan and in Eskimo-Aleut languages (see the respective sections of Part III).

4.4 Morpho-semantico-syntactic attribution

[4] This is true, however, only with the indefinite adjective. The definite adjective, by contrast, does not need a dummy head but is unbound (and equipped with the definite marker): *stor-en* [big(M) DEF:M.SG] 'the big one (masculine)', *stor-et* [big(M) DEF:N.SG] 'the big one (neuter)'.

[5] Applying the same test, it turnes out that English incorporates even other modifiers of nouns, such as possessive pronouns: *give me her book – give me her-s.*

marking: agreement

Sorbian, Lower

Agreement (aka *concord*) is a common type of overt attribution marking device. Agreement is commonly understood as a systematic covariance between a semantic or formal property of one element and a formal property of another **steele1978** In other words, agreement can be defined as the spread of semantic or morphological properties across constituents of a syntactic phrase. The agreement properties (or *agreement features*) spread from "trigger constituents"[6] and are formally, i.e. morphologically expressed on "target constituents".

The primary syntactic function of agreement is to relate phrase constituents to each other. Agreement thus serves the formal licensing of dependency in the given phrase. As compared to construct marking, however, the licensing of dependency by means of agreement is more the indirect result of morphological copying of agreement features across phrase constituents.

In principle, agreement features can be triggered by both syntactic heads and syntactic dependents, as will be shown in the following sections. Based on where the agreement features originate, the terms *head-driven* and *dependent-driven agreement*, first proposed by Balthasar Bickel and Johanna Nichols in 2001 (**bickel-etal2007**), will be used in the following.

4.4.1 Head-driven agreement

Typical morphosyntactic agreement features triggered by syntactic heads are GENDER, NUMBER and CASE, as in Lower Sorbian.

(6) Sorbian, Lower (Indo-European; **janas1976**)

a. *dobr-y* *człowjek*
good-SG:M person(M)
'a good person'

b. *dobr-e* *człowjek-y*
good-PL person-PL
'good people'

c. *k* *dobr-emu* *człowjek-oju*
to good-SG:M:DAT person-SG:M:DAT
'to a good person'

[6] In other terms, the trigger of agreement can be called *controller*, cf. **corbett2006**

Sorbian, Lower
Sorbian, Lower
Sorbian, Lower

Note, however, that **kibort2010a** following **corbett2006** does not list CASE as a prototypical agreement feature. In Kibort's and Corbett's view, the matching of a case value on the noun phrase head and its adjectival (or other) modifier(s) does not count as "canonical agreement" but is simultaneously imposed on the noun phrase constituents as the result of government by a syntactic element outside the noun phrase. Consider the Lower Sorbian example (6c) in which both the adjective 'good' and the noun 'person' are marked with the dative case suffix.

The question is whether the case value in such examples is imposed on both noun phrase constituents through government (in example 6c by the preposition *k* 'to') as argued by Corbett and Kibort, or if the dative case on the modifying adjective is imposed by its head by means of agreement, similar to gender and number agreement which are also imposed by the head noun. Adopting Mel'čuk's (**melcuk1993**) dependency view of syntax instead of Corbett's (**corbett2006**) "constituency", the dependent constituent in the adposition phrase is a noun phrase. The dependent constituent in the noun phrase, again, is an adjective phrase (i.e. the attributive adjective) which depends on the noun head of the phrase and inherits its case marking. In this view, the morphosyntactic mechanisms of assigning a head's morphological features to dependent constituents are similar for case and other agreement categories (like gender and number). Consider (6c) 'to a good person' in Lower Sorbian.

(7) Sorbian, Lower (Indo-European; **janas1976**)
$[_{AdP}\ k\ [_{NP}\ dobremu_{agr}\ człowjeku_{gender:number:case}]]$

Another possible agreement feature beside GENDER, NUMBER and CASE is SPECIES, typical values of which are DEFINITE and INDEFINITE. Consider, for instance, the agreement paradigm of adjectives in Icelandic (Table 4.1) in which indefinite and definite forms are distinguished.

Cross-linguistically, head-driven agreement seems to be a wide-spread attribution marking device across the world's language families. The actual morphological appearance of agreement marking, however, is highly diverse across languages and depends on several parameters.

One such parameter concerns the form of the agreement marking morphemes in comparison to the morphemes marking the respective values on the head noun. In fact, adjective agreement paradigms in many languages are different from the respective inflectional paradigms of nouns. This is true, for instance, for Slavic and Germanic languages, as mentioned, but also for other Indo-European languages. In other languages, however, inflectional suffixes might simply reoccur on the modifier, as in Finnish.

Table 4.1: Adjective declension paradigm for Icelandic (Indo-European; **kress1982**) Finnish

		M.SG	F.SG	N.SG	M.PL	F.PL	N.PL
INDEF	NOM	−ur	−Ø	−t	−ir	−ar	−Ø
	ACC	−an	−a	−t	−a	−ar	−Ø
	DAT	−um	−ri	−u	−um	−um	−um
	GEN	−s	−rar	−s	−ra	−ra	−ra
DEF	NOM	−i	−a	−a		−u	
	ACC	−a	−u	−a		−u	
	DAT	−a	−u	−a		−u	
	GEN	−a	−u	−a		−u	

(8) Finnish (Uralic; own knowledge)

a. *iso-t* *talo-t*
large-PL house-PL
'large houses'

b. *iso-i-ssa* *talo-i-ssa*
large-PL-INESS house-PL-INESS
'in large houses'

Adjectives and nouns in Finnish (and in most other Uralic languages) differ in syntactic function rather than in morphological properties. Consequently, adjectives and nouns in Finnish exhibit similar inflectional paradigms. Probably, such a weak distinction between adjectival and nominal inflections was also true for proto-Indo-European (cf. **comrie1998** **kuriaki2007**). But the declensions of both adjectives and nouns in Indo-European languages have undergone radical changes and have become clearly distinct from each other. This is evident, e.g., in the Lower Sorbian example (6) on page 33 where the adjective suffix *-emu* and the noun suffix *-oju* both mark the dative masculine singular.

Head-driven agreement marking also deviates across languages in respect to the inventory of morphological categories involved. Many languages exhibit head-driven agreement paradigms which exclude certain inherent or assigned morphological categories of the head noun, as in Finnish, where nouns inflect for NUMBER, CASE and POSSESSION. The latter feature, however, never spreads through the noun phrase.

Finnish
Danish
Chechen

(9) Finnish (Uralic; own knowledge)

a. *iso talo-ni*
large house-POSS:1SG
'my large house'

b. **iso-ni talo-ni*
large-POSS:1SG house-POSS:1SG

Finally, agreement paradigms can be "defect" in the sense that certain agreement categories do not show up on all members of the paradigm. In Danish, for example, gender as an agreement feature is marked on the attributive adjective only in indefinite noun phrases. In noun phrases marked for definite species, the attributive adjective is marked with an invariable definite agreement suffix. Consider (10) and Table **??** with the respective paradigm in §**??**.

(10) Danish (Indo-European; own knowledge)

a. *en* ***stor*** *mand*
INDEF.COM big.UTR man(UTR)
'a tall man'

b. *ett* ***stor-t*** *hus*
INDEF.N big-N house(N)
'a large house'

c. *den* ***stor-e*** *mand*
DEF.COM big-DEF man(UTR)
'the tall man'

d. *det* ***stor-e*** *hus*
DEF.N big-DEF house(N)
'the large house'

An extreme case of a defective agreement paradigm is found in Chechen where adjectives only partially agree with the head noun and show only one single case distinction between nominative versus all other cases, as in the (incomplete) paradigm (11).[7]

(11) Chechen (Nakh-Daghestanian; **nichols1994a**)

a. *dika^{n} stag^{3} 'good person'* *NOM:SG*

b. *dikaču stega^{n}* *GEN:SG*

[7] A similar defective agreement paradigm with only one case distinction is found in Ingush, cf. §**??**.

c.	*dikaču stagana*	*DAT:SG*	Oroch
d.	*dikaču staga*	*ERG:SG*	Oroch
e.	*dikaču stagie*	*ALL:SG*	
f.	*dika^{n} na:x*	*NOM:PL*	
g.	*dikaču ne:xa^{n}*	*GEN:PL*	

4.4.2 Dependent-driven agreement

In many languages spoken inside and outside the northern Eurasian area, head-driven agreement is attested as a device for licensing attributive modification. The reverse agreement type, *dependent-driven agreement*, is also wide-spread among the world's languages. Among the languages of my sample, however, dependent-driven agreement marking is attested only as a device for the licensing of (possessor) noun attributes. An example of a language with dependent-driven agreement marking in possessive noun phrases is Oroch.

(12) Oroch (Tungusic; **malchukov2000**) *nia d'uu-ni*
man house-POSS:3SG
'a man's house'

The possessed noun 'house' in example (12) obligatorily agrees with the 3SG possessor 'man'. This type of dependent-driven agreement is usually called *possessor agreement*.[8]

4.4.2.1 Modifier-headed possessor agreement

The term *modifier-headed possessor agreement* is derived from *modifier-headed agreement* introduced in **AUTOTYP-NP** It is a subtype of dependent-driven agreement characterized by reverse semantic and syntactic dependency relations between attribute and head.

Structurally similar to example (12), Oroch also exhibits dependent-driven agreement marking by means of possessive affixes on attributive adjectives.

(13) Oroch (Tungusic; **malchukov2000**)
a. *nia aja-ni*
man good-POSS:3SG
'a GOOD man'

[8] Another commonly used term is *cross reference marking*.

Saliba
Saliba

b. *nia-sa aja-ti*
man-PL good-POSS:3PL
'GOOD men'

In the Oroch example, the semantic head of the noun phrase 'man' is syntactically "degraded" to the (dependent) possessor function, and the semantic dependent is "upgraded" to the function of the syntactic head of the phrase, i.e. the possessed. According to **malchukov2000** the expression still has an attributive reading: 'a man, a property of whom is "to be good"', rather than a possessive one: *"a man's goodness". Thus, the semantic attribute is rendered as the head (i.e. the possessed) and the semantic head of the possessive noun phrase takes the slot of the dependent (i.e. the possessor).

Whereas modifier-headed possessive agreement constitutes a marked structure in Oroch, it can be the universal type of attributive marking on adjectives in other languages. This kind of adjective attribution marking device is not very common in the northern Eurasian area under investigation, but it is pervasive, for instance, in Oceanic languages (cf. **ross1998**). In Saliba, for example, attributive adjectives as a rule are marked by means of 3rd person possessive suffixes.

(14) Saliba (Austronesian; **mosel1994**)

a. *sine natu-na*
woman child-POSS:3SG
'a woman's child / the child of the woman'

b. *sine-o natu-di*
woman-PL child-POSS:3PL
'women's children / the children of the women'

In Saliba, possessor nouns are licensed as modifiers in a noun phrase by means of (dependent-driven) possessor agreement on the head noun. Similar to the marked noun phrase in Oroch (13), attributive adjectives are marked by means of modifier-headed possessor agreement.

(15) Saliba (Austronesian; **mosel1994**)

a. *mwaedo gagili-na*
eel small-POSS:3SG
'a small eel'

b. *mwaedo gagili-di*
eel small-POSS:3PL
'small eels'

The adjectival attribute 'small' in example (15) occurs in a possessive-like construction (similar to 14) where the adjective takes the slot of the possessed and is subsequently marked with a possessive agreement suffix.[9] Unlike in Oroch, however, modifier-headed possessor agreement is the default type of attributive connection of adjectives in Saliba.

