Michael Rießler. 2016. *Adjective attribution* (Studies in Diversity Linguistics 2).
Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog

ISBN: 978-3-944675-65-7 (Digital)
978-3-944675-66-4 (Hardcover)
978-3-944675-49-7 (Softcover)
978-1-530889-34-1 (Softcover US)
ISSN: 2363-5568

Cover and concept of design: Ulrike Harbort
Typesetting: Felix Kopecky, Sebastian Nordhoff, Michael Rießler
Proofreading: Martin Haspelmath, Joshua Wilbur
Fonts: Linux Libertine, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
Habelschwerdter Allee 45
14195 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

За най-любимите ми Алма, Ива и Кристина

Contents

Preface

This is a thoroughly revised version of my doctoral dissertation *Typology and evolution of adjective attribution marking in the languages of northern Eurasia*, which I defended at Leipzig University in January 2011 and published electronically as **riesler2011a** I am indebted to my family members, friends, project collaborators, data consultants, listeners, supporters, sources of inspiration, opponents and other people who assisted in completing my dissertation.

I am very thankful to the series editors who accepted my manuscript for publication with this prestigious open-access publisher, to the technical staff at Language Science Press, as well as to proofreaders and other individuals who have spent their valuable time producing of this book.

My sincere thanks are due to my thesis supervisor and *Doktorvater* Balthasar Bickel, to my second thesis supervisor and mentor throughout my whole career as a linguist Jurij Kusmenko as well as to Martin Haspelmath, who provided particularly valuable comments after his careful reading of my manuscript. Other important comments on the manuscript where provided by Rogier Blokland, Ciprian Gerstenberger, Martin Kümmel and Joshua Wilbur as well as two anonymous reviewers.

Freiburg, 1st July 2016 Michael Rießler

Abbreviations and notational conventions

1 Morphological glosses

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

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

2 Syntactic classes and phrase constituents

A	adjective
AdP	adposition phrase
AP	adjective phrase
ART	(attributive) article
CASE	case (clitic)
DEF	definite article

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

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

3 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)		

4 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, like in English *the house* (*the* ⇐ Old English *þæt*) and Swedish *hus-et* (*-et* ⇐ Old Norse *hið*).

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 book 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 as well as the WALS project (**walsOnline2013**). 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 (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.

Part IV (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.

The book's last Part V (Conclusions) summarizes my findings. In addition, there is an appendix, containing maps and the sample of languages used for my study, as well as indices with references to names, languages and subjects.

[1] Cf. http://www.autotyp.uzh.ch 19.07.2016

[2] Cf. http://www.degruyter.com/view/serial/16329 19.07.2016

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 a 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 "nominal attributes" (or noun phrases), "adjectival attributes" (or adjective phrases), "adpositional attributes" (or adposition phrases) and "clausal attributes" (or 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 father's*), adjective phrases (*brand new*), adposition phrases (*in the village*) or relative clauses (*which was expensive*).

2.2 Adjectival modifiers

This book presents a cross-linguistic comparison of "adjectival attributes", or ATTRIBUTIVE ADJECTIVES. It investigates the syntactic and morpho-syntactic 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

[1] Possible syntactic dependencies between modifying constituents inside this noun phrase are ignored in this illustrating example.

form a distinct word class, in other languages adjectives may not be clearly distinguishable from other parts of speech and constitute a flexible category together with nouns or with verbs. In a third group of languages, adjectives do not exist as a distinct word class at 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. On the comparative concept of adjectives, see also also **haspelmath2010b**

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**).

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

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

LATEX soll die Tabelle nicht lieber zwischen Absätzen liegen?

If a language does not exhibit a distinct class of adjectives, inherent 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 an adnominal modifier (**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 Figure **??**.

In the "flexible" language types 1–2 in Figure **??**, 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 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 in their morpho-syntax. The

[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 **??** 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**

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.

2.3.1 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; personal 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; personal knowledge)

a. *a large one*
INDEF large $_{\text{HEAD}}$:SG
'a large one'

b. *large ones*
large $_{\text{HEAD}}$: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; personal 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 (??) exemplifies a covert syntactic device because the relative clause is simply juxtaposed, (??) is an overt syntactic device because the relative clause is marked by an invariable formative. In (??, ??), 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.

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; personal 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 morpho-syntactic device. Other morphological marking in German occurs on syntactic units or on constituents of syntactic units without belonging to morpho-syntax. For instance, the plural inflection on the head noun (*Häus-er*) or the inflectional circumfix yielding a participle (*ge-bau-t*) in (??) belongs exclusively to the level of (inflectional and derivational) morphology but not to syntax.

(6) German (Indo-European; personal 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

[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'.

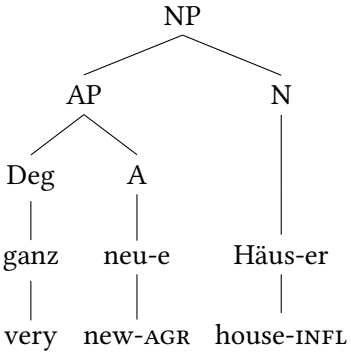
similar attribution marking. The attribution marking device (i.e., the agreement inflection) attaches to the participle stem as such (in boldface in example **??**). The participle derivation of the verb root *bau-* yielding this new stem does not belong to the sphere of syntax. Similarly, category-changing 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.

2.3.2 Headless noun phrases

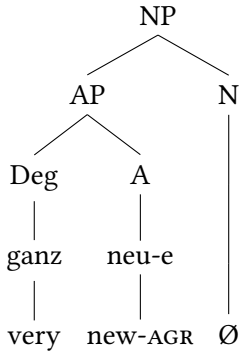
Adjectives as well as various other modifiers can also occur in noun phrases without a noun. Normally, this is the case with adjectives in elliptical constructions or adjectives which are made to nouns by means of a derivational process ("substantivized"). In many languages, noun phrases with and without an overtly expressed head noun exhibit a similar phrase structure, as in the following examples from German. The syntactic structure of the two examples in (**??**) is

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

a. *ganz neue Häuser*



b. *ganz neue* (viz. *Häuser*)



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 morpho-syntactic agreement features in both examples. Even though the adjective in the headless phrase is semantically a noun and used referentially, it is still syntactically the modifier of the (elliptic) 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 of a noun phrase 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 (**??**) and (**??**). Only when attributive adjectives occur in elliptical noun phrases are they inflected identically to nouns, as in (**??**) and (**??**).[6]

(8) Kildin Saami (Uralic; personal 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'

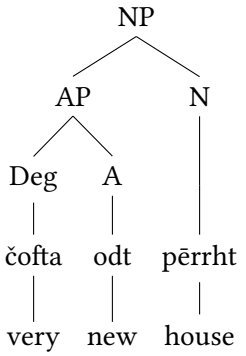
If the elliptical construction in Kildin Saami is analyzed as having an empty syntactic head position, as in German, an explanation for the different behavior of the (nominal) case inflection is needed. Unlike in German, where (nominal) inflection is always bound to the noun, inflection in Kildin Saami 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. 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. Being a grammatical word, hence a constituent in the phrase structure, *one* is sometimes described as "dummy head" in English (**rijkhoff2002**) replacing the noun at the syntactic

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

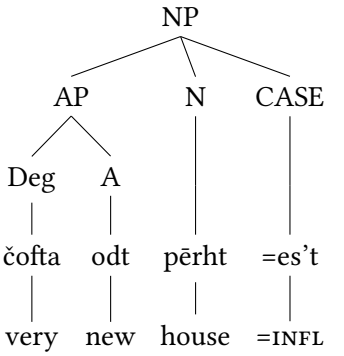
[6] The stem alternation in the adjective *odt : od-* is due to a regular morpho-phonological process.

(9) Kildin Saami (Uralic; personal knowledge)

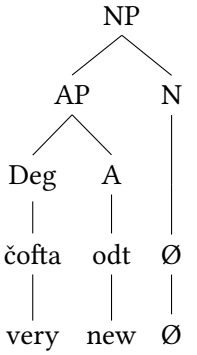
a.



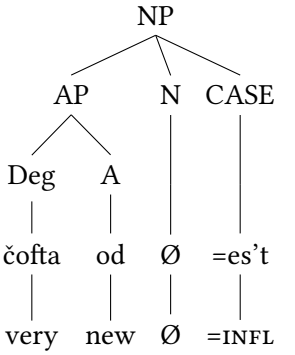
b.



c.

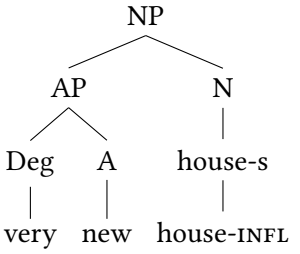


d.

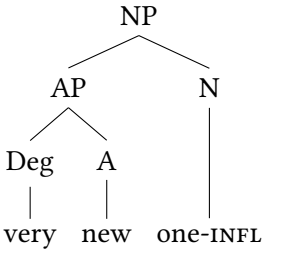


(10) English (Indo-European; personal knowledge)

a. very new houses



b. very new ones (viz. houses)



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

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 (**??**).[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 **??**). In postposition (marking emphasis), however, the adjective inflects for the full set of cases and numbers (**??**). 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 §**??**.