4.5 Overt morphosyntactic construct marking: attributive state marking

Due to a lack of better terminology the feature STATE was earlier defined as assigned through *syntactic government* (in §3.2). Unlike the common notion of *government*, which requires a trigger inside the phrase, true syntactic government considered in this study has no other trigger than the syntactic construction as such.

In order to avoid the misleading term *government*, all overtly marked attribution devices with the exclusive function of licensing the syntactic relation between constituents of a noun phrase are defined here as *attributive state marking*. "Overtly marked" means that (at least one) additional attribution marking morpheme is present in the noun phrase.

Attributive state is adopted from "Construct state" or "Status constructus" which are commonly used in syntactic descriptions of languages exhibiting head-marking state (e.g. Persian). Since construct state marking morphemes may occur on different loci inside the noun phrase, *attributive state* will be used as superordinate term, subsuming the subtypes with the following loci of their respective attributive markers:[10]

- on-head (construct)
- on-dependent (anti-construct)
- neither on-head nor on-dependent (floating construct)

[9] An alternative account of noun phrase structure in Saliba could claim that the verbal attribute is marked by head-driven agreement, analyzing the suffixes *-na* and *-di* as singular and plural markers, respectively. This analysis is obviously underlying the descriptions of Saliba (e.g. **mosel1994** **margetts1999**), which leave the homophony of *-na* POSS:3SG and *-di* POSS:3PL with *-na* SG and *-di* PL undiscussed .

[10] Other logically possible loci of attributive state markers would result from simultaneous marking on head- and/or on dependent+floating. I am, however, not aware of any language exhibiting such noun phrase types.

Persian
Kirmanji

- simultaneously on-head and on-dependent (double construct)

Among the northern Eurasian languages considered in the present study, only the first two types of attributive state marking, i.e. head-marking state and dependent marking state, are attested as devices for licensing attributive adjectives. These two types are dealt with in more detail below in §4.5.1 and §4.5.2.

4.5.1 Head-marking attributive state

The attributive construction in Persian, commonly known as *Ezafe* (or *Izafe*), illustrates a typical case of head-marked attributive state.

(16) Persian (Indo-European; **mahootian1997**)
xane-ye *bozorg*
house-ATTR big
'a large house'

The only function of the attributive suffix *-(y)e*[11] on the noun 'house' is to show that "I am a noun phrase and I have a dependent."[12] The traditional term for the morphological value given by the head-marking attribution device in Persian is *construct state* (or *status constructus*). What is meant hereby is that the noun displays different "states" depending on the presence of a modifier in the noun phrase.

Obligatory attribution marking by means of an Ezafe-construction is also characteristic for other Iranian languages. In Kurmanji, a variety of Kurdish spoken in the northern Eurasian area, the Ezafe-formative is not an invariable suffix – unlike the cognate suffix *-(y)e* in Persian – but also indicates morphological values of NUMBER (SG/PL), GENDER (M/F) and SPECIES (DEF/INDEF). Consider example (17) and the paradigm in Table 4.2.

(17) Kirmanji (Indo-European; **ortmann2002b**)
a. *kur-ê* *mezin*
boy-ATTR:DEF.M.SG big
'the tall boy'
b. *keç-a* *baç*
girl-ATTR:DEF.F.SG nice
'the nice girl'

[11] The allomorph *-e* appears ofter consonants.

[12] The attributive construct state marking in Persian is polyfunctional in the sense that its function is not restricted to the licensing of adjectives as modifier in a noun phrase, but also of noun attributes, adpositional phrases and infinitives.

c. *kur-ên / keç-ên baç* Saamic, Kildin
boy-ATTR:DEF.PL girl-ATTR:DEF.PL nice
'the nice boys / girls'

Table 4.2: Paradigm of the Ezafe in KURMANJI (**schroder2002**)

	M.SG	F.SG	PL
DEF	-(y)ê	-(y)a	-(y)ên
INDEF	-î	-e	

Note that the values of true morphological features (NUMBER, GENDER, SPECIES) of the noun are combined with the morphosyntactic feature ATTRIBUTIVE in the differentiated forms of the Ezafe in Kirmanji. But agreement is not involved here because gender, number and species marking is not triggered within the noun phrase but is inherited to the head noun morpho-semantically.

4.5.2 Dependent marking attributive state

4.5.2.1 Anti-construct state

In some languages there is an attributive construction corresponding to the Iranian Ezafe, which however does not mark the head but the adjectival dependent for "state" (i.e., indicating the availability of a head in the present noun phrase). This type of marking occurs, for instance in Saamic languages.

(18) Saamic, Kildin (Uralic; own knowledge)

a. *Predicative state Tedt pērrht lī ēll.*
DEM house COP high
'This house is high.'

b. *Attributive state*

i. *Tedt lī ēl'l'-**es'** pērrht.*
DEM COP high-ATTR house
'This is a high house.'

ii. *Tegk liev ēl'l'-**es'** pērht.*
DEM COP high-ATTR house\PL
'These are high houses.'

Whereas the predicatively used adjective 'high' is represented by its pure stem form (18a), it is marked with the attributive suffix *-es'* if used as modifier (18b-i+18b-ii). Attributive marking on adjectives in Kildin and other Saamic languages is highly irregular due to the strong tendency to merge predicative and attributive adjective forms. Other adjective marking devices also occur. The default type in most Saamic languages, however, is that attributive adjectives exhibit an attributive inflection (**riesler2006b** see also below §**??**).

The attribution marker in Saamic is invariable, i.e. the adjective does not show agreement with its head noun. The host of the Saamic attributive suffix is the adjective. Its only function is to specify the syntactic relation between head noun and adjectival modifier ("my host is dependent in the present syntactic structure"). Since the construction in Saamic constitutes dependent marking opposite to the Persian construct state, it can be labeled *anti-construct*.[13]

Anti-construct state marking seems not uncommon cross-linguistically, even if Saamic and the Iranian language Northern Talysh (cf. §**??**) provide the only examples of European languages with anti-construct state marking on adjectives. Note that typological descriptions and grammars use quite different terms for anti-construct state markers, such as *attributive particle*, *relator*, *associative marker*, *linker*, etc. If anti-construct marks the attribution of possessor nouns (besides adjectives) it is also often called *attributive case* or *genitive*.

4.5.2.1.1 Possessive case marking From a purely syntactic point of view, possessive case marking is similar to anti-construct state marking. Both are syntactically governed dependent marking devices. In fact, anti-construct state marking of adjectives is sometimes described as "genitive" if the device is polyfunctional and marks possessor nouns as well.[14] Rather than extending the terminological domain of possessive case marking to adnominal modifiers beyond noun possessors, the term *possessive case* (or *possessor case*) will be used here only for describing a special subtype of anti-construct state. Whereas the latter is a purely morphosyntactic device, possessive case additionally specifies a semantic relation (i.e. possession).

[13] The term was introduced during Bickel's and Nichols' earlier work on the AUTOTYP Noun Phrase Structure Database, cf. **bickel-etal2002 AUTOTYP-NP**

[14] Even other construct marking devices, such as the linker in Tagalog (33) or the construct state marker in Persian (16) are often described as "genitives" because they mark possession. Unlike prototypical genitives, however, the construct markers in Tagalog and Persian do not constitute dependent marking devices.

4.5.2.2 Anti-construct state agreement marking

Russian

Construct state markers such as the linker in Tagalog, the head-marking construct state marker *-(y)e* in Persian, or the dependent marking anti-construct state marker *-es'* in Kildin Saami are proper construct state markers in the sense that they are exclusively used as licensee of an attributive syntactic relation between modifying and modified constituents in the noun phrase. The respective formatives thus have morphologically unalterable shapes.

In other languages, however, certain adnominal modifiers marked for anti-construct state may additionally be the target of either head- or dependent-driven agreement. Such combined agreement and construct marking devices should consequently be characterized as simultaneously marking the syntactic and the morphological relation between the noun modifier and the modified noun.

This subtype of anti-construct state marking, characterized by (adjectival or other) adnominal modifiers being marked simultaneously for anti-construct state and for head-driven agreement, will be labelled *anti-construct state agreement marking* in the following.[15]

A typical example of a language with anti-construct state agreement marking is Russian.

(19) Russian (Indo-European; own knowledge)

a. *Anti-construct state agreement*

i. *vysok-**ij*** *dom*
high-ATTR:M.NOM house(M)
'(a/the) high house'

ii. *vysok-**aja*** *bašn'a*
high-ATTR:F.NOM tower(F)
'(a/the) high tower'

b. *Predicative agreement*

i. *ėtot* *dom* *vysok*
DEM:M house(M) high:M
'this house is high'

ii. *ėta* *bašn'a* *vysoka*
DEM:F tower(F) high:F
'this house is high'

[15] The extended label ***head-driven** anti-construct state agreement marking* seems obsolete because the agreement is self-evidently triggered by the head noun in this type.

In Russian, attributive as well as predicative adjectives show agreement in GENDER and NUMBER. Attributive adjectives agree additionally in CASE. The agreement suffixes of the attributive and predicative paradigms, however, have different shapes; consider Table **??**.

Traditionally, the two inflection paradigms of the adjective in Russian have been contrasted to each other as "short" and "long" forms. These terms, however, describe the form rather than the function of the different agreement inflections and are thus less useful for the classification of the Russian noun phrase type from a morphosyntactic typological perspective. The "long" adjectives of Russian do not simply belong to a different declension paradigm as compared to their "short" counterparts. The formal distinction between the two adjective declensions is connected to attribution marking. Whereas the predicative ("short") forms show "pure" agreement, the agreement suffixes on attributive adjectives mark agreement and the attributive state of the adjective simultaneously.

Historically, the attributive adjective inflection consists of two morphemes: a pronominal stem plus the original "short" agreement suffix.[16] Synchronically, the attributive adjective suffixes in Russian are thus best analyzed as portmanteau suffixes marking anti-construct and head-driven agreement simultaneously.

One could argue against the analysis of the "long" adjective declension in Russian as attributive state marking saying that "long form adjectives" also occur in predicative position. In fact, the semantic difference between the use of "short" versus "long" forms in adjective predication in Russian could be described as an opposition between temporal and permanent properties denoted by the adjective. Nonetheless, the marking of the predicative adjective is rather irrelevant here. What is crucial, however, is the use of the "long" forms, which occur in attributive position as a rule. The "short" (i.e. predicative) form cannot occur in attributive position.

Furthermore, it could even be argued that "long" form adjectives in predicative position are instances of adjective attribution marking rather than of adjective predication. This is the case if one analyses the "long form adjectives" as headless noun phrases in an appositive construction, as the "long" predicative form in (20b) denoting a permanent property apposed to the "short" predicative form in (20a) denoting a temporal property.[17]

[16] In the forms for nominative (cf. Table **??**) the two morphemes for ATTR and GENDER/NUMBER/CASE are still separable. In the remaining cases, however, they are merged into one portmanteau suffix.

[17] Russian examples of morphologically differentiated predicative adjectives also often reflect an opposition in the subject's denotative status. The "short" form is used for denoting reference to a class of objects: *krasavicy* ***kaprizn-y*** [capricious-PRED:AGR] 'beautiful women are

Russian
Latvian

(20) Russian (Indo-European; own knowledge)

a. *"short" predicative adjective*
on bolen
3SG ill:PRED:M
'he is ill'

b. *"long" predicative adjective*
on bol'nyj
3SG ill:ATTR:M
'he is a sick one (i.e. he is mentally sick)'

The origin of anti-construct state agreement marking in Russian is dealt with in Chapter **??**. It is worth mentioning that remnants of an Old Slavic anti-construct adjective inflection are found in other modern Slavic languages as well, especially in the South-Slavic languages Slovenian and Serbian where the "long" adjective forms occur in definite noun phrases (cf. §**??**).

Similar to South-Slavic but much more regular is the occurrence of a cognate anti-construct adjective inflection in the Baltic languages Latvian and Lithuanian.

(21) Latvian (Indo-European; **dahl2007**)

a. *liel-a māja*
big-F.NOM.SG house(F)
'a large house'

b. *liel-ā māja*
big-ATTR:F.NOM.SG house(F)
'the large house'

Unlike in Russian where attributive adjectives are marked with the anti-construct state agreement suffixes as a rule, the use of the cognate attributive forms in the Baltic languages is usually described as depending on the referential status of the head noun. Whereas the "short form" agreement suffix is used with adjectives modifying indefinite nouns (21a), the attributive adjective in definite noun phrases is obligatorily marked with the "long form" agreement suffix (21b).

The anti-construct state agreement marking suffixes in the Baltic languages is often described as a definiteness marker. Note, however, that the definite noun

capricious'), the "long" form is used for denoting reference to an individual: *oni* ***kaprizn-ye*** [capricious-ATTR:AGR] 'they are capricious' (or 'they are (the) capricious ones', e.g. two sisters known from the discourse) (cf. **mendoza2004**).

Latvian

never exhibits definite marking itself. If no attributive adjective is present the definite noun remains unmarked. The analysis of the "long form" agreement suffix in Baltic as definite marker would thus presuppose the assumption that the definite marker is selective and shows up only on attributive adjectives.

Markers which are selective according to their host's parts-of-speech membership are indeed attested.[18] The Latvian and Lithuanian examples, however, could be compared to selective marking in other languages only if one assumes a zero-allomorph of the definiteness marker attaching to non-modified definite nouns.

(22) Latvian (Indo-European; **dahl2007**)

a. *māja*
house
'a house'

b. *māja-?Ø*
house-DEF
'the house'

c. *liel-ā* *māja*
big-DEF:F.NOM.SG house(F)
'the large house'

melcuk1998 introduced the term *displaced category* (Russian *smeščennaja kategorija*) for the type of marking found in Baltic. It has also been argued by Dahl (**dahl2003** see also **dahl2007**) that definite noun phrases often show special behavior in languages depending on whether or not they exhibit attributive adjectives (or other modifiers).[19]

An alternative analysis is preferred here: Since the "long form" agreement suffix only attaches to attributive adjectives, the formative could well be analyzed as an anti-construct state agreement marker (similar to Russian) which is, however, restricted to occurring in semantically definite noun phrases.

[18] Consider, e.g., the two allomorphs of the definite marker in Danish *hus-**et*** [house-DEF.N] 'the house', ***det** store hus* [DEF.N big.DEF.N house] 'the large house'. The suffix *-et* DEF.N. attaches to bare nouns, whereas the free form *det* DEF.N attaches to noun phrases with adjective modifiers, cf. also Table **??**.

[19] **dahl2003** compares the "long form" adjectives in the Baltic languages with attributive articles in Romance languages (such as in Latin *Babylon illa magna*) and Yiddish, among others. A structural and even historical connection is indeed plausible, as will be shown in Part **??** of this study, especially in §**??**.

Several examples of languages are attested where the occurrence of different noun phrase types is restricted to certain subsets of noun phrase constituents. In the case of the Latvian example given above (and similar to Lithuanian) attributive adjectives are marked differently depending on the referential status of the whole phrase. The choice between the head-driven agreement versus the anti-construct state agreement type would thus be constrained by the semantically defined subsets of the noun head (i.e. indefinite versus definite).

As a consequence of the suggested analysis of the "long form" agreement suffixes in Baltic as anti-construct state agreement markers, Latvian and Lithuanian could be described as lacking definiteness as morphological category. In fact, several authors have questioned the existence of morphologized definite marking at least in Lithuanian, where the occurrence of the anti-construct state agreement suffix is clearly not restricted to definite noun phrases (cf. **wissemann1958** cit. **kramsky1972**). **trost1966** argues that permanent versus non-permanent properties are marked rather than definite versus indefinite, for example (Lithuanian) *aukštoji mokyla* 'college (lit. 'high school')'.[20]

In Chapter **??**, diachronic arguments will be presented in favor of the assumption that a morphological feature SPECIES (with the values DEFINITE / INDEFINITE) was not present in Baltic languages, at least until the most recent stages in their language history. The anti-construct state agreement inflection is clearly older than the morphologization of definiteness in Baltic (and similarly in certain Slavic languages). In older stages of Baltic (and Slavic) the "long" adjective inflection was connected to attributive rather than to definiteness marking. To a certain extent, this holds true for the modern Baltic languages Latvian and Lithuanian.

Thus, in the ontology presented here anti-construct state agreement marking in Baltic belongs to the same noun phrase type as the one described for Russian (cf. example 19 on page 43). This analysis seems justified regardless of the question as to whether the device constitutes the default type of adjective attribution marking (as in Russian) or is restricted to a given semantically restricted subset of the head noun (as in Latvian and Lithuanian).

Also in German (similar to the other West-Germanic languages, except English), attributive and predicative adjectives are morphosyntactically differentiated. Whereas attributive adjectives show head-driven agreement, predicative adjectives are used in an invariable form. Given the definition of dependent marking attributive state which was applied here (see also Chapter 3), German

[20] For Latvian, however, **trost1966** accepts the analyses of the "long" suffix as definite marker because it occurs regularly after possessive pronouns.

Endo

thus exhibits a similar type of obligatory anti-construct state agreement marking as Russian. Note, however, that the adjective inflection suffixes in German are merged to a relatively high degree: Only the five single forms *-e, -en, -em, -er, -es* are formally distinguished.

What is even more interesting in German is the fact that the agreement feature SPECIES exhibits a third value for which a grammatical label is hard to find. Whereas indefinite agreement shows up on adjectives in semantically indefinite noun phrases (formally marked by the indefinite marker *ein* in Table 4.3) and definite agreement on adjectives occurs in semantically definite noun phrases (formally marked by the definite marker *der* in Table 4.3), the "third species" agreement forms show up in semantically indefinite or definite noun phrases marked, for instance, by possessive pronouns and the indefinite pronoun *kein* 'no(t any)'. Whereas the "third species" agreement forms are similar to the indefinite forms in singular, they are similar to the definite forms in plural. Accordingly, three species values thus have to be distinguished in the morphological paradigm.

It is worth mentioning that adjectives which are simultaneously marked for attributive state (i.e. anti-construct) and head-driven agreement are also attested in languages outside the northern Eurasian area. Similar to Russian, adjectives in Endo, a Nilotic language of Kenya, require different agreement suffixes depending on their use as modifiers of a noun or as predicates.

apply better spacing between columns

(23) Endo (Nilotic; **zwarts2003**)

a. *karaam inyeentee*
good(SG) 3SG
'S/he is good.'

b. *laakwa nyaa karaam*
child ATTR:SG good(SG)
'a good child'

c. *karaam-a akwaaneek*
good-PRED:PL 3PL
'They are good.'

d. *piich chaa karaam-een*
people ATTR:PL good-ATTR:PL
'good people'

The example illustrates that adjectives in Endo show agreement in number. The singular is unmarked and the plural is marked by the suffix *-a* for predicative

Table 4.3: Agreement paradigm for the GERMAN adjective 'good' ('good man' M, 'good woman' F, 'good child' N, 'good people' PL)

		M.SG			F.SG			N.SG			pl	
INDEF	NOM	(ein)	gut-**er**	(Mann)	(ein-e)	gut-**e**	(Frau)	(ein)	gut-**es**	(Kind)	gut-**e**	(Leute)
	GEN	(ein-es)	gut-**en**	(Mannes)	(ein-er)	gut-**en**	(Frau)	(ein-es)	gut-**en**	(Kind-es)	gut-**er**	(Leute)
	DAT	(ein-em)	gut-**en**	(Mann)	(ein-er)	gut-**en**	(Frau)	(ein-em)	gut-**en**	(Kind)	gut-**en**	(Leuten)
	ACC	(ein-en)	gut-**en**	(Mann)	(ein-e)	gut-**e**	(Frau)	(ein)	gut-**es**	(Kind)	gut-**e**	(Leute)
DEF	NOM	(der)	gut-**e**	(Mann)	(die)	gut-**e**	(Frau)	(das)	gut-**e**	(Kind) (die)	gut-**en**	(Leute)
	GEN	(des)	gut-**en**	(Mannes)	(der)	gut-**en**	(Frau)	(des)	gut-**en**	(Kind-es) (der)	gut-**en**	(Leute)
	DAT	(dem)	gut-**en**	(Mann)	(der)	gut-**en**	(Frau)	(dem)	gut-**en**	(Kind) (den)	gut-**en**	(Leuten)
	ACC	(den)	gut-**en**	(Mann)	(die)	gut-**e**	(Frau)	(das)	gut-**e**	(Kind) (die)	gut-**en**	(Leute)
IN/DEF	NOM	(mein)	gut-**er**	(Mann)	(meine)	gut-**e**	(Frau)	(mein)	gut-**es**	(Kind) (meine)	gut-**en**	(Leute)
	GEN	(meines)	gut-**en**	(Mannes)	(meiner)	gut-**en**	(Frau)	(meines)	gut-**en**	(Kind-es) (meiner)	gut-**en**	(Leute)
	DAT	(meinem)	gut-**en**	(Mann)	(meiner)	gut-**en**	(Frau)	(meinem)	gut-**en**	(Kind) (meinen)	gut-**en**	(Leuten)
	ACC	(meinen)	gut-**en**	(Mann)	(meine)	gut-**e**	(Frau)	(mein)	gut-**es**	(Kind) (meine)	gut-**en**	(Leute)

Mandarin Chinese

adjectives and by *-een* for attributive adjectives.[21]

4.5.2.3 Attributive nominalization

Nominalization is often understood very broadly as a word-class changing morphological operation deriving nouns from other syntactic classes. This definition stresses the lexical-semantic side of nominalization. But the term is sometimes also used for a syntactic operation in which a verbal (single or complex) constituent, like a verb, a verb phrase, a sentence, or a portion of a sentence (including a verb) is converted into a nominal (single or complex) constituent (**li-etal1981**). In this latter sense, nominalization is a means of licensing nominal constituency.

Mandarin Chinese illustrates a language in which syntactic nominalization is a highly polyfunctional device for the licensing of different modifying phrase constituents (cf. **li-etal1981** see also example 1 in Chapter 5). Adjectives in Mandarin are used in attributive position (24a), in predicative position (24b) and when used as adverbial modifiers (24c).

(24) Mandarin Chinese (Sinotibetan; **li-etal1981**)

a. *Adjectival attribute*
[$_{NP}$ *xīn* ***de***] *shū*
new NMLZ book
'new book'

b. *Adjectival predicate*
wǒ-de *shū* *shì* [$_{NP}$ *xīn* ***de***]
1SG-NMLZ book COP new NMLZ
'My book is new (i.e. a new one).'

c. *Adjectival adverb*
wǒ [$_{NP}$ *yánli-**de***] *zébèi* *tā* *le*
1SG stern-NMLZ reproach 3SG CRS
'I sternly (i.e. as a stern one) reproached him/her.'