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

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; personal 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 (**??**) could simply be analyzed as an integral noun phrase differentiated from other non-emphasized noun phrases by constituent order. Georgian, however, is different from Bulgarian. The emphasized noun phrase in (**??**) exhibits different morpho-syntactic marking due to the additional agreement features 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 (see §**??**). Attributive nominalization can be illustrated with the epithet construction in German.

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

3 The syntax-morphology interface

3.1 Morpho-syntax

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 morpho-syntactic features, which is grounded in other work, for instance by **aronoff1994**; **corbett1987**; **carstairs-mccarthy2000a**; **corbett2006**; **corbett-etal2006**; **bickel-etal2007**; **kibort2008a** will be evaluated in the following sections. It will be shown that true morpho-syntactic 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 "morpho-syntax" (or "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 morpho-syntax is here understood as the interface between syntax and morphology, i.e., syntactic structure assigning morphology on one or more of its constituents.

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.

Morphosemantic features Morphosemantic features also only have inherent values whose assignment is not sensitive to syntax. The values of morphosemantic features are selected from a range of values. However, unlike purely mor-

phological features, the selection is based on semantic criteria. A prototypical example of the assignment of a morphosemantic feature is definiteness marking.

Morpho-syntactic features Morpho-syntactic 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 morpho-syntactic feature belongs per definition both to morpho-syntax – due to the feature's contextual assignment to the agreement target – and simultaneously to pure morphology (or morphosemantics) – due to the feature's status inherent in the agreement trigger.

The difference between morpho-syntactic and purely morphological (or morphosemantic) 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 (**??**, **??**, **??**). 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=DEF.M.SG boy(M)
'the GOOD boy'

(3) Bulgarian (personal knowledge)

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

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

The feature SPECIES,[1] however, does not belong to morpho-syntax 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 morpho-syntactic only in Albanian. In Bulgarian and Rumanian definiteness is purely morphological.

3.2 Morpho-syntactic features

As shown in the previous section, MORPHO-SYNTACTIC 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 morpho-syntax, 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, the most accurate term would be 'morpho-semantico-syntactic' features" (**kibort2010a**).

Both agreement and government require a syntactic constituent as trigger and another constituent as target of morpho-syntactic 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, Kibort's GOVERNMENT covers only morpho-syntactic marking assigned

[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.

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 morpho-syntactic features. A prototypical example of morpho-syntactic 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 morpho-syntactic marking.

Morpho-syntactic attributive construct state marking 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 (personal 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 others), 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 morpho-syntactic 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.

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 (**??**, **??**), 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 **??**), a similar morpho-syntactic 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) Kildin Saami (Uralic; personal knowledge)
tuel' *al'n*
chair\GEN on
'on the chair'

It could be argued that the genitive inflection of 'chair' in example (**??**) is a morphological value of the feature CASE assigned to the dependent noun phrase by the mechanism of GOVERNMENT. But since genitive is the obligatory and only

[2] Bulgarian *brojna forma*

possible marker of the dependent noun in postposition 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 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 adposition phrase), a more appropriate gloss in this construction could in principle be CONSTRUCT. However, since there is no formal difference between the possessive genitive from the genitive assigned by postpositions there are no good arguments to dissociate them into two different morpho-syntactic categories.

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) Kildin Saami (Uralic; personal knowledge)

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

b. Predicative adjective (cf. "predicative 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., dependency inside an adposition 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

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

inflection of attributive adjectives in Russian, for instance, marks the syntactically governed feature STATE simultaneously with the morpho-syntactically governed features NUMBER/GENDER/CASE.

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

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

b. Predicative adjective inflection (cf. "predicative state")[4]
olen' bel
deer white:PRED:M.SG
'the (rein)deer is white'

3.3 An ontology of morpho-syntactic 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 morpho-syntactic features (i.e., features not interfacing with 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 morpho-syntactic feature and which should definitely be added to Kibort's list.

Figure **??** shows the morpho-syntactic 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 MORPHO-SYNTACTIC 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

[4] Note that in Russian the use of the so-called "short adjective" (*bel*) in predicative constructions is highly marked stilistically because it implies a temporary property, which is rather unexpected for the color of a reindeer. Using the "long adjective" even in predicative constructions (*olen' belyj*) is the default. However, the example, which is not ungrammatical, is used here for better comparison to Kildin Saami. On attributive and predicative adjectives in Russian, see in more detail §?? and §??.

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 morpho-syntax (through agreement on adjectives) or in morphology (through the assignment of either grammatical or semantic cases on head nouns). In the following Part II (Typology) of this book, dependent

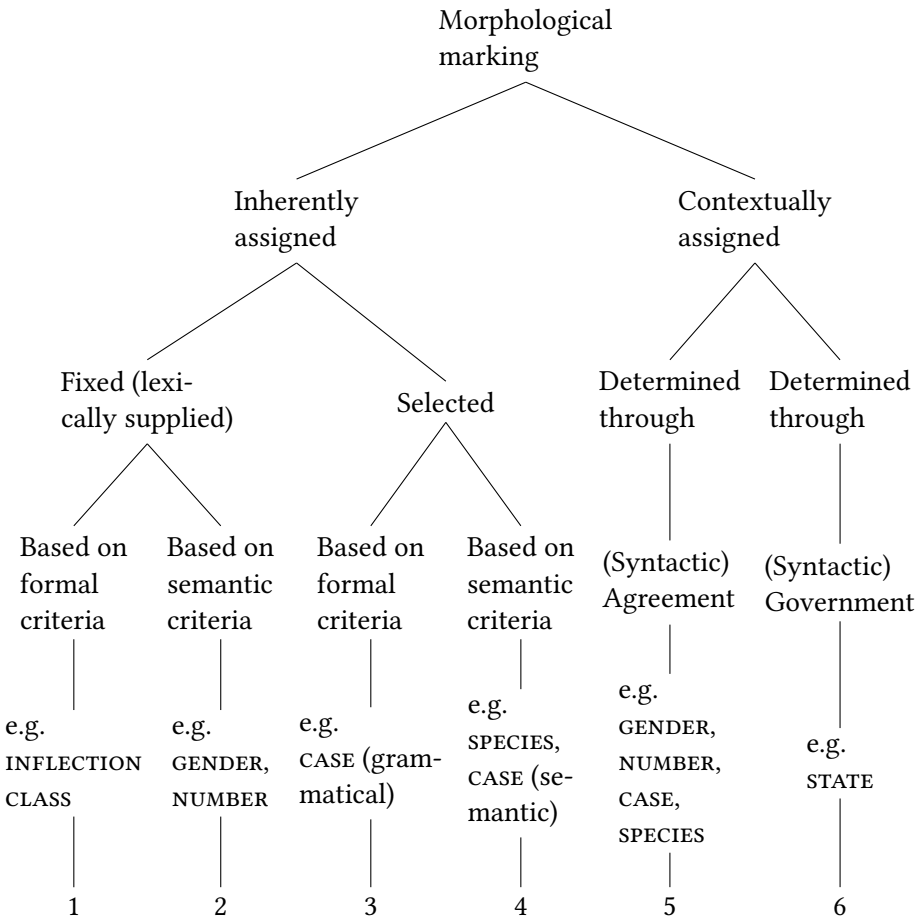


Figure 3.1: An ontology of morpho-syntactic features relevant to the present inventory of noun phrase types (adapted from **kibort2010a kibort2008a** and extended with the feature STATE)

marking STATE will be dealt with in more detail since this type occurs in several languages of the geographical area under investigation.

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 sections is to typologize noun phrases and to present a comprehensive ontology of different syntactic, morpho-syntactic, and morpho-semantico-syntactic attribution marking devices attested in the languages spoken in 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 morpho-syntactic 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 subordinate 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.

Attribution marking Minimally, an attribution marking device will simply license the syntactic structure without licensing any of the constituents as head or dependent, i.e., without ranking single constituents. 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 morpho-syntactic 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 attribution 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).

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 syntactic 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 current chapter. The focus of the ontology presented here is on morphological and morpho-syntactic parameters, especially with regard to the absence or presence of additional attribution marking morphemes, as well as to their kind and

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

syntactic behavior. An overall picture of the ontology of attribution devices relevant to this study is given in Figure **??** at the end of §**??**.

Noun phrase types can also be defined on a polyfunctionality scale with regard to the class of modifying elements: attributive adjectives and other, non-adjectival adnominal modifiers (demonstratives, bare nouns or noun phrases, adposition 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 §**??**).

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 morpho-syntax. In English, for instance, adjectives and clauses behave syntactically differently 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 morpho-syntax 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*** [POSS:3SG.ACC] 'the black one (viz. pencil)'; 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 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 §**??** 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 **??** on page **??** 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* 'purple' and *rosa* 'pink' behave morpho-syntactically differently and do not agree with the modified noun (**schafer2015a**).

Another example of 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 partic-

ular 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.