Interestingly, nominal constituents can also be nominalized, i.e. they can be syntactically licensed as constituents in larger syntactic units. In some languages, such syntactic licensing is obligatory for certain types of nominals. The respective markers (i.e. nominalizers of nominals) are labelled with quite different

[21] Unlike in Russian, however, there is a second attributive marker present in Endo, an attributive article *nyaa* ATTR:SG, *chaa* ATTR:PL. The noun phrase type would thus better be characterized as a combination of attributive article+anti-construct state agreement, hence "double agreement".

terms, such as, for instance, "articles", "noun phrase articles" or "noun (phrase) markers" (cf., e.g., **dryer2007** **rijkhoff2002**). Prototypical examples of such markers come from Oceanic languages where noun phrases contain an obligatory nominalizer deriving from a demonstrative. German

Due to lack of a conventionalized terminological distinction, "nominalization" is here used for denoting the purely syntactic operation by which a noun or noun phrase is marked as a syntactic constituent by making it syntactically more complex, i.e. by projecting a full noun phrase. This use of the term *nominalization* is also consistent with the fact that "nominal" is most often used as a homonym for "noun phrase" rather than for "noun". "Substantivation", on the other hand, will be used for the purely morpho-semantic (derivational) process yielding a noun (substantive). Whereas substantivation belongs to the spheres of morpho-semantics and lexicon, nominalization belongs to syntax: Nominalizers function exclusively for the licensing of noun phrases as constituents in larger syntactic units.

Attributive nominalization has already been discussed as "appositive modification" in §2.3.3. Attributive nominalization is a special subtype of dependent marking construct state. Similar to the latter, attributive nominalization represents a covert dependent-marking morphosyntactic device and is triggered either by purely syntactic government (as, e.g., anti-construct state marking in Kildin Saami, see §4.5.2) or by syntactic government in combination with head-driven agreement (as, e.g., anti-construct state agreement marking in Russian, see §4.5.2.2). The special distinguishing characteristic of attributive nominalization lies in the syntactic structure: Whereas true anti-construct state markers attach directly to the dependent constituent (as, e.g., the respective inflectional suffixes in Kildin Saami or Russian), attributive nominalizers attach to an intermediate dependent phrasal constituent between the head noun and the modifier.

Epithet-constructions with attributive articles in Germanic languages illustrate a prototypical case of attributive nominalization by means of an article.[22]

(25) German (Indo-European; own knowledge)
Friedrich der Große *'Frederick the Great'*

Following **himmelmann1997** the syntactic structure of this example can be described as follows:

(26) $[_{NP}$ *Friedrich* $[_{NP'}\ _{ART}$*der* $_{A}$*Große* $]]$

[22] The examples are from **himmelmann1997** Note that attributive nominalization in German is restricted to noun phrases with proper names as heads. This restriction is, however, irrelevant to the following argumentation.

German

The intermediate phrasal constituent between the noun phrase (NP) and the adjective is labeled as NP', leaving open the rather theoretical question about what constitutes the syntactic head of this phrasal projection.[23]

Note that the attributive marker *der* in example 25 is homophone with the definite marker *der* but clearly has a different function in this construction. For instance, the attributive marker *der* cannot be exchanged with a possessive or a demonstrative pronoun and is thus not a marker of definiteness. The proper noun *Friedrich*, on the other hand, can be further determined by means of a demonstrative (***jener*** *Friedrich der Große* 'that Frederick the Great') or a possessive pronoun (***unser*** *Friedrich der Große* 'our Frederick the Great'). In fact, (in-)definiteness marking of the whole noun phrase does not affect the attributive nominalizer, consider the following example:

(27) German (Indo-European; own knowledge)

a. *Irgendein* [*Friedrich der Große*]$_{INDEF.NOM}$ *soll das gesagt haben.*

b. *Dieser* [*Friedrich der Große*]$_{DEF.NOM}$ *soll das gesagt haben.*

c. *Ich sehe mir irgendeinen* [*Friedrich den Großen*]$_{INDEF.ACC}$ *an.*

d. *Ich sehe mir diesen* [*Friedrich den Großen*]$_{DEF.ACC}$ *an.*

The attributive adjective forms a complex constituent together with the article. This complex constituent is subordinated to the noun phrase head (i.e. the proper name *Friedrich*) whom it modifies. Agreement in gender/number/case is triggered (through agreement) by the head noun (*Friedrich*)The agreement pattern in the German epithet-construction also show that the nominalizer *der* has not only to be distinguished from the homophone definite marker but also from the relativizer *der*. Consider the following examples (cf. also **himmelmann1997**). According to Lehmann (**lehmann1984** cf. also **himmelmann1997**) true relative pronouns represent the syntactic head for the predicate of the embedded clause. The syntactic function of the relative pronoun is determined by the predicate, but it is independent from the syntactic function of the head noun. Consequently, the relativizer *der* (similar to the adjective *groß*) in example (**??**) agrees only in gender and number with the head noun *Friedrich*. Case is alloted according to the function of *der* as argument in the embedded clause. This is different from the syntactic function of the attributive nominalizer *der*. The nominalizer does agree in case with the head noun. The article's syntactic function is thus dependent of the head noun's function in the superordinate construction.

[23] "Article phrase" (similar to "Determiner phrase" in X-bar syntax) would imply the nominalizer (in this case the article *der*) is the head.

4.5.2.4 Attributive articles

Attributive nominalizers similar to *der* in German epithet-constructions will be labeled *attributive articles* in the following. Attributive articles are similar to anti-construct state agreement markers in that they mark the syntactic relation of attribution and agreement simultaneously. Prototypically, attributive articles are grammatical words and hence syntactic constituents on their own. In the case of the German attributive article *der*, the constituency of the marker becomes evident in the fact that both the adjective and the article are the target of head-driven agreement.

Even though "article" is often used for many different types of grammatical markers, this term (<Latin *artus/articulus* 'joint, small connecting part') originally referred to the metaphor of a joint between the constituents in a noun phrase, hence a true attribution marker. Interestingly, **dryer1989a** and **rijkhoff2002** distinguish two types of "articles": (1) words indicating (in-) definiteness (or some related discourse notion) and (2) words serving as a noun phrase marker "in the sense that noun phrases in that language [...] typically occur with one of the words in question" (**rijkhoff2002**). Attributive articles could nicely be subsumed under type (2) "Noun phrase marker" if the definition would be extended: "a marker which occurs with noun phrases **and/or phrasal dependent constituents of noun phrases**".

The term *attributive article* used here matches Himmelmann's (**himmelmann1997**) *Gelenkartikel* 'linking article', which in turn is borrowed from Gamillscheg's (**gamillscheg1937**) description of the "linking function" (*Gelenksfunktion*) of articles in different Indo-European languages.[24]

Even though the use of the term *article* by Indo-Europeanists is often applied in grammatical descriptions of different languages and even in theoretical linguistic studies, the present study prefers to use *article* only for denoting an attributive marker. On the basis of examples from Greek (with the so-called repeated article) and from Latin (with the so-called linking demonstrative), **gamillscheg1937** characterizes the attributive article as exhibiting "a disjunctive and linking function simultaneously"[25] by marking the adjective as "physically independent."[26] The articles *ille* in Latin and *tó* in Greek thus have different functions than the

[24] In Himmelmann's **himmelmann1997** terminology, however, the attributive or linking article is a subtyp of a class of grammatical words (which he calls "operators"), which are labeled *articles*. Other subtypes of this class are definite, indefinite and other types of (non-attributive) grammatical markers.

[25] "[...] zugleich trennende und verbindende Funktion [...]"

[26] "[...] physisch selbständig [...]"

Yiddish
Greek (Modern)

homophone demonstratives/definite markers in that the article nominalizes an adnominal constituent in order to function as attribute of a certain kind. The homophone demonstrative/definite marker on the other hand, marks the whole noun phrase for certain values of the feature SPECIES.

While the use of attributive articles in German, English and several other Indo-European languages is restricted to epithet-constructions, a similar construction with an attributive article occurs much more unrestrictedly in Yiddish.

(28) Yiddish (Indo-European; **jacobs-etal1994**)

a. *di grin-e oyg-n*
DEF.PL green-DEF.PL eye-PL
'the green eyes'

b. *di oyg-n di grin-e*
DEF.PL eye-PL ATTR.DEF.PL green-DEF.PL
'the GREEN eyes'

c. *'n grin-et oyge*
INDEF.N green-INDEF.N eye(N)
'a green eye'

d. *'n oyge 'n grin-et*
INDEF.N eye(N) ATTR.INDEF.N green-INDEF.N
'a GREEN eye'

In the default attributive construction in Yiddish, the adjective precedes the noun which also triggers agreement on the adjective (28a+28c). In an emphatic construction and postponed to the head noun, however, the attributive adjective is marked with an article (28b+28d) (**plank2003**).

Yiddish thus shows that attributive articles can have a much broader use than for example in German. But even in Yiddish the use of the attributive article is subject to restrictions. In this case, the restriction is of a semantic nature and is due to the referential status of the adjective. In order to occur in an attributive nominalization construction the adjective must be in contrastive focus.

A similar rule applies to Greek, where the so-called repeated article also occurs in contrastive-focus constructions.

(29) Greek (Modern) (Indo-European; **ruge1986**)

a. *i kondés fústes*
DEF short skirts
'the short skirts'

b. *i fústes i kondés*
DEF skirts ATTR short
'the SHORT skirts'

Note that the the two phrases in the attributive apposition constructions (i.e. attributive nominalization) of German (§4.5.2.3), Yiddish (28) and Greek (29) cannot be re-arranged unless the whole construction yields a different reading. In the case of the epithet-construction in German, re-arrangement of adjective and noun would result in a simple noun phrase with an attributive adjective which is, however, no longer an epithet. Re-arrangement of the constructions in Yiddish and Greek would result in true noun phrase appositions.

4.5.2.4.1 Attributive articles as subtype of attributive nominalizers Attributive articles have been characterized as grammatical words and agreement targets. In accordance with the common practice of labelling an unchangeable, non-bound grammatical marker "particle", the attributive nominalizer *the* in English (epenthet constructions) would fall into this category because it is not an agreement target.[27]

In the present survey, however, there are only a few examples of languages with attributive, non-article nominalizers attested, among them Ket (cf. §**??**) and Dungan (cf. §**??**) where the respective markers seem to constitute affixes rather than particles.

In the present ontology, attributive articles are defined as a subclass of attributive nominalizers. Whereas attributive nominalizers are construct markers (belonging to pure morphosyntax), articles have an additional semantic component because they undergo agreement.

4.5.2.4.2 D-Elements which are not nominalizers In the previous section, attributive articles and other attributive nominalizers have been described and attributive nominalizers have been characterized as a special subtype of anti-construct state markers which attaches to an intermediate dependent phrasal constituent between the head noun and the modifier.

Somewhat similarly, Himmelmann **himmelmann1997** describes attributive articles and other attributive nominalizers as D(eterminer) elements between head and attribute[28]. Illustrating attributive nominalization with examples from

[27] Consider also Himmelmann's (**himmelmann1997**) "Gelenkartikel" versus "Gelenkpartikel".
[28] "D(eterminer)-Element zwischen Kopf und Attribut"

Swedish
Albanian

several languages, the author shows that these markers prototypically originate from adnominally grammaticalized local deictic pronouns used as functional heads of nominalizer phrases. Himmelmann does not, however, clearly distinguish between synchronic and diachronic evidence and considers both attributive nominalizers (such as the "repeated article" in Greek), agreement markers (such as the so-called "adjective article" in Albanian and even linkers (as in Tagalog) as D-elements.