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 appositional 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 §**??**. 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 nor 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

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

of constituents alone, i.e., that an adnominal modifier and a head noun occur next 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. When an adjective is used as the predicate in English (*the man **is** good, the men **are** good*), the copula is obligatory. However, constituent order may be relevant, too. In English, again, juxtaposed attributive adjectives precede the noun as a rule, whereas predicative adjectives follow it.

Constituent 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 constituent 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 constituent order.

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

4.3 Covert morpho-syntactic construct marking: adjective incorporation

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 a 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 Swedish (and Norwegian) dialects is syntactically and semantically distinguishable from prototypical compounding it is often referred to as ADJECTIVE-NOUN-INCORPORATION (for instance by **sandstrom-etal2003** **dahl2015a** or **julien2005**).

Phonological versus 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 roots, such as *tré*, 'tree', *bắt* 'boat', *bắt-er* 'boats', *bắt-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. Bisyllabic roots or stems, including compounds, by contrast have pitch accent (marked with an additional ` on the second root). Compare *tré-bå̀t-en* 'the wooden boat' or *stór-bå̀t-en* with the noun phrase *bắt-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, for instance, 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

where the attributive adjective occurs obligatorily as a (syntactically) bound morpheme. To prove syntactic boundness one has to show that the adjective cannot occur unbound. In Västerbotten Swedish (and other northern Swedish dialects), for instance, the adjective stem cannot occur unbound unless alternative morpho-syntactic 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 an 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 boundness 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 Adyghe and in Chukchi, Itelmen and in Eskimo-Aleut languages (see the respective sections of Part III (Synchrony); on the typology of adjective incorporation see also **dahl2004a** and **dahl2015a**).

[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-DEF:M.SG] 'the big one (masculine)', *stor-et* [big-DEF:N.SG] 'the big one (neuter)'.

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

4.4 Morpho-semantico-syntactic attribution marking: agreement

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

(6) Lower Sorbian (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**

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 (**??**) 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 **??** 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 morpho-syntactic mechanisms of assigning a head's morphological features to dependent constituents are similar for case and other agreement categories (like gender and number). Consider (**??**) 'to a good person' in Lower Sorbian.

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

Another possible agreement feature beside GENDER, NUMBER and CASE is the feature SPECIES, typical values of which are DEFINITE and INDEFINITE. Consider, for instance, the agreement paradigm of adjectives in Icelandic (Table **??**) 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 corresponding values on the head noun. In fact, adjective agreement paradigms in many languages are different from the corresponding 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**)

		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; personal 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**). 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, for instance, in the Lower Sorbian example (**??**) on page **??** where the adjective suffix *-emu* and the noun suffix *-oju* both mark the dative masculine singular.

Head-driven agreement marking also surfaces in different ways across languages with 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.

(9) Finnish (Uralic; personal 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 "defective" 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 (**??**) and Table **??** with the corresponding paradigm in §**??**.

(10) Danish (Indo-European; personal 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 paradigm (**??**).[7]

(11) Chechen (Nakh-Daghestanian; **nichols1994a**)[8]

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

[7] A similar defective agreement paradigm with only one case distinction is found in the closely related language Ingush, see §**??**. Another, non-related language exhibiting defective agreement is Burgenland Romani, see §**??**.

[8] The paradigm includes only selected forms.

b. *dikaču stega*^{n} [GEN:SG]
c. *dikaču stagana* [DAT:SG]
d. *dikaču staga* [ERG:SG]
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 (**??**) obligatorily agrees with the 3SG possessor 'man'. This type of dependent-driven agreement is usually called POSSESSOR AGREEMENT.[9]

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.

Oroch also exhibits dependent-driven agreement marking by means of possessive affixes on attributive adjectives, which is structurally similar to example (**??**).

(13) Oroch (Tungusic; **malchukov2000**)

[9] Another commonly used term is CROSS-REFERENCE MARKING.

a. *nia aja-ni*
man good-POSS:3SG
'a GOOD man'

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 (**??**), 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 (??) occurs in a possessive-like construction (similar to ??) where the adjective takes the slot of the possessed and is subsequently marked with a possessive agreement suffix.[10] I propose that attributive adjectives in Saliba occur in "headstand" noun phrases and are marked by means of modifier-headed possessor agreement. Unlike in Oroch, however, modifier-headed possessor agreement is the default type of attributive connection of adjectives in Saliba.

4.5 Overt morpho-syntactic construct marking: attributive state marking

Due to a lack of better terminology the feature STATE was earlier defined as assigned through SYNTACTIC GOVERNMENT (in §??). 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.

The term 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:[11]

[10] An alternative account of noun phrase structure in Saliba could claim that a verbal adjective used as an 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.

[11] 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.

- on-head (construct)
- on-dependent (anti-construct)
- neither on-head nor on-dependent (floating construct)
- 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 §?? and §??.

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*[12] on the noun 'house' is to show that "I am the head of a noun phrase and I have a dependent."[13] 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 the Northern 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 (??) and the paradigm in Table ??.

(17) Northern Kurdish (Indo-European; **ortmann2002b**)

[12] The allomorph *-e* appears after consonants.

[13] 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, adposition phrases and verb infinitives.

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'

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

Table 4.2: Paradigm of the Ezafe in NORTHERN KURDISH (**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 morpho-syntactic feature ATTRIBUTIVE in the differentiated forms of the Ezafe in Northern Kurdish. 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) Kildin Saami (Uralic; personal knowledge)

a. Predicative state

i. *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 (**??**), it is marked with the attributive suffix *-es'* if used as modifier (**??**, **??**). 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 in contrast to the Persian construct state, it can be labeled ANTI-CONSTRUCT.[14]

Anti-construct state marking seems not uncommon cross-linguistically, even if Saamic and the Iranian language Northern Talysh (see §**??**) 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 affix", "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".

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.[15] Rather than extending the terminological

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

[15] Even other construct marking devices, such as the linker in Tagalog (**??**) or the construct state marker in Persian (**??**), are often described as "genitives" because they mark possession. Unlike

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 morpho-syntactic device, possessive case additionally specifies a semantic relation (i.e., possession).

4.5.2.2 Anti-construct state agreement marking

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 a licenser 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 labeled ANTI-CONSTRUCT STATE AGREEMENT MARKING in the following.[16]

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

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

a. Attribution

i. ***krasiv-yj*** *mal'čik*
beautiful-ATTR:M.NOM boy(F)
'a handsome boy'

ii. ***krasiv-aja*** *devuška*
beautiful-ATTR:F.NOM girl(F)
'a pretty girl'

prototypical genitives, however, the construct markers in Tagalog and Persian do not constitute dependent marking devices.

[16] 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.

b. Predication (“short”)
   i. *Etot* *mal’čik* ***krasiv***
   DEM:M boy(M) beautiful:M
   ‘this boy is handsome’
   ii. *Eta* *devuška* ***krasiv-a***
   DEM:F tower(F) high-F
   ‘this girl is pretty’
c. Predication (“long”)
   i. *Etot* *mal’čik* ***krasiv-yj***
   DEM:M boy(M) beautiful:ATTR?:M
   ‘this boy is handsome’ (a handsome one)
   ii. *Eta* *devuška* ***krasiv-aja***
   DEM:F tower(F) high-ATTR?:F
   ‘this girl is pretty’ (a pretty one)

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 morpho-syntactic 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.[17] Synchronically, the attributive adjective suffixes in Russian are thus best analyzed as portmanteau suffixes marking anti-construct and head-driven agreement simultaneously.

[17] 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.

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. The semantic difference between the use of "short" versus "long" forms in adjective predication in Russian can be described as an opposition between temporal and permanent properties denoted by the adjective. In fact, the use of the "short" adjective in predicative position – implying a temporary property – is stylistically marked in contrast to the "long" form, which has become the default in contemporary Russian.

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 appositional construction, as the "long" predicative form in (**??**) denoting a permanent property is in contrast to the "short" predicative form in (**??**) denoting a temporal property.[18]

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

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

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

The origin of anti-construct state agreement marking in Russian is dealt with in §**??**. 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 (see §**??**).

[18] 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 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**).

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; **dahl2015a**)

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 (**??**), the attributive adjective in definite noun phrases is obligatorily marked with the "long form" agreement suffix (**??**).

The anti-construct state agreement marking suffixes in the Baltic languages are often described as a definiteness markers. Note, however, that the definite noun 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.[19] 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; **dahl2015a**)

a. *māja*
house
'a house'

[19] Consider, for instance, 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 **??**.

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 **dahl2015a**) that definite noun phrases often show special behavior in languages depending on whether or not they exhibit attributive adjectives (or other modifiers).[20]

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.

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 mokykla* 'college (lit. 'high school')'.[21]

[20] **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 IV (Diachrony) of this study, especially in §**??**.

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

In §??, 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 (see §??). 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 ?? on page ??). 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 morpho-syntactically 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 §??), German thus exhibits a similar type of obligatory anti-construct state agreement marking as Russian. Note, however, that the inherited adjective inflection suffixes are merged to a relatively high degree in Modern German: 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 ??) and definite agreement on adjectives occurs in semantically definite noun phrases (formally marked by the definite marker *der* in Table ??), 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 – traditionally labeled "mixed adjective inflection" (**schafer2015a**) – 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.

LATEX: apply better spac ng 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 adjectives and by *-een* for attributive adjectives.[22]

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**). The present study uses the term nominalization in the latter sense, i.e., a licenser of 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 **??** in §**??**). Adjectives in Mandarin

[22] 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".

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)

are used in attributive position (??), in predicative position (??) and as adverbial modifiers (??).

(24) Mandarin Chinese (Sino-Tibetan; **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 corresponding markers (i.e., nominalizers of nominals) are labeled with quite different 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.

Due to the 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 process yielding a lexical noun ("substantive") as the result of a word class changing operation, i.e., derivation. 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 "appositional

modification" in §??. Attributive nominalization is a special subtype of dependent marking construct state. Similar to the latter, attributive nominalization represents a covert dependent marking morpho-syntactic device and is triggered either by purely syntactic government (as, for instance, anti-construct state marking in Kildin Saami, see §??) or by syntactic government in combination with head-driven agreement (as, for instance, anti-construct state agreement marking in Russian, see §??). 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, for instance, 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.[23]

(25) German (Indo-European; personal 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* $]]$

The intermediate phrasal constituent between the noun phrase (NP) and the adjective is labeled NP', leaving open the question about what constitutes the syntactic head of this phrasal projection.[24]

Note that the attributive marker *der* in (??) is homophonous with the definite marker *der* but clearly has a different function in this construction. For instance, the attributive marker *der* cannot be replaced by a possessive or a demonstrative pronoun and is thus not a marker of definiteness. The (proper) noun phrase *Friedrich der Große*, on the other hand, can be further modified 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, species marking of the whole noun phrase (i.e., in/definiteness) does not affect the attributive nominalizer; consider the following example:

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

[23] 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.

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

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. The agreement pattern in the German epithet construction also show that the nominalizer *der* must not only be be distinguished from the homophonous definite marker, but also from the relativizer *der*. Consider the following examples (cf. also **himmelmann1997**).

(28) German (Indo-European; personal knowledge)
a. * *ein Jagdhund Friedrichs der Große*
b. *ein Jagdhund Friedrichs des Großen*
c. *die Jagdhunde Friedrichs, der seine Sommerresidenz in Potsdam hatte*
d. *die Jagdhunde Friedrichs, den man auch den Alten Fritz nannte*

According to Lehmann (**lehmann1984** cf. also **himmelmann1997**) true relative pronouns represent the syntactic head in relation to 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 allotted 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.

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 species (i.e., 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 were 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.[25]

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 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 separating and linking function simultaneously"[26] by marking the adjective as "physically independent."[27] The articles *ille* in Latin and *tó* in Greek thus have different functions than the homophonous demonstratives/definite markers in that the article nominalizes an adnominal constituent in order to function as attribute of a certain kind. The homophonous 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 less restrictedly in Yiddish.

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

[25] In Himmelmann's **himmelmann1997** terminology, however, the attributive or linking article is a subtype 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.

[26] "[…] zugleich trennende und verbindende Funktion […]"

[27] "[…] physisch selbständig […]"

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 (**??**, **??**). In an emphatic construction and postponed to the head noun, however, the attributive adjective is marked with an article (**??**, **??**) (**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 Modern Greek, where the so-called repeated article also occurs in contrastive focus constructions.

(30) Greek (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 (§**??**), Yiddish (**??**) and Greek (**??**) 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.

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 (epithet constructions) would fall into this category because it is not an agreement target.[28]

In the present survey, however, there are only a few examples of languages with attributive, non-article nominalizers attested, among them Ket (see §**??**) and Dungan (see §**??**) 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 morpho-syntax), articles have an additional semantic component because they undergo agreement.

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, **himmelmann1997** describes attributive articles and other attributive nominalizers as D(eterminer) elements between head and attribute[29]. Illustrating attributive nominalization with examples from several languages, he 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 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

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

not project a noun phrase (see §?? in Part II Typology). Examples of agreement marking "D-Elements" come from Swedish and Albanian.

(31) a. Swedish (Indo-European; personal 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 a nominalizer, the functionally related markers in Swedish (and other languages) are 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 (see §?? for Albanian and §?? for Swedish). From a diachronic point of view, however, these markers clearly originate from very similar attributive nominalizers. Consequently, the grammaticalization path suggested by **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 IV (Diachrony).

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

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.

(32) Hopi (Toreva) (Uto-Aztecan; **whorf1946**)

a. *ca·va*
is_short

b. *pǫ̀yo*
knife

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

According to **whorf1946** both the adjective modifier (which is a stative verb in Hopi) and the noun head alter their phonological shapes depending on whether they are used in predication or as constituents in a noun phrase. Consider the noun phrase in example (**??**) where the modifier *ca·va* 'is short' occurs with a shortened stem form (compared to **??**) and the noun is marked by means of lenition of the word-initial consonant (*pǫ̀yo* 'knife' versus *vǫ̀yo* [knife\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 above-mentioned examples of different types of state markers, the corresponding 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.

(33) Northern Saami (Uralic; personal knowledge)

a. Diminutive derivation
guolli / guolá-š / guolá-ža-t
fish fish-DIM fish-DIM-PL
'fish' / 'little fish' / 'little fishes'

b. Anti-construct state marking ('big')[30]

[30] 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'.

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')[31]
unna *guolá-**š** / guolá-**ža**-t*
small: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 (**??**, **??**). However, diminutive can in fact also be a morpho-syntactic 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 (**??**). However marginal these examples seem to be, diminutive is assigned syntactically on the head by the dependent and thus also belongs to the morpho-syntactic features in Northern Saami.

4.5.4 Neutral attributive state (Linker)

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.

(34) Tagalog (Austronesian; **rubin1994**)

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

b. Attribution (adjective-noun)

[31] 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'.

*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 *na/-ng*.[32] The marker occurs with attributive adjectives (**??** and **??**) but not with predicative ones (**??**).[33]

The two types of adjective attribution in Tagalog (**??** and **??**) are distinguished from each other only by constituent 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. 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 §**??**), it seems rather relevant to characterize this noun phrase type here.

4.6 An ontology of adjective attribution marking devices

§§**??**–**??** of this chapter were aimed at typologizing adjective attribution marking devices. The attested devices described so far belong to the following noun phrase types:

- Juxtaposition
- Incorporation
- Construct state

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

[33] The state marker in Tagalog is polyfunctional in the sense that it also marks attribution of demonstratives, numerals and other modifiers (**himmelmann1997**). See also §**??**.

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

Table ?? on page ?? summarizes the typology presented in §§??–?? and presents short definitions (including bracketed syntactic templates) and an example for each type.[34] 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.[35]

Table ?? on page ?? 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.
- *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.

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

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

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

Figure **??** on page **??** presents a similar ontology in a tree diagram. The order of types (from left to right) is similar to Table **??** (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 morpho-syntactic 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 **??**. 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")

Tables ??–?? on the following pages present definitions and diagrams for the ontology of adjective attribution marking devices used in the present study.[36]

[36] The following type abbreviations are used in these tables: ACAgr=Anti-construct state agreement, AConstr=Anti-construct state, AHDAgr=Appositional 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

ungünstige Positionierung: Since this footnote 1. is too big to appear on one single page, 2. it is divided by 5 pages with tables, and 3. it is important to understand all tables, it is recommendable either to force the footnote to stay together on one page, or to include the footnote into the text (and not be included as a footnote).

Table 4.5: Attested adjective attribution marking devices with definitions. I.

Type	Definition	Syntactic dependency	Commonly used label	Example language
Juxt	Unmarked sequence of constituents; Test: no additional morphemes available in NP	[$_{\text{NP}}$[$_{\text{AP}}$ A] (N)]	Juxtaposition	Komi-Zyrian
Inc	No additional morphemes available in NP, but dep is syntactic compound; Test: dep cannot occur unbound (headless)	[$_{\text{NP}}$ 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	[$_{\text{NP}}$[$_{\text{AP}}$ A (N:ATTR)]]	Ezafe	Kurdish
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	[$_{\text{NP}}$[$_{\text{AP}}$ A] [ATTR] N]	Linker	Tagalog

Table 4.6: Attested adjective attribution marking devices with definitions. II.

Type	Definition	Syntactic dependency	Commonly used label	Example language
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	Skolt Saami
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 duplicated	[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

Table 4.7: Attested adjective attribution marking devices with definitions. III.