The linker in Tagalog is not an article (not even an attributive nominalizer) according to the present ontology of attribution marking devices because the marker is floating, with a locus neither on-dependent or on-head, and it does not project a noun phrase (cf. §4.5.4 in Part II). Examples of agreement marking "D-Elements" come from Swedish and Albanian.

(30) a. Swedish (Indo-European; own knowledge)
den *goda* *vännen*
ATTR:DEF.SG.UTR good:DEF.SG.COM friend:DEF.SG.COM

b. Albanian (Indo-European; **himmelmann1997**)
shoku ***i*** *mirë*
friend:DEF:NOM.SG.M **NMLZ:NOM.SG.M** good:NOM.SG.M
'the good friend'

Whereas the agreement marking "D-Element" in Albanian is perhaps a nominalizer, the markers in Swedish (and other languages) might simply be construct-state agreement markers from a purely synchronic point of view because they do not occur in attributive apposition constructions, i.e. they do not project noun phrases (cf. §?? for Albanian and §?? for Swedish). From a diachronic point of view, however, these markers clearly originate from absolutely similar attributive nominalizers. Consequently, the grammaticalization path suggested by Himmelmann **himmelmann1997** can even be extended with an additional stage: from "D-elements" to attributive articles (or other attributive nominalizers) to construct-state markers, as will be shown in the diachronic Part **??**.

From a purely synchronic point of view, however, the different types of *anti-construct state agreement* and *attributive article* might not always be easily distinguishable from each other or from *head-driven agreement*. The first two often include some "article notion" (sometimes connected to definiteness or other referential values), and all three types include agreement marking. "Pure" agreement marking, however, cannot include the feature STATE (construct marking). A simple test is whether or not attributive adjectives show different agreement marking than predicative adjectives. If they do, as, e.g. in Russian, construct

marking is involved. If construct marking undergoes agreement and additionally projects a full noun phrase, as, e.g. the article in Germanic epithet constructions, than the type of marking is best characterized as attributive article.

Hopi, Toreva
Saamic, Northern

4.5.3 Head+dependent marking attributive state

This combined type refers to state marking which has two loci: on-head and on-dependent simultaneously. A language spoken outside the northern Eurasian area which gives an example of this noun phrase type is the Toreva dialect of Hopi.

(31) Hopi, Toreva (Uto-Aztecan; **whorf1946**)

a. *ca·va*
is_short

b. *pǫ̀yo*
knive

c. *ca·v* *vǫ̀yo*
is_short\ATTR knive\ATTR
'a short knive'

According to **whorf1946** both the adjective modifier (which is a stative verb in Hopi) and the noun head alter their phonological shapes regarding whether or not they are used in predication or as constituents in a noun phrase. Consider the noun phrase in example (31c) where the modifier *ca·va* 'is short' occurs with a shortened stem form (compared to 31a) and the noun is marked by means of lenition of the word-initial consonant (*pǫ̀yo* 'knive' versus *vǫ̀yo* [knive\ATTR]).

The noun phrase type in Hopi is thus best analyzed as attributive state marking in which both the noun head and the adjective dependent are construct marked. Note, however, that in contrast to the other mentioned examples of different types of state markers, the respective formatives in the noun phrase of Hopi are non-concatenative morphemes represented by stem alternations.

Double (head+dependent) construct state marking is also attested as adjective attribution marking device in one language of northern Eurasia. In Northern Saami, two adjectives meaning 'little' govern diminutive marking on the head noun. Noun phrases with these two adjectives are ungrammatical if diminutive marking on the noun is missing.

(32) Saamic, Northern (Uralic; own knowledge)

a. *Diminutive derivation*

Tagalog

guolli / guolá-š / guolá-ža-t
fish fish-DIM fish-DIM-PL
'fish' / 'little fish' / 'little fishes'

b. *Anti-construct state marking ('big')*[29]
stuorra *guolli / guoli-t / guolá-š / guolá-ža-t*
big:ATTR fish fish-PL fish-DIM fish-DIM-PL
'big fish' / 'big fishes' / 'big little-fish' / 'big little-fishes'

c. *Double-construct state marking ('little')*[30]
unna *guolá-**š** / guolá-**ža**-t*
smal:ATTR fish-DIM fish-DIM-PL
'small fish' / 'small fishes'

d. * ***unna*** *guolli / guoli-t*
small:ATTR fish fish-PL

Diminutive is a derivational category in Northern Saami. Normally it is assigned semantically to the noun and thus belongs to the morphological features, as in (32a+32b). However, diminutive can in fact also be a morphosyntactic feature in Northern Saami, namely when it is obligatorily governed by one of the two attributive adjectives *unna* or *uhca* 'little, small (attr.)', as in (32c). However marginal these examples seem to be, diminutive is assigned syntactically on the head by the dependent and thus also belongs to the morphosyntactic features in Northern Saami.

4.5.4 Neutral attributive state

The term *neutral marking* was introduced by Nichols (**nichols1986**) in her typology of head marking versus dependent marking grammar. *Neutral marking* refers to a marker's locus neither on-head nor on-dependent. This means that the marker floats in the noun phrase depending on the actual order of constituents. A floating state marker occurs, for instance, in Tagalog.

(33) Tagalog (Austronesian; **rubin1994**)

[29] State marking of 'big' is non-concatenative and affects the quantity of the stem consonants and the quality and quantity of the stem-final vowel, cf. the same adjective inflected for predicative state (agreement): *guolli/guoláš lea **stuoris*** [PRED:SG] 'the fish/little fish is big'; *guolit/guolážat leat **stuorrát*** [PRED:PL] 'the fishes/little fishes are big'.

[30] State marking of 'little' is non-concatenative and affects the quantity of the stem consonants and the quality and quantity of the stem-final vowel, cf. the same adjective inflected for predicative state (agreement): *guolli/guolá-š lea **unnni*** [PRED:SG] 'the fish/little fish is little'; *guolit/guolážat leat **unni*** [PRED:PL] 'the fishes/little fishes are small'.

a. *Predication*
Maganda ang bahay.
beautiful TOP house
'The house is beautiful.'

b. *Attribution (adjective-noun)*
*maganda-**ng** bahay*
beautiful-ATTR house
'beautiful house'

c. *Attribution (noun-adjective)*
*bahay **na** maganda*
house ATTR beautiful
'beautiful house'

In the Tagalog noun phrase, the combination of noun and modifier is licensed by the attributive state marker *-ng*.[31] The marker occurs with attributive adjectives (33b and 33c) but not with predicative ones (33a).[32]

The two types of adjective attribution in Tagalog (33b and 33c) are distinguished from each other only by word order of the head noun and the modifying adjective. The attribution marker follows the first constituent, regardless of whether this is the modifier or the noun. The attribution marker in Tagalog behaves thus like a second-position clitic (**nichols1986** see also **himmelmann1997**).

In the typology presented here only a floating state marker, i.e. an overt state marker which behaves neutrally with regard to its locus and is neither head- nor dependent marking, is considered to be a true *linker*. The occurrence of such an attribution marking device is not attested among the northern Eurasian languages investigated for the present study. However, since *linkers* and *articles* (but even other attribution marking devices) are sometimes not clearly distinguished in terminology (see below §4.5.2.4), it seems rather relevant to characterize this noun phrase type here.

4.6 An ontology of adjective attribution marking devices

The previous sections § 4.2–§ 4.5.4 of this chapter were aimed at typologizing adjective attribution marking devices. The attested devices described so far belong

[31] After consonants the allomorph *na* is used.

[32] The state marker in Tagalog is polyfunctional in the sense that it also marks attribution of demonstratives, numerals and other modifiers **himmelmann1997** See also below Chapter 5.

to the following noun phrase types:

- Juxtaposition
- Incorporation
- Construct state
- Linker
- Anti-construct state
- Attributive nominalization
- Attributive article
- Anti-construct state agreement
- Head-driven agreement
- Apposed head-driven agreement
- Modifier-headed possessor agreement

Table 4.5 on page 63 summarizes the typology presented in sections § 4.2–§ 4.5.4 and presents short definitions (including bracketed syntactic templates) and an example for each type.[33] Note that a lexical head is required only in certain noun phrase types. Note also that the constituent order (e.g. [$_{\text{NP}}$ A N] or [$_{\text{NP}}$ N A]) and the morpho-phonological fusion of formatives (e.g. (free) [$_{\text{NP}}$ A NMLZ], (cumulative) [$_{\text{AP}}$ A:ATTR:AGR] or (affixal) [$_{\text{AP}}$ A-ATTR]) is not relevant for the presented ontology.[34]

Table 4.1 on page 65 presents an ontological cross-classification of all devices defined earlier. This ontology has three main dimensions:

- *Syntactic source*, i.e. the central syntactic operation which constitutes attribution and belongs either to *agreement marking* or *government*. But note that syntactic government can include secondary, i.e. non-constitutional agreement.

[33] This overview is derived from the definition file of general noun phrase patterns included in **AUTOTYP-NP**

[34] The presented ontology is defined by (mostly) morphosyntactic parameters. But grammatical word-hood could be relevant for definitions of subtypes in the leaves of Figure 4.2. For instance *head-driven agreement* could perhaps be sub-divided into types exhibiting agreement affixes vs. grammatical agreement words.

- *Syntactic pattern*, i.e. devices projecting adjective phrases versus devices projecting full noun phrases (by means of attributive apposition or, in the case of modifier-headed possessor agreement, by converting the attribute to the "possessed" noun phrase).
- *Syntactic locus* of the respective formatives.

Figure 4.2 on page 66 presents a similar ontology in a tree diagram. The order of types (from left to right) is similar to Table 4.5 (from top to bottom). The left branch of the tree consists of a purely syntactic device (*juxtaposition*) with the subtype (*incorporation*); the middle branch consists of three overt morphosyntactic types differentiated by the locus of the respective formatives: on-head (*construct state*), floating (*linker*) and on-dependent. "Dependent-marking" again can be divided further into the three subtypes: *attributive nominalization*, *anti-construct state agreement* and *attributive article* (a subtype of *attributive nominalization*). The right branch of the tree, finally, comprises morpho-semantico-syntactic devices, i.e. devices primarily connected to head- (*head-driven agreement*) or dependent-driven agreement (*modifier-headed possessor agreement*). A dashed line combines the types of *head-driven agreement*, *anti-construct state agreement* and *attributive article* because (morpho-semantico-syntactic) agreement marking is involved in all of them.