Type	Definition	Syntactic dependency	Commonly used label	Example language
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 duplicated, often with reversed constituent order	[$_{NP}$[$_{NP'}$ 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]($_{PSR}$N)]	Possessive-like attribute	Saliba

<table>
<tr><td></td><td colspan="6">Syntactic source</td></tr>
<tr><td rowspan="3">Locus</td><td colspan="4">Government [+GOV] [±(secondary)AGR]</td><td colspan="2">Agreement [−GOV] [+(primary)AGR]</td></tr>
<tr><td colspan="4">Syntactic pattern</td><td colspan="2">Syntactic pattern</td></tr>
<tr><td>[±AGR]</td><td>Embedded</td><td>Non-embedded</td><td>Incorporated</td><td>Embedded</td><td>Non-embedded</td></tr>
<tr><td>no marking</td><td></td><td></td><td>Juxtaposition
[NP A (N)]</td><td>Incorporation
[NP A *(N)]</td><td></td><td></td></tr>
<tr><td rowspan="2">floating</td><td>[−AGR]</td><td></td><td>Linker
[NP A [ATTR] *(N)]</td><td></td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>[+AGR]</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">dep-marking</td><td>[−AGR]</td><td>Nominalization
[NP[NP' A:NMLZ] (N)]</td><td>Anti-Construct State
[NP A:ATTR (N)]</td><td></td><td rowspan="2">Appositional Head-Driven Agreement
[NP [NP' A:AGR] (N)]</td><td rowspan="2">Head-Driven Agreement
[NP A:AGR (N)]</td></tr>
<tr><td>[+AGR]</td><td>Article
[NP[NP' A:NMLZ:AGR] (N)]</td><td>Anti-Construct Agreement
[NP A:ATTR:AGR (N)]</td><td></td></tr>
<tr><td rowspan="2">head-marking</td><td>[−AGR]</td><td></td><td>Construct State
[NP A *(N:ATTR)]</td><td></td><td rowspan="2">Modifier-headed Possessor Agreement
[NP[PSD A:POSS:AGR] (PSRN)]</td><td rowspan="2"></td></tr>
<tr><td>[+AGR]</td><td></td><td></td><td></td></tr>
</table>

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 cells filled in gray color are marked for logically impossible types; other cells were left open because the corresponding types were not detected in noun phrases with attributive adjectives; glosses and tags are: 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

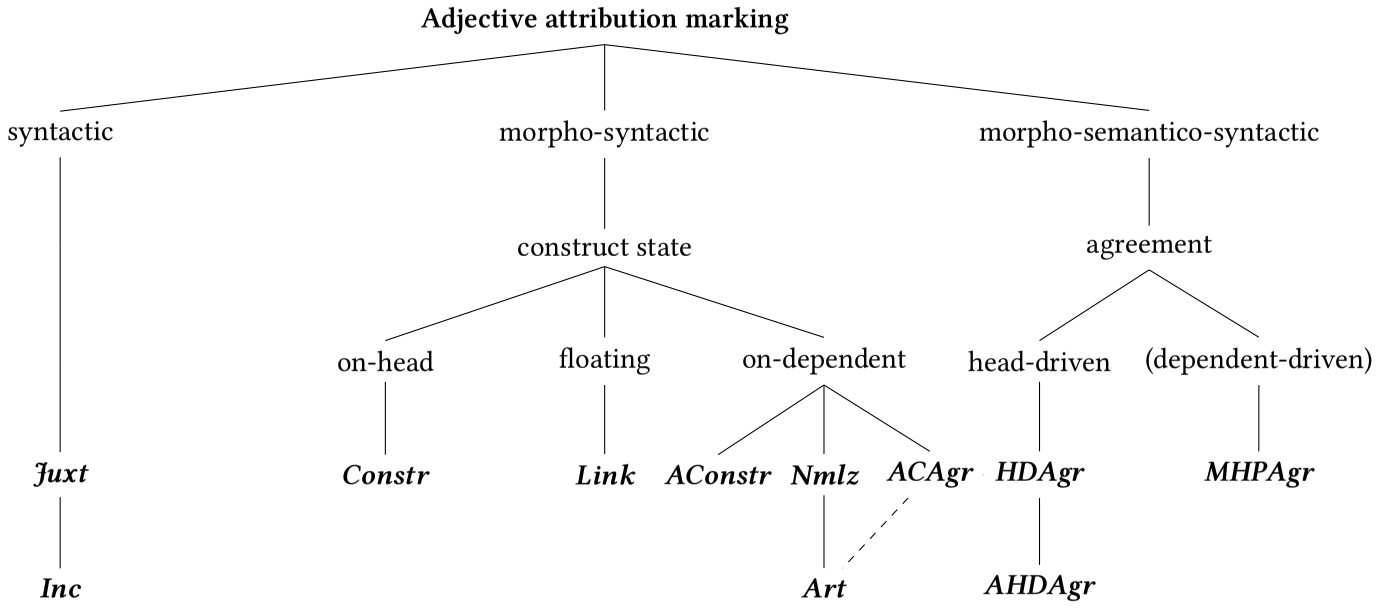


Figure 4.2: Ontological tree of attested adjective attribution marking devices; type abbreviations are: ACAgr=Anti-construct state agreement, AConstr=Anti-construct state, AHDAgr=Appositional head-driven agreement, Art=Attributive article, Constr=Construct state, HDAgr=Head-driven agreement, Inc=Incorporation, 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 with respect to polyfunctionality and regarding to the class of attributed elements beyond adjective attribution: attributive adjectives may or may not be used in similar noun phrase structures like other adnominal modifiers (such as demonstratives, adposition phrases, clauses, etc.).

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 adposition 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.

In example (**??**) from Mandarin Chinese, the anti-construct state marker *de* illustrates a highly polyfunctional attribution marking device. It licenses adjectival

(??), nominal (??) and verbal attributes (??).[1]

(1) Mandarin Chinese (Sino-Tibetan; **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
'the book I bought yesterday' *wǒ zuótiān mǎi* ***de*** *shū*
1SG buy yesterday ATTR book

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 always 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 focus or stress – resembling contrastive focus marking – is put on the denoted property, like in *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 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
      'a/the mango that Jojo bought'

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

(4) Persian (Indo-European; **mahootian1997**)
   a. Adposition 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 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.

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

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

(5) Västerbotten Swedish (Indo-European; **gil2005**)

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

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

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 Chinese (**??**) and Minangkabau (**??**)

- **Moderately differentiating languages** using polyfunctional devices for attribution of two syntactic categories, for instance:
  - adjectives and relative clauses, as in Tagalog (**??**)
  - possessor nouns and adjectives, as in Västerbotten Swedish (**??**) and Persian (**??**)
- **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 adnominal adjectives and relative clauses).[4] No languages of the "weakly differentiated" type are known to occur in the northern Eurasian area. Figure **??** illustrates the polyfunctionality

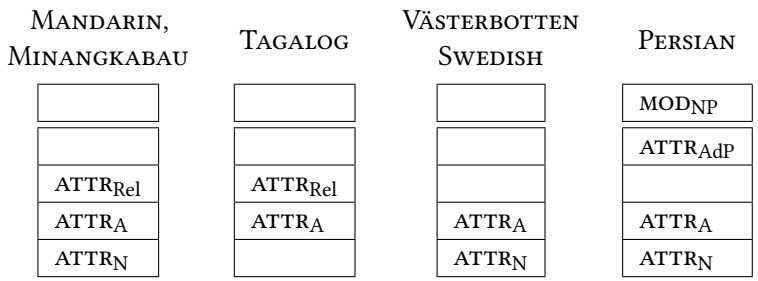


Figure 5.1: Functional maps for modification markers: the anti-construct state marking in Mandarin Chinese and juxtaposition in Minangkabau, the linker in Tagalog, adjective incorporation in Västerbotten Swedish and construct state marking in Persian

of modification markers in the languages mentioned in this chapter.[5] The true

[4] Note that English is not coded as "moderately differentiating" by **gil2005** although juxtaposition can be used polyfunctionally 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.

attributive 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 **??**. The cell extending upwards shows the additional function of the marker as licenser 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]

‼Andi: Chinese names usually have the opposite order

‼Andi: link is defective, wrong format, perhaps it is sufficient to cite the actually published work by Foley?

(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, all categories 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

[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 **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.

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 (**??**), nominal (**??**) and verbal attributes (**??**). In addition, the marker *ve* in Lahu is used as a nominalizer (**??**) and as a focus marker (**??**).[7]

(7) Lahu (Sino-Tibetan; **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 **??**. 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 **??**. 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 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 (Synchrony) 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 the Caucasus), over central, northern, and northeastern Asia (including the whole of Siberia, the adjacent parts of northern Mongolia) 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 inventory of languages will be followed here. However, the present survey strictly follows the geography of northern Eurasia and consequently also includes Siberian Yupik Eskimo, Ainu, the Sino-Tibetan language Dungan, and some Semitic languages.

6.1 The languages of northern Eurasia

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

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

Even though some of these genealogical units have been assumed to combine to larger stocks (such as Altaic, Chukotko-Kamchatkan, North Caucasian and others) 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 a 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 taxon 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 corresponding geographic coordinates have either been taken from **walsOnline2013** or were included using the language coordinates provided by **AUTOTYP** or on Ljuba Veselinova's website.[2] For some languages, which were 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 – like in the maps used by typological surveys of the EU-

[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

ROTYP-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-continuos languages seems rather inadequate for drawing such isoglosses.[4]

6.3.2 Data points for type coding

In several languages more than one default attribution marking device occurs, for example in Albanian (see §??) 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 in Chuvash (see §??), where attributive adjectives are normally juxtaposed but can alternatively be marked for attributive nominalization in contrastive focus constructions: Chuvash *Juxt(Nmlz)*.

Square brackets are used for languages where the occurrence of a given type of attribution marking device seems even more restricted or if the device's characteristics remain uncertain due to inadequate data. Consider for example Turkish (see §??), 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

[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 diverse as Romance languages.

[5] Type abbreviations are explained in the Appendix.

types (juxtaposition, agreement, attributive state, incorporation) in Figure **??**. 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 are actually coded in the language sample in the Appendix.

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 taxon, this was always done in favor of diversity rather than uniformity. One taxon 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*) (§**??**).

The choice to let the maps illustrate the highest possible diversity instead of displaying a genealogically and geographically balanced picture is justified 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 IV (Diachrony) will inspect this diversity form a diachronic perspective.

7 Adjective attribution marking in the languages of northern Eurasia

The following chapter contains an overall survey of adjective attribution marking devices which occur in the languages of northern Eurasia. For each genealogical unit, both the prototypical and the known minor noun phrase type(s) will be characterized and illustrated with examples. A complete list of adjective attribution marking devices in over 200 single languages considered for the present survey is found in a table starting on page **??** in the Appendix. The geographic spread of the different noun phrase types is shown on several maps starting on page **??** in the Appendix.

7.1 Eskimo-Aleut (Central Siberian Yupik)

Whereas most languages of the Eskimo-Aleut family are spoken on islands in the Bering Strait or on the North American continent, a few varieties of the Yupik subbranch of Eskimo can be localized to north-easternmost Siberia. But only one of these languages, Central Siberian Yupik, is still spoken (**salminen2007**).

In Central Siberian Yupik, only one adjective attribution marking device is attested:

- incorporation.

Adjective incorporation in Central Siberian Yupik Items that correspond to property-denoting words in other languages ("adjectives") are phonologically bound nominal roots in Central Siberian Yupik. Adjectival modification is thus expressed by means of polysynthetic morphology and can be characterized as adjective incorporation according to the ontology presented in Part II (Typology).

(1) Central Siberian Yupik (**de-reuse1994**)
   a. *qawaagpag-**rukutaagh**-**ghllag**-Ø*
      legendary_large_bird-huge.NOUN-big.NOUN-ABS
      'huge big (legendary large) bird' (54)

b. *mangteghagh-**ghllag**-lgu-uq*
house-big.NOUN-have.NOUN-IND(3s)
'He has a large house.' (55)

c. *mangteghagh-**ghrugllag**-ngllagh-yug-nghit°e-unga*
house-big.NOUN-make.NOUN-want_to.VERB-NEG-IND(1s)
'I did not want to make a large house.' (56)

7.2 Chukotkan

The Chukotkan language family (aka Chukchi-Koryak) consists of two branches. The first branch, Chukchi, is represented by only one language, Chukchi, which has the same name as the branch itself. The second branch, Koryak-Alutor, is represented by the two languages Alutor and Koryak proper. A third branch, Kerek, is probably extinct (**salminen2007**) and consequently not considered here.

Constituent order inside the noun phrase of Chukotkan languages is strictly head-final. Adjective attribution marking is also similar in all Chukotkan languages. Two types are attested:

- incorporation
- head-driven agreement.

7.2.1 Chukchi

Adjective incorporation in Chukchi The use of the bound adjective morpheme in the polysynthetic structure (similar to Yupik) is illustrated in the following examples.[1]

(2) Chukchi (**skorik1960**)

a. ***elg**-ə-qoranə*
white-ə-deer:ABS.SG
'white reindeer'

b. ***elg**-ə-qorat*
white-ə-deer:ABS.PL
'white reindeer (pl.)'

[1] The vowel -ə- in these and the following examples is epenthetic.

7.2.2 Koryak

Adjective incorporation in Alutor Similar to Chukchi, adjective incorporation is the default adjective attribution marking device in Alutor.

(3) Alutor (**nagayama2003**)

a. ***meŋ**-ə-rara-ŋa*
big-ə-house-ABS.SG
'large house'

b. ***meŋ**-ə-rara-wwi*
big-ə-house-ABS.PL
'large houses'

Head-driven agreement in Chukchi and Alutor Whereas adjective incorporation is the default and unmarked type of adjective attribution marking in Alutor and Chukchi, several descriptions of the Chukotkan languages mention that adjectives can also occur in an unbound form (for Alutor, see **nagayama2003** for Chukchi, see **skorik1960** and **comrie1981**). As unbound morphemes, adjectives take the stative marker *n-* as well as agreement markers for person, number and case.

(4) a. Chukchi (**skorik1960**)

i. ***n-ilg-ə-qin-Ø*** *qoranə*
STAT-white-ə-3SG deer:ABS.SG
'white reindeer'

ii. ***n-ilg-ə-qine-t*** *qorat*
STAT-white-ə-3-PL deer:ABS.PL
'white reindeer (pl.)'

b. Alutor (**nagayama2003**)

i. ***n-ə-meŋ-ə-qin*** *rara-ŋa*
STAT-ə-big-ə-ABS:3SG house-ABS.SG
'large house'

ii. ***n-ə-meŋ-ə-laŋ*** *rara-wwi*
STAT-ə-big-ə-ABS:3PL house-ABS.PL
'large houses'

The number/person/case-agreement suffixes of adjectives and the suffixes marking possessive inflection of nouns belong to one and the same paradigm. Consequently, one could also interpret the Alutor and Chukchi data as another instance of modifier-headed possessor agreement (as in Oroch, described in §**??**). If so, the examples in (**??**) should be translated literally as 'reindeer's whiteness', 'house's bigness'. An analysis avoiding syntactic dependency reversal between noun and adjective (**malchukov2000**), however, is preferred here for two reasons: the first reason is the constituent order inside the noun phrase. The assumed head shift to a modifier-headed possessor agreement construction would violate the otherwise strictly head-final constituent order rule in Alutor and Chukchi.

The other reason arguing against syntactic head shift between noun and adjective is that in order to use non-incorporating constructions as in the examples in (**??**), the adjective is first transformed into a stative verb by means of a verbalizing prefix (*-n*, glossed as STAT in example **??**).

The verbalizer together with the agreement affix is sometimes glossed as an adjectivizing circumfix (ADJZ>-…-<ADJZ:AGR), for instance in Nagayama's (**nagayama2003**) grammatical description of Alutor. The given noun phrase type should then perhaps be analyzed as attributive state marking (as in Russian, see §§**??**, **??**). Unlike in Russian, however, the same agreement marking as in attributive constructions shows up on predicates as well.

(5) Alutor (**nagayama2003**)
   a. ***n*-ə-*tur*-*iɣəm***
      STAT-ə-**young**-1SG
      'I'm young'

Consequently, an analysis of adjective attribution marking in Alutor and Chukchi as belonging to the state-marking type is rejected.

The semantic difference between the two constructions, with adjective incorporation on the one hand and head-driven agreement marking on the other hand, is not clear. Whereas adjective incorporation is often described as the main or even only possible type (for Chukchi, see **kampfe-etal1995**), **kibrik-etal2000** state that this type indicates the corresponding quality or property as referring to background information in Alutor.

The following example from Chukchi, on the other hand, indicates that the non-incorporated adjective is used in an emphasized construction. Sentence (**??**) was elicited by Vladimir Nedjalkov (**rijkhoff2002**) in order to find examples of multiple modifiers in one noun phrase, which seems to be avoided by speakers of Chukchi. In sentence (**??**) with the incorporated adjective, the speaker simply

left out the demonstrative when translating into Chukchi.

(6) Chukchi (**rijkhoff2002**)

a. *əngena-t ngəroq **n-ilg-ə-qine-t** qora-t*
this-PL three STAT-white-ə-3-PL deer-PL
'these three white reindeer'

b. *ətlon ga-twetcha-twa-len ga-ngəron-**elg**-ə-qaa-ma*
3SG PFCT-stand_up-be-3SG COM-white-ə-deer-COM
'He stood next to (these) three white reindeer.'

bogoras1922 states that the circum-positioned marker of the unbound adjective "sometimes corresponds to the definite article or designates an object as referred to before." The unbound adjective, on the other hand, can only occur in absolutive case which is inherently connected to semantic definiteness (**dunn1999**).

7.3 Kamchatkan

The only surviving member of the Kamchatkan language family is Itelmen (aka Western Kamchadal) (**salminen2007**).

The only attested type of adjective attribution marking in Itelmen is:[2]

- anti-construct state agreement.

Anti-construct state agreement in Itelmen Constituent order inside the noun phrase of Itelmen is head-final. Adjectives form a class that is clearly syntactically distinguished from nouns: unlike the latter, adjectives are never represented by their root morphemes alone. Unlike verbs, which take TAM markers, adjectives take adjectival morphology and are licensed either by an attributive or predicative (adverbal) suffix (**volodin1997 georg-etal1999**).