Whereas construct- and agreement marking in the types of *anti-construct state agreement* and *attributive article* are combined in portmanteau morphemes (e.g. in the anti-construct state agreement marking suffixes in Russian) other devices can (or must) co-occur without being combined into one formative. Attested and non-attested combinations of adjective attribution marking devices are illustrated in Table 4.4. The attested co-occurring adjective attribution marking devices are:

- Anti-construct state agreement + Head-driven agreement ("Double agreement")
- Anti-construct state + construct state ("Double construct")
- Anti-construct state + attributive article ("Double construct")
- Attributive article + head-driven agreement ("Double agreement")

Table 4.4: Attested combined adjective attribution marking devices

Device 1	Device 2	Note
Juxt	–	No logical combination possible
Inc	?	No attestation of any combination
Constr	AConstr	Northern Saami (“Double construct”)
Nmlz (Art)	AConstr	Endo (“Double construct”)
ACAgr	HDAgr	Swedish (“Double agreement”)
Nmlz (Art)	HDAgr	Albanian (“Double agreement”)
Link	?	No attestation of any combination
MHPAgr	?	No attestation of any combination

Table 4.5: Attested adjective attribution marking devices with definitions

Type	Definition	Syntactic dependency	Commonly used label	Example language
Juxt	Unmarked sequence of constituents; Test: no additional morphemes available in NP	$[_{NP}[_{AP}$ A] (N)]	Juxtaposition	Komi
Inc	No additional morphemes available in NP, but dep is syntactic compound; Test: dep cannot occur unbound (headless)	$[_{NP}$ A-N]	Incorporation	Chukchi
Constr	Head-marking formative that only registers presence of dep; Test: formative does not undergo agreement and is not present without head (in predication) or without dep	$[_{NP}[_{AP}$ A (N:ATTR)]]	Ezafe	Kurmanji
Link	Floating formative (neither ad-head nor ad-dep, but truly ad-phrase) that only registers presence of head-dep relation; Test: formative not present without head (in predication) or without dep	$[_{NP}[_{AP}$ A] ATTR N]	Linker	Tagalog
AConstr	Dep-marking formative that only registers presence of head; Test: formative does not undergo agreement and is not present without head (in predication	$[_{NP}[_{AP}$ A:ATTR] (N)]	Attributive suffix	Saamic

Nmlz	Dep-marking formative that only registers presence of head by projecting full NP; Test: formative does not undergo agreement and is used in focus construction where inflection of the head is reduplicated	[NP[NP[AP A:NMLZ]] (N)]	Nominalizer	Udmurt
Art	Subtype of nominalizer that undergoes agreement	[NP[NP[AP A NMLZ:AGR]] (N)]	Double article	Yiddish
ACAgr	Dep-marking formative that registers presence of head and undergoes agreement triggered by the head; Test: not present without head (in predication)	[NP[AP A:ATTR:AGR] (N)]	Long-form adjective	Russian
HDAgr	Dep-marking formative that duplicates morpho-semantic features of the head	[NP[AP A:AGR] (N)]	Agreement suffix	Finnish
AHDAgr	AP marked with HDAgr but projecting a full NP in apposition; Test: AP is used in focus construction where inflection of the head is reduplicated, often with reversed constituent order	[NP[NP[AP A:AGR] (N)]	Appositional agreement	Georgian
MHPAgr	Head-marking formative that duplicates morpho-semantic features of (adjectival) dep by means of possessor agreement in a modifier-headed NP	[NP[PSD[A:POSS:AGR](PSRN)]	Possessive-like attribute	Saliba

Locus	Syntactic source					
	Government [+GOV] [±(secondary)AGR]				Agreement [−GOV] [+(primary)AGR]	
	Syntactic pattern				Syntactic pattern	
	[±AGR]	Embedded	Non-embedded	Incorporated	Embedded	Non-embedded
no marking	\/ /\		**Juxtaposition** [NP A (N)]	**Incorporation** [NP A *(N)]	\/ /\	\/ /\
floating	[−AGR]		**Linker** [NP A ATTR *(N)]			
	[+AGR]					
dep-marking	[−AGR]	**Nominalization** [NP [NP' A:NMLZ] (N)]	**Anti-Construct State** [NP A:ATTR (N)]		**Appositive Head-Driven Agreement** [NP [NP' A AGR] (N)]	**Head-Driven Agreement** [NP A AGR (N)]
	[+AGR]	**Article** [NP [NP' A:NMLZ:AGR] (N)]	**Anti-Construct Agreement** [NP A:ATTR:AGR (N)]			
head-marking	[−AGR]		**Construct State** [NP A *(N:ATTR)]		**Modifier-Headed Possessor Agreement** [NP [PSD A:POSS:AGR] (PSR N)]	
	[+AGR]					

Figure 4.1: Multidimensional ontology of noun phrase structures according to the parameters *syntactic source* (true agreement marking or governed [± additional agreement]) and *syntactic pattern* of the device (projects noun phrase, projects adjective phrase) as well as *syntactic locus* of the respective markers (on-head, on-dependent, floating). Note that some cells in the table are marked for logically impossible types, others are left open because the respective types were not detected in noun phrases with attributive adjectives. *Glosses and tags:* A=adjective, AGR=agreement, AGR=agreement marker, AP=adjective phrase, ATTR=Attribution marker, GOV=government, N=Noun, NMLZ=(attributive) nominalizer, NP=Noun phrase, POSS=possessive marker, PSD=possessed noun phrase, PSR=possessor noun phrase

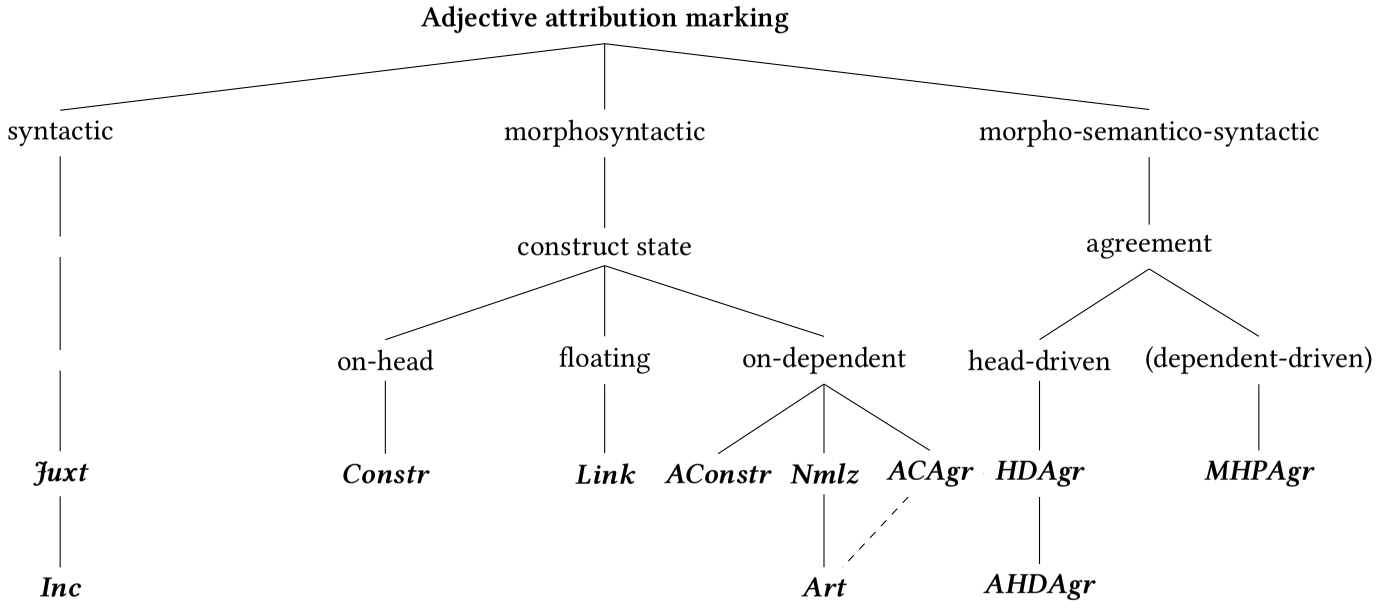


Figure 4.2: Ontological tree of attested adjective attribution marking devices. **Type abbreviations**: ACAgr=Anti-construct state agreement, AConstr=Anti-construct state, AHDAgr=Appositive head-driven agreement, Art=Attributive article, Constr=Construct state, HDAgr=Head-driven agreement, Inc=Incorporation, Juxt=Juxtaposition, Link=Linker, MHPAgr=Modifier-headed possessor agreement, Nmlz=Attributive nominalization

5 Excursus: Polyfunctionality of attribution marking devices

In a typological survey, noun phrases with adjectival modifiers can be examined from different perspectives. In the previous chapter, noun phrases with attributive adjectives were described according to their syntactic, morpho-syntactic, and/or morpho-semantico-syntactic structure. But noun phrase types of a given language can also be defined on a polyfunctionality scale with regard to the class of attributed elements beyond adjective attribution: Attributive adjectives and other adnominal modifiers (such as demonstratives, possessor nouns, adpositional phrases, clauses, etc.) may or may not be used in similar noun phrase structures.

Moreover, polyfunctionality is also relevant in languages where one and the same device is used as a nominal modification marker beyond attribution: for modification inside an adjective phrase (licensing, for instance, a degree word as modifier of an adjective) or as a modification marker inside an adpositional phrase (licensing, for instance, an adposition as determined by a noun phrase). *Attribution marker* should thus be understood as a term denoting a subset of *modification markers* relevant to nominal phrase structure in general.

Finally, the polyfunctionality concerns even the semantic content (or function) of certain devices beyond modification marking.

In the present chapter, polyfunctionality of adjective attribution marking devices will be illustrated with examples from a few languages.

5.1 Polyfunctionality of modification markers

In many languages, more than one class of attributes belong to one and the same noun phrase type. Some languages exhibit even highly polyfunctional noun phrase types and use one and the same device for licensing verbs, nouns, adjectives and even other syntactic classes as attributive modifiers inside noun phrases.

check zu possessor compounding Dahl 2004. The growth and maintenance of linguistic complexity. Studies in language companion series, 71. Amsterdam: John Benjamins.

In example (1) from Mandarin Chinese, the anti-construct state marker *de* illus-

Mandarin Chinese
Minangkabau

trates a highly polyfunctional attribution marking device. It licenses adjectival (1b), nominal (1a) and verbal attributes (1c).[1]

(1) Mandarin Chinese (Sinotibetan; **li-etal1981**)

a. *Noun (possessor) attribute*
Zhāngsān ***de*** *shū*
Zhangsang ATTR book
'Zhangsang's book'

b. *Adjectival attribute*
xīn ***(de)*** *shū*
new (ATTR) book
'new book'

c. *Verbal (relative clause) attribute*
wǒ zuótiān mǎi ***de*** *shū*
11SG buy yesterday ATTR book
'the book I bought yesterday'

In Minangkabau, an Austronesian language spoken on Sumatra in Indonesia, juxtaposition is polyfunctional to a similar degree.

(2) Minangkabau (Austronesian; **gil2005**)

a. *Noun (possessor) attribute*
batiak Kairil
papaya Kairil
'Kairil's papaya'

b. *Adjectival attribute*
batiak kuniang
papaya yellow
'a/the yellow papaya'

c. *Verbal (relative clause) attribute*
batiak Kairil bali
papaya Kairil buy
'a/the papaya that Kairil bought'

[1] Note, however, that the attributive marker is not obligatory. In noun phrases with pronominal and adjectival attributes, it can also be omitted. If *de* is used with adjectives, a certain clarifying or delineating stress is put on the denoted property, e.g. *hóng hūa* [red flower] 'a red flower', *hóng de hūa* [RED ATTR flower] 'a flower that is red (and not of a different color)' (**li-etal1981**).

Tagalog
Persian

Tagalog is another language with a polyfunctional attribution marker. The Tagalog linker, however, is less polyfunctional than juxtaposition in Minangkabau or anti-construct state marking in Mandarin Chinese. It marks only verbal and adjectival attributes.[2]

(3) Tagalog (Austronesian; **gil2005**)

a. *Adjectival attribute*
*pula=**ng*** *mangga*
red=ATTR mango
'red mango'

b. *Verbal (relative clause) attribute*
*binili ni Jojo=**ng** mangga*
bought PERS.GEN Jojo=ATTR mango
'mango that Jojo bought'

Highly polyfunctional attribution marking by means of a head-marking construct suffix is found even in Persian.[3]

(4) Persian (Indoeuropean; **mahootian1997**)

a. *Modification inside an adpositional phrase*
*tu-**ye** ašpæzxune*
in-MOD kitchen
'in the kitchen'

b. *Nominal attribution*

i. *Noun (non-possessor) attribute*
*æŋgoštær-**e** ælmas*
ring-MOD diamond
'diamond ring'

ii. *Noun (possessor) attribute*
*æŋgoštær-**e** pedær*
ring-MOD father
'father's ring'

[2] Note that that the constituent order of attribute and head noun is free in Tagalog: The relative clause and the adjective can also occur in a head-initial phrase type. In this case, the linker *=ng* attaches phonologically to the noun (**gil2005** **himmelmann1997**).