(7) Itelmen (**volodin1997**)

a. Attributive state of adjectives
*thun-**lah***
dark-ATTR
'dark'

[2] According to Volodin (**volodin1997**), a few adjectives (among them Russian loan adjectives) occur in juxtaposition.

b. Predicative state of adjectives
*thun-**k***
dark-PRED
'(is) dark'

Since attributive adjectives also agree in case (though restricted to instrumental case), the noun phrase type can be characterized as anti-construct state agreement, structurally similar to the type found in Russian. Consider the following example.[3]

(8) Anti-construct state agreement in Itelmen (**georg-etal1999**)
Kəmma çasit t'-nu-qz-al-kiçen ***teŋ-lan'ļ*** ***thalthe-ļ***, *min kn-anke*
1SG now 1SG>-eat-IPFV-FUT-<1SG good-ATTR:INS meat-INS REL 1SG-DAT
t-zapasa-qzo-çen.
1SG-keep-IPFV-3SG-PRTC
'Now I will eat the good meat which I kept for you.'

7.4 Nivkh

Nivkh (aka Gilyak) is an isolated language spoken in the far east of the Eurasian continent on Sakhalin Island in easternmost Russia (**salminen2007**).

The only type of adjective attribution marking attested in Nivkh is:

- head-driven agreement.

Head-driven agreement in Nivkh Property words in Nivkh are verbal roots. As modifiers in noun phrases these adjectival verbs occur to the left of the head noun in a construction which is sometimes described as a polysynthetic structure (cf. **gruzdeva1998 jakobson1971** quoted by **rijkhoff2002**). The reason for analyzing adjectives in Nivkh as being incorporated into the modified noun is the phonological boundedness of the constituents evidenced by regular alternations in the initial segments of the noun stem **gruzdeva1998**

(9) Nivkh (**gruzdeva1998**)
a. *tu 'lake'*

[3] Note that the shape of the state marking suffix *-lan'ļ* (← *-lah-ļ*) is the result of a regular morphophonological process (**georg-etal1999**).

b. *pily-du*
be_big-lake
'large lake'

In her sketch grammar of Nivkh, however, **gruzdeva1998** writes adjectival words consistently as morphologically unbound words.[4]

Interestingly, the phonological stem alternation rules also apply to the plural inflection of nouns and their adjectival attributes by means of reduplication. The reduplicated stem of the participle *t'osḳ* in (**??**) 'destroyed' is therefore realized as *-zosḳ*.

(10) Nivkh (Ekaterina Gruzdeva, p.c.)

a. *tuin* ***t'osq-mu*** *hum-d'*
here break.PTCP-boat be-IND
'there is a destroyed boat here'

b. *tuin* ***t'osq~zosḳ-mu-ɣu*** *hum-d'[-ɣu]*
here break.PTCP~PL-boat-PL be-IND[-PL]
'there are destroyed boats here'

Note that the number agreement of the attributive forms of adjectives by means of reduplication is archaic. According to Ekaterina Gruzdeva (p.c.), attributive adjectives practically never reduplicate any more. Examples of reduplicating adjectives are, however, included in the older grammar by Panfilov (**panfilov1965**).

7.5 Ainu

Ainu is an isolate spoken on Hokkaido Island in northern Japan.

The only type of adjective attribution marking attested in Ainu is:

- juxtaposition.

Juxtaposition in Ainu Ainu does not exhibit morphological differences between adjectives and verbs (**refsing1986**). Words expressing states (**??**) or properties (**??**) in Ainu are best described as stative verbs. They form a subclass of intransitive verbs and are only semantically distinguished from verbs denoting an action

[4] For instance *čuz pitɣy-Ø* [new book-NOM] (19), *kyla n'iɣvṇ* [high man] (33), *pila eri* [big river] (38).

(**refsing1986**). As modifiers of a noun, these property words are juxtaposed to the left.

(11) Ainu (Shizunai) (**refsing1986**)

a. "State adjective"
mokor *cep*
sleep fish
'a sleeping fish' (141)

b. "Quality adjective"
pirka *cep*
be_good fish
'a fine fish' (142)

7.6 Japanese

The noun phrase structure in Japanese, an isolated language, is strictly head-final. Two types of adjective attribution marking devices are attested:

- juxtaposition
- anti-construct state marking.

Juxtaposition in Japanese Two distinct lexical classes of words describe the state that an entity is in. Verbal adjectives belong to the first class. These adjectives are distinguished from stative verbs by the adjectivizer suffix *-i*. Used as predicates, the adjectivized verbs marked with *-i* follow the noun but do not require any copula. Attributive adjectives, on the other hand, are juxtaposed to the left of the modified noun.

(12) Verbal adjectives in Japanese (**backhouse1984**)

a. Adjective predication
kono rombun=wa ***naga-i***
this article-TOP long-ADJZ
'This article is long.'

b. Adjective attribution
naga-i *rombun*
long-ADJZ article
'long article'

Since the adjectivizer suffix *-i* simply marks stative verb roots as (attributive and predicative) adjectives, it is not considered an attribution marking device. Hence, the class of verbal adjectives in Japanese is merely attributed by juxtaposition. Constituent order is crucial for differentiating attributive from predicative adjectives.[5]

Anti-construct state in Japanese Unlike "verbal adjectives", which were described in the previous section, the few members of the second adjectival subclass, i.e., "nominal adjectives" require a special attributive form marked by the invariable attributive suffix *-na*.

(13) Japanese (**pustet1989**)

a. Attribution: verbal adjective
waka-i *hito*
young-ADJZ person
'a young person'

b. Attribution: nominal adjective
kirei-na *hito*
beautiful-ATTR person
'a beautiful person'

Note that the word class boundary between nominal adjectives and nouns in Japanese is not always clear because some words take either the noun attribution marker *-no* (**??**) or the adjective attribution marker *-na* (**??**) when modifying a noun. The arbitrary behavior of attribution marking of nouns and nominal adjectives in Japanese indicates the continuous nature of these two word classes in this language (**pustet1989**).

(14) Japanese (**pustet1989**)

a. Noun attribution
wazuka-na *okane*
little-ATTR money
'little money'

[5] Note that the description if the suffix *-i* as an adjectivizer is simplified here. There is also overlap with TENSE MARKING, cf. *rombun-wa naga-i* [PRS] 'the article is long' versus *rombun-wa naga-kat-ta* [PST] 'the article was long'.

b. Adjective attribution
wazuka-no *okane*
little-ATTR money
'little money'

7.7 Korean

Korean is an isolated language spoken on the Korean peninsula in northeastern Asia. The only type of adjective attribution marking attested in Korean is:

- anti-construct state marking.

Note, however, that Korean does not have a distinct class of adjectives but adjectival notions are expressed by verbs.

Anti-construct state in Korean The constituent order in the noun phrase of Korean is strictly head-final. Modifying "property words" are verbs equipped with a special attributive suffix *-(u)n* (**martin-etal1969**).

(15) Korean (**chang1996**)
a. i. *i* ***ppalka-n*** *chayk*
this be_red-ATTR book
ii. *i* ***ppalka-n*** *chayk i*
this be_red-ATTR book SUBJ
'this red book'
b. i. *ce* ***khu-n*** *namwu*
that be_big-ATTR tree
ii. *ce* ***khu-n*** *namwu lul*
that be_big-ATTR tree OBJ
'that big tree'

7.8 Sino-Tibetan (Dungan)

The Sino-Tibetan language family is represented in northern Eurasia only by one language, Dungan (aka Dunganese), which is a Gansu variety of Chinese spoken in the Kyrgyz Republic in Inner Asia (cf. **yuo2003 kalimov1968**).

Two types of adjective attribution marking are attested in Dungan:

- juxtaposition
- attributive nominalization.

Juxtaposition in Dungan Adjective attribution marking in the unmarked noun phrase in Dungan is characterized by juxtaposition. Hereby, the adjective either precedes or follows the noun.

(16) Dungan (**kalimov1968**)

a. ***da*** *fonzy*
big house

b. *fonzy* ***da***
house big

'large house'

Attributive nominalization in Dungan A second noun phrase type with the adjectival modifier marked by a suffix *-di*^{1} occurs in Dungan as well. Whereas juxtaposition constitutes the general and unmarked type of adjective attribution marking, the attributive suffix *-di*^{1} seems to be much more restricted and occurs for example in connection with a comparative (**??**) or negated attribute (**??**).

(17) Dungan

a. Negated attribute (**zevachina2001**)
gubuə^{3} *bu*^{1} ***da***^{3}***-di***^{1} *gun*^{1}*fu*^{1}
went NEG big-ATTR time

'Not much (lit. 'not big') time passed.'

b. Comparative attribute (**kalimov1968**)
da-ščer-di *fonzy*
bigger-COMPAR-ATTR house

'a somewhat larger (i.e., different) house'[6]

The marker *-di*^{1} is clearly cognate with the functionally similar nominalizer *-de* in Mandarin Chinese (cf. example **??** in §**??**). In Dungan, however, *-di*^{1} is sometimes also described as a marker of predicative adjectives, as in (**??**).

[6] Note that the quoted transcriptions of the two authors differ from each other.

(18) Attributive nominalization in Dungan (**zevachina2001**)
*žy³gə¹ mə¹mə² **gan¹-di¹***
this bread stale-ATTR
'This bread is STALE (i.e., different).'

zevachina2001 labels the function of the marker as an "emphasizing-predicative". But looking at her other examples it becomes obvious that *-di¹* does not mark predicative adjectives but rather nominalized attributive adjectives.