[3] Note the consistent glossing MOD instead of ATTR. The Persian construct marker licenses modification beyond attribution.

Västerbotten-Swedish

c. *Adjectival attribute*
*ælmas-**e*** *bozorg*
diamond-MOD big
'a big diamond'

d. *Adpositional attribute*
*miz-**e*** *tu-ye* *ašpæzxune*
table-MOD in-MOD kitchen
'the table in the kitchen'

e. *(Infinite) verbal attribute*
*væqt-**e*** *ræftæn*
time-MOD to_go
'time to go'

While nominal, adjectival, adpositional and (infinite) verbal attributes are marked by the same device, finite verbal attributes (relative clauses) never occur in a similar noun phrase type in Persian.

In Västerbotten-Swedish, a language of the northern Eurasian area under investigation, attribution marking by means of adjective incorporation is also considered to be polyfunctional (cf. Sections 4.3 and **??**). Beside adjective attribution, the device marks attribution of (human) possessors.

(5) Västerbotten-Swedish (Indoeuropean; **gil2005**)

a. *Noun (human possessor) attribute*
Pelle-äpple
Pelle-apple
'Pelle's apple'

b. *Adjectival attribute*
rö-äpple
red-apple
'red apple'

bessere Referenz zu VB-Schwedisch??

Gil (**gil2005**) surveyed the polyfunctionality of attribution markers licensing possessor nouns, adjectives and relative clauses in a world-wide sample of languages. According to the number of morpho-syntactically differentiated classes of attributes Gil grouped the languages of his sample into the following types:

- **Weakly differentiating languages** using polyfunctional devices for attribution of all three syntactic categories, as in Mandarin (1) and Minangkabau (2)

- **Moderately differentiating languages** using polyfunctional devices for attribution of two syntactic categories, for instance:
  - adjectives and relative clauses, as in Tagalog (3)
  - possessor nouns and adjectives, as in Västerbotten-Swedish (5) and Persian (4)
- **Highly differentiating languages** are not polyfunctional at all, as in German where the three syntactic classes are marked differently.

In Gil's sample, Europe and adjacent parts of Asia and Africa stand out as an area with predominantly non-polyfunctional languages, while almost all languages of Southeast Asia are of low differentiation (**gil2005**).

Northern Eurasian languages of the "moderately differentiating" type included in Gil's sample are Japanese and Västerbotten-Swedish (with polyfunctional attribution marking of possessor nouns and adjectives) as well as Ainu, Nivkh and Tatar (with polyfunctional attribution marking of adjectives and relative clauses).[4] No languages of the "weakly differentiated" type are known to occur in the northern Eurasian area. Figure 5.1 illustrates the polyfunctionality of modifi-

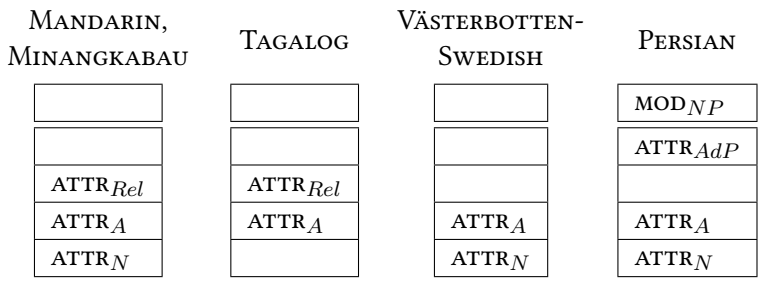


Figure 5.1: Functional maps for modification markers: the anti-construct state marking in Mandarin Chinese and juxtaposition in Minangkabau, the linker in Tagalog, adjective incorporation in Västerbotten-Swedish and construct state marking in Persian

cation markers in the languages mentioned in this chapter.[5] The true attributive

[4] Note that English is not coded as "moderately differentiating" by **gil2005** although juxtaposition can be polyfunctionally used as a device for attribution of adjectives and relative clauses (with reverse constituent order though: *The woman I saw.*)

[5] Cf. **haspelmath2003** for a systematic and historiographic description of functional (or semantic) maps.

functions of the marker, i.e. licensing of adpositional, verbal, and adjectival attributes, are found in the middle cells of the left column in Figure 5.1. The cell extending upwards shows the additional function of the marker as licensee of modification above the noun phrase level (i.e. inside an adposition phrase).

The order of ATTR$_{Rel}$ through ATTR$_{N}$ in these functional maps corresponds to the hierarchical alignment of polyfunctional attribution marking suggested by Bingfu Lu and Zhenglin Qu.[6]

gibt es eine Publikation inzwischen?

(6) *Noun < Adjective < Verb*

The hierarchy is to be read as follows: The highest category of attributive modifiers are verbs (i.e. relative and other attributive clauses), the next lower categories are adjectives and nouns. If one attributive category is marked with a polyfunctional attribution marker, the less bounded category adjacent to the left side in the hierarchy should be marked by the same device, too.

ATTR$_{Rel}$	NMLZ	FOC
ATTR$_{A}$		
ATTR$_{N}$		

Figure 5.2: Functional map for the modification marker *ve* in LAHU

5.2 Polyfunctionality of modification markers and additional content

Polyfunctional modification marking devices with semantic content (or function) beyond attribution are also attested in several languages. Lahu is an example of a Southeast Asian language of the "weakly differentiating" type according to Gil's (**gil2005**) classification. Syntactically similar to Mandarin Chinese, Lahu exhibits an anti-construct state marker *ve* that licenses adjectival (7a-i), nominal

[6] Lu's and Qu's hierarchy, cited from a LingTyp posting ("The alignment of modification coding", LingTyp Item #2580, 6 May 2009, 01:36, http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0905A&L=LINGTYP&P=R146) is based on a similar hierarchy for Austronesian languages by Foley (**foley1980**). Note that Foley's hierarchy is proposed to be cross-linguistically valid and even includes two more syntactic classes than considered here: Determiner > Numeral > Noun > Adjective > Verb.

(7a-ii) and verbal attributes (7a-iii). In addition, the marker *ve* in Lahu is used as nominalizer (7b-i) and as focus marker (7b-ii).[7] Lahu

(7) Lahu (Sinotibetan; **matisoff1973**)

a. *Attribution*

i. *Adjectival attribute*
dàʔ ***ve*** *ŋâʔ*
pretty ATTR bird
'pretty birds' (194)

ii. *Noun (possessor) attribute*
Càlɔ̂ ***ve*** *ɔ̀ha*
Jalaw ATTR picture
'Jalaw's picture' (141)

iii. *Verbal (relative clause) attribute*
cɔ́ *câ* ***ve*** *ŋâʔ*
boil eat ATTR bird
'birds one boils to eat' (194)

b. *Additional semantic content*

i. *Nominalization (of a complement clause)*
nɔ̀ *qôʔ* ***ve*** *thàʔ ŋà mâ* *na* *ɣa* *qôʔ-ma!*
you say NMLZ ACC I NEG understand be_able INTERJ
'I can't catch what you're saying!' (157)

ii. *Focusing (of a clause)*
mâ *qay* ***ve***
NEG go FOC
'I am certainly not going.' (362)

The functions of the marker *ve* in Lahu can also be summarized in a functional map, see Figure 5.2. The true attributive functions of the marker, i.e. licensing of verbal, nominal and adjectival attributes, are found in the cells of the left column in Figure 5.2. The cells extending to the right show the additional content of the attributive marker, i.e. as a nominalizer and focus marker of a clause.

[7] See **bickel1999** on the "Standard Sino-Tibetan Nominalization pattern" which in some languages include even additional content beyond attribution, nominalization and focus.

5.3 Conclusion

From a purely synchronic point of view, polyfunctionality of adjective attribution marking devices seems less relevant to the area under investigation, northern Eurasia. Most languages of the area exhibit highly differentiated attribution marking devices. Languages of the "moderately differentiating" type are rare; no languages of the "weakly differentiated" type are known to occur in the northern Eurasian area at all.

However, polyfunctionality can perhaps indicate historical change if additional semantic content of attribution marking devices across related languages is taken into consideration. The topic of polyfunctional attribution markers across languages of one family will thus be taken up again in Part III of this study.

Part III

Synchrony

6 Introduction

The geographic area covered in the present survey stretches from Europe (including the Mediterranean Islands Malta and Cyprus as well as the regions Anatolia and Caucasia), over central, northern, and northeastern Asia (including the whole of Siberia, the adjacent parts of northern Mongolia and Manchuria) to the Islands of the northwestern Pacific Ocean. The language families represented in this area are genealogically categorized by **salminen2007** in his chapter on the endangered languages of "Europe and North Asia". By and large, Salminen's index of languages will be followed here. However, the present survey strictly follows the geography of northern Eurasia and consequently also includes Siberian Yupic Eskimo, Ainu, the Sinotibetan language Dungan, and some Semitic languages.

6.1 The languages of northern Eurasia

Adopting Salminen's rather cautious genealogical classification the following families are considered (roughly from Northeast to Southwest):

1. Eskimo-Aleut
2. Chukotkan
3. Kamchatkan
4. Nivkh
5. Ainu
6. Japanese
7. Korean
8. Sinotibetan
9. Mongolic
10. Tungusic
11. Yukagir
12. Yeniseian
13. Turkic
14. Nakh-Daghestanian
15. Abkhaz-Adyge
16. Kartvelian
17. Semitic
18. Uralic
19. Indoeuropean
20. Basque

Even though some of these genealogical units have been assumed to combine to larger stocks (as Altaic, Chukotko-Kamchatkan, North-Caucasian and other) the restriction to uncontroversial units seems adequate for the present areal typological investigation. This is especially true since an attempt is made to map variation inside genealogical units rather than to evaluate a statistically balanced genealogical sample of languages.

6.2 The language sample

All attested adjective attribution marking devices of languages mentioned in the present study are coded in Table **??** in the appendix.[1] This table thus includes a relatively complete list of languages from the northern Eurasian area. At least one representative of each existing genera is found in that sample. Additionally, several languages from within or outside the area (all of which are mentioned in other chapters of this investigation) or even other languages on which information was easily accessible are coded.

All languages are sorted alphabetically according to their genealogical affiliation. For each of the languages, the attested noun phrase type(s) relevant to adjective attribution marking are listed.

6.3 The language maps

The language maps have been generated using the interactive reference tool for the World Atlas of Language Structures (**bibiko2005**).

6.3.1 Data points for geographic coding

Each language is displayed as one data point. The respective geographic coordinates have either been taken from **WALS** or were included using the language coordinates provided by **AUTOTYP** or on Ljuba Veselinova's website.[2] For some languages missing in the mentioned databases new coordinates had to be defined based on the main geographic location where the respective languages are spoken.

Displaying the distribution of a given feature by means of a borderline around a group of languages – as in the maps used by typological surveys of the EUROTYP-

[1] The table is derived from **AUTOTYP-NP** where these languages are coded for noun phrase patterns.

[2] http://www.ling.su.se/staff/ljuba/16.02.2014

project[3] – was not preferred because these maps might imply the existence of isoglosses around continuous language and dialect areas. A typological survey of non-continous languages seems rather inadequate for drawing such isoglosses.[4]

6.3.2 Data points for type coding

Farben markieren verschiedene Typen; Formen markieren Untertypen (secondary types)

In several languages more than one default attribution marking device occurs, for example in Albanian (cf. (??) in § ??) where two lexical classes of adjectives exist: one of them marked for head-driven agreement, the other simultaneously marked for head-driven agreement and attributive nominalization. In the map's legend, a slash marks the occurrence of multiple basic types in one language: Albanian *HDrAgr/Nmlz+HDrAgr*.[5]

Parentheses denote secondary types of attribution marking devices with additional semantic content, as Chuvash (cf. (??) in § ??) where attributive adjectives are normally juxtaposed but can alternatively be marked for attributive nominalization in contrastive-focus constructions: Chuvash *Juxt(Nmlz)*.

Square parenthesis will be used for languages where the occurrence of a given type of attribution marking devices seems even more restricted or if the device's characteristics remain uncertain due to inadequate data. Consider for example Turkish (cf. (??) in § ??) where attributive nominalization occurs as a secondary type but is restricted to headless noun phrases in direct object position (marked for accusative): Turkish *Juxt[Nmlz]*. Secondary and tertiary types are not coded in the maps.

6.3.3 The maps

The maps in Figure ?? and Figure ?? show the distribution of different adjective attribution marking devices across those world's languages mentioned in the present study. Whereas all types are coded with different colors or shapes in Figure ??, a similar language sample is coded only for the main morpho-syntactic types (juxtaposition, agreement, construct state, incorporation) in Fig-

[3] http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/eurotyp-guidelines/16.02.2014

[4] Cf. also Van Pottelberge's **van-pottelberge2001** critique of EUROTYP's "name maps". Furthermore, the Eurotyp sample of languages are somewhat arbitrary. The western Romance varieties, for instance, are represented in large number whereas varieties of Balkan Romance (Megleno-Rumanian, Aromunian, etc.) are missing completely. Also the whole Saamic branch is represented in the Eurotyp sample as one single language only even though Saamic languages are as comparably diverse as Romance languages.

[5] Type abbreviations are explained in Table ?? in the appendix.

ure **??**. Note that these world maps do not reflect systematic sampling but are rather the result of random choice due to my work with data coded for the noun phrase structure module of AUTOTYP (**AUTOTYP-NP**). Note also that the maps show fewer languages from the northern Eurasian area than actually coded in Table **??**.

The other pairs of maps are coded similarly but zoom in on northern Eurasia (Figure **??** and Figure **??**), on North-Asia (Figure **??** and Figure **??**) and on Europe (Figure **??** and Figure **??**). Whereas the maps of northern Eurasia and North-Asia show only representatives of each genera, the maps of Europe present a more complete picture. The reason for displaying a deeper resolution in the European map is the easier accessibility of data for almost all existing languages of that area. Displaying a similar deep resolution on the whole northern Eurasian area was not possible due to lack of data for several languages.

In order to present a balanced picture, several European languages are thus not displayed in the larger map of northern Eurasia. When a choice had to be made whether or not to keep a language inside a given genera, this was always done in favor of diversity rather than unity. One genera can even be represented by more than one language in order to display extraordinary diversity inside that group of closely related languages. Consequently, the northern branch of Germanic is represented by Icelandic (with *HDrAgr*), Swedish (with *ACAgr+HDrAgr/HDrAgr*) and Västerbotten-Swedish (with *Inc/HDrAgr*) (cf. § **??**).

The choice to let the maps illustrate the highest possible diversity instead of displaying a genealogically and geographically balanced picture is legitimated by the general goal of the present investigation, namely the synchronic and diachronic mapping of cross-linguistically attested adjective attribution marking devices in a geographically restricted area. Whereas the mapping of synchronically attested diversity is the aim of the present part, Part **??** will inspect this diversity form a diachronic perspective.

7 Areal uniformity and diversity in northern Eurasia

In the previous chapter, the prototypical and the known minor noun phrase types occurring in the languages of northern Eurasia were characterized and illustrated with examples. This survey thus provides an overall picture of the degree of typological uniformity or divergence with regard to adjective attribution marking marking within both the whole area and each genealogical unit.

7.1 Attested attribution marking devices

Fourteen (simple and combined) types of adjective attribution marking devices are attested in the languages of northern Eurasia:

- Anti-construct state (as in Saamic)
- Anti-construct state + head-driven agreement ("double agreement", as in Swedish)
- Anti-construct state + construct state ("double-construct state", as in Northern Saami)
- Anti-construct state agreement (as in Russian)
- Appositive head-driven agreement (as in Georgian)
- Attributive article (as in Yiddish)
- Attributive article + head-driven agreement ("double agreement", as in Albanian)
- Incorporation (as in Chukchi)
- Attributive nominalization (as in Dungan)
- Attributive nominalization (as in Udmurt)

- Construct state (as in Kurmanji)
- Juxtaposition (as in Komi)
- Head-driven agreement (as in Finnish)
- Modifier-headed possessor agreement (as in Oroch)

Only one type attested in the world-wide sample (Table ??) does not occur in the northern Eurasian area: the floating construct state marker (*linker*) found, for instance, in Tagalog (Austronesian).

The Indoeuropean family has the largest absolute number of attested adjective attribution marking devices (nine), followed by Nakh-Daghestanian and Uralic (five each), and Kartvelian and Tungusic (four each). The Mongolic family has the lowest possible number with only one attested device, just as with Kamchatkan and the isolates Ainu, Basque, Korean and Nivkh.

The most rare types are: (1) modifier-headed possessor agreement, which is attested only as a secondary device in a few Tungusic languages, and (2) the combined construct device (i.e. "double-construct state"), which is attested only marginally in one single language, Northern Saami (Uralic). Attributive nominalization (as an article combined and with head-driven agreement) is also very rare. This type occurs as the primary device only in the Albanian branch of the Indoeuropean family, but it is also attested as a secondary or tertiary device in other languages. Head-marking construct state is also relatively uncommon in the northern Eurasian area as it is attested only in Iranian languages (Indoeuropean).

The most common type is juxtaposition, followed by head-driven agreement.

7.2 Prototypes of attribution marking devices

Several language families of northern Eurasia exhibit clear prototypes of adjective attribution marking devices: All Mongolic and Turkic languages have juxtaposition as the default device, as is the case for the languages of most branches of Uralic as well. Head-driven agreement occurs as another prototype in many branches of the Indoeuropean family. Even though the attested deviation from the prototype is much higher in Indoeuropean than in Mongolic, Turkic and Uralic, head-driven agreement marking can be shown to occur prototypically in most Indoeuropean genera.

For Abkhaz-Adyge, Chukotkan, Kartvelian, Nakh-Daghestanian and Tungusic families, synchronic prototypes are not very easy to find because a predominant

type does not occur inside these families. The other language families of northern Eurasia are either isolates (Nivkh, Ainu, Japanese, Korean, Basque) or they exhibit rather shallow genealogical diversity (Kamchatkan, Yukagir, Yeniseian). Together with a few other families, predominantly spoken outside the investigated area (Eskimo-Aleut, Sino-Tibetan, Semitic), these families are excluded from generalizations about prototypes.

Larger language families with strikingly high diversity in regard to the attested absolute number of adjective attribution marking devices are Indoeuropean, Nakh-Dagestanian, Uralic and Tungusic. A strikingly high degree of unity is found in Mongolic and Turkic.

7.3 Diachronic implications of uniformity and diversity inside and across genera

Measuring the degree of diversity (or unity) from a synchronic point of view may help identify diachronic processes. A significantly high degree of diversity inside a given genus as compared to its proto-stage is likely to manifest pervasive linguistic changes and the innovation of new types. Similarly, the synchronic attestation of a high degree of unity inside a given genus indicates the inheritance of original types without significant innovations.

A genus is defined as a group of related languages which go back to a reconstructed (or documented) proto-form. The East-Saamic languages, for instance, form a group of sister languages which derived from proto-East-Saamic. Proto-East-Saamic is derived together with its Saamic sister languages from a more distant proto-stage, i.e. proto-Saamic which again is derived together with its Uralic sister languages from proto-Uralic. Since the proto-stages of languages are normally reconstructed as single languages it can be assumed that most of them had only one single type of adjective attribution marking (similar to the prevailing number of languages spoken today, cf. the sample in Table ?? in the appendix). It is self-evident that daughter languages which descend from a proto-language will either inherit the original adjective attribution marking devices, innovate secondary (or tertiary etc.) devices or replace the original devices with new ones. The proto-Saamic daughter language of proto-Uralic, for instance, has replaced the original Uralic juxtaposition with anti-construct state marking (see § ??). The proto-Slavic daughter language of proto-Indoeuropean inherited the original Indoeuropean head-driven agreement marking but innovated a secondary type, i.e. anti-construct state agreement marking (see § ??). All modern Mongolic languages, by contrast, exhibit juxtaposition uniformly and have

obviously inherited this device from their proto-languages (proto-Dagur, proto-Moghol, proto-Mongolian etc.) which in turn must have inherited juxtaposition from proto-Mongolic. A comparison of synchronically attested diversity inside and across genera might thus have diachronic implications.

The simple statistics in Table ?? tries to scope the degree of diversity in the investigated families of northern Eurasia. Column 1 lists all families, branches and subbranches in alphabetical order. Isolates and genera with only one member language are not included in the table, and neither are genera which are not spoken predominantly in northern Eurasia, with only two exceptions: the Iranian and Indo-Aryan subbranches within the Indoeuropean family. Since the highest possible diversity is of interest here, the number of all attested devices devices (including secondary and tertiary types restricted to special noun phrase types) are counted.

The second column in Table ?? ("Units (abs.)") gives the absolute number of coded languages from each genus. The third column ("Types (abs.)") gives the absolute number of attested types. The next two columns 4 and 5 present ratio figures. The first of them ("Ratio (gen.)") results from dividing the number of attested types in the given genus by the number of types attested for the higher branch:

$$Diversity_{genus} = \frac{Types_{genus}}{Types_{family}}.$$

For instance, West-Saamic has a ratio of 1.00 because it exhibits all 4 types attested in the whole Saamic branch. The Saamic branch as such has a ratio of 1.25 because 4 types are found in Saamic compared to 5 attested for the whole Uralic family. Similarly, South-Slavic has also a ratio of 1.00 because it exhibits all 3 types attested in Slavic. But the Slavic branch as such has a higher ratio of 3.00 (meaning a lesser degree of diversity) because only 3 types are attested in this branch out of 9 types for the whole Indoeuropean family.

The last ratio figures ("Ratio (lgs.)") result from dividing the overall number of languages by the number of attested types in the given genus:

$$Diversity_{languages} = \frac{Languages_{genus}}{Types_{genus}}.$$

For instance, 5 West-Saamic languages are coded for 4 different types, resulting in a ratio of 1.25. For the whole Saami branch all together 9 languages are coded for 5 types, resulting in a somewhat higher ratio figure of 1.80. South-Slavic has the ratio of 1.33 because the 4 South-Slavic languages are coded for 3 types; Slavic, however, has 4.33 because 13 Slavic languages are coded for only 3 different types.

[h]l l l || c || c | c | c || c c c | c