(19) Attributive nominalization in Dungan (**zevachina2001**)
*žy³gə¹ fu¹ bu¹cy¹ **xun¹-di¹**, **zy²-di¹***
this book NEG red-ATTR bordeaux-ATTR
'This book is not RED, but BORDEAUX.'
(lit. 'This book is not a red one, but a bordeaux one.')

The nominalizing function of the suffix is also described by **kalimov1968**

(20) Attributive nominalization Dungan (**kalimov1968**)
***ščin-di** gŭjdixyn*
new-ATTR expensive
'The new (one) is expensive.'

Attributive marking with the suffix *-di¹* in Dungan needs to be investigated in more detail, especially in connection to constituent order. The head-initial structure seems to be used in order to emphasize the property denoted by the adjective.

However, according to the descriptions of Dungan taken into account here (i.e., **kalimov1968** and **zevachina2001**), the language exhibits two adjective attribution marking devices: juxtaposition and attributive nominalization by means of the article *-di¹*. While juxtaposition (with the order adjective-noun) seems to be the unmarked type, attributive nominalization is restricted to certain pragmatically marked constructions.

7.9 Mongolic

The Mongolic language family consists of five branches (cf. **salminen2007**). The core branch, Mongolian, includes the languages Kalmyk, Khalkha, Khamnigan Mongol, and Oyrat (aka Oirat). Kalmyk is spoken in easternmost Europe (in the Republic of Kalmykia of the Russian Federation). The other Mongolian languages are all spoken in Inner Asia, along with Dagur which belongs to a satellite branch

of the Mongolic family. Languages of the remaining three satellite branches of Mongolic are not considered here since they are all spoken outside the northern Eurasian area.

With regard to their principal noun phrase structure, all Mongolic languages of northern Eurasia exhibit the inherited Proto-Mongolic features, including strictly head-final constituent order and juxtaposition of attributive adjectives ("adjectival nouns") as the only attribution marking device.

Note, however, that adjectives in Mongolic languages do not differ formally from regular nouns but are distinguishable from the latter only by their syntactic behavior and specific derivational patterns (cf. **janhunen2003b** for Proto-Mongolic and **svantesson2003** for Khalkha).[7]

The only type of adjective attribution marking attested in Mongolic languages of northern Eurasia is:

- juxtaposition.

7.9.1 Mongolian

Juxtaposition in Khalkha The only attested adjective attribution marking device in the languages of the Mongolian branch of Mongolic is juxtaposition, similar to the following example.

(21) Khalkha (**svantesson2003**)

a. *sayin nom*
good book
'good book'

b. *sayin nom-uud*
good book-PL
'good books'

7.9.2 Monguor, Moghol, Dagur

The only attested adjective attribution marking device in the languages of the Monguor, Moghol and Dagur branches of Mongolic is juxtaposition (**slater2003**;

[7] In the two Mongolic languages Moghol (spoken in Afghanistan) and Mangghuer (spoken in China) there is a distinct class of adjectives (cf. **weiers2003** for Moghol and **slater2003** for Mangghuer). However, these languages are not considered since they are spoken outside the northern Eurasian area.

weiers2003; **tsumagari2003**), similar to example (??) from Khalkha Mongolian.

7.10 Tungusic

The Tungusic language family (aka Manchu-Tungus) comprises several single languages belonging to the three branches North Tungusic, Amur Tungusic and Manchu, all spoken in southern Siberia (Russia), northern Mongolia and northern China.

The constituent order inside the noun phrase in all Tungusic languages is relatively strictly head-final. In several Tungusic languages, attributive adjectives ("adjectival nouns") are simply juxtaposed with the modified noun. This type is also mentioned as being prototypical of adjective attribution marking devices in Tungusic languages (e.g., **sunik1968a kormusin2005**). However, several other types occur as well. The following adjective attribution marking devices are attested in Tungusic:

- juxtaposition
- head-driven agreement
- attributive nominalization
- modifier-headed possessor agreement.

7.10.1 North Tungusic

Languages belonging to the northern branch of Tungusic are Even (aka Lamut), Evenki (aka Oroqen in China), Negidal and Solon (aka Ewenke in China).

The major North Tungusic languages, Even and Evenki, deviate from the Tungusic prototype and exhibit head-driven agreement as their general type (**malchukov1995 bulatova-etal1999**). Attributive nominalization and modifier-headed possessor agreement occur in these two languages as well, even though these devices are restricted to specially marked noun phrase types.

Head-driven agreement in Even According to **malchukov1995** the occurrence of head-driven agreement marking of adjectives in Even is determined by discourse-

pragmatic factors: attributes in the rhematic (focus) position always agree with their heads, whereas agreement is optional in non-focus positions **malchukov1995**[8]

(22) "Attribute raising agreement" in Even (**malchukov1995**)

a. Juxtaposition
(A N-NUMBER-CASE)
Eŋi *beji-l-bu* *emu-re-m.*
strong man-PL-ACC bring-NONFUT-1SG

b. Incomplete head-driven agreement
(A-NUMBER N-NUMBER-CASE)
Eŋi-l *beji-l-bu* *emu-re-m.*
strong-PL man-PL-ACC bring-NONFUT-1SG

c. Complete head-driven agreement
(A-NUMBER-CASE N-NUMBER-CASE)
Eŋi-l-bu *beji-l-bu* *emu-re-m.*
strong-PL-ACC man-PL-ACC bring-NONFUT-1SG

'I have brought back only strong men.'

malchukov1995 describes the attributive agreement patterns in Even in a hierarchical way: the adjective modifier can agree in all morphological features of the head-noun (**??**) or just in number (**??**). Juxtaposition is also possible but restricted to adjectives in non-focus position (**??**).

Attributive nominalization in Even (I) The "attribute raising agreement" illustrated in the previous section (§**??**) can be extended with a fourth step, specifically with adjective attributes marked by the "restrictive" (i.e., contrastive focus) marker *=takan/=teken* (here glossed as a nominalizer).

(23) Even (**malchukov1995**)

a. ***Eŋi-l-bu=tken*** *beji-l-bu* *emu-re-m.*
strong-PL-ACC=NMLZ man-PL-ACC bring-NONFUT-1SG

b. * ***Eŋi=tken*** *beji-l-bu* *emu-re-m.*
strong=NMLZ man-PL-ACC bring-NONFUT-1SG

'I have brought back only strong men.'

[8] According to **malchukov1995** regular head-driven agreement occurs as the default type of adjective attribution marking only in literary Even and hence in prescriptive grammars. This does not reflect, however, the actual language use.

Attributes marked as "restrictive" obligatorily agree with the head noun (**malchukov1995**). Noun phrases marked by means of *=takan / =teken* thus resemble the attributive nominalization type, i.e., the attribute is marked as a syntactically complex constituent (i.e., as an embedded complement to the head noun) by means of nominalization.

Attributive nominalization in Even (II) A second attributive nominalization strategy by means of the possessive suffix 3rd person singular (in "determinative" function; here glossed as a nominalizer) is attested in an investigation of the non-possessive use of the possessive marker in different Turkic and Tungusic languages (**benzing1993b**).

(24) Even (**benzing1993b**)
*hagdiŋata-**n** orolcemŋā*
oldest-NMLZ reindeer_herder
'the OLDEST reindeer herder'

According to **benzing1993b** the "determinative" suffix *-n* (⇐ POSS:3SG) can be used as a marker of contrastive focus in Even.

Modifier-headed possessor agreement in Evenki Evenki follows the general Tungusic rule of head-final constituent ordering inside the noun phrase. In constructions emphasizing the property denoted by the attributive adjectives, however, the unmarked adjective-noun order can be reversed. In these constructions, the adjective is obligatorily equipped with the possessive suffix 3rd person (singular or plural).

(25) Evenki (**bulatova-etal1999**)
a. ***aja*** *bəjə*
good man
'good man'
b. *bi: bəjə **aja-βa:-n*** *sa:-m*
1SG man good-ACC-POSS:3SG know-1SG
'I know the GOOD man'

According to **bulatova-etal1999** the phrase final adjective 'good' marked with the possessive suffix is used as a true possessive noun in (**??**) and they translate the example like this: 'I know the man's goodness'. This construction, however,

is similar to the modifier-headed possessor agreement described for Oroch (**??**) and Udege (**nikolaeva-etal2001**).[9]

7.10.2 Amur Tungusic

The Amur (aka South) branch of Tungusic consists of five languages. According to **salminen2007** however, it is better to assume two separate subbranches, one of them comprising Udege and Oroch and the other comprising Nanay (aka Hejen in China), Ulcha and Orok (aka Uilta).

7.10.2.1 Oroch-Udege

Head-driven agreement in Udege Head-driven agreement in Udege is restricted to the feature NUMBER. Morphologically plural head nouns obligatorily trigger plural marking on the attributive adjective.

(26) Udege (**nikolaeva-etal2001**)
*uligdig'a-**ŋku** moxo-ziga bi-si-ti*
beautiful-PL cup-PL be-PST-3PL
'There were beautiful cups.'

Modifier-headed possessor agreement in Oroch Similar to Evenki from the northern branch of Tungusic, the Udege-Oroch languages from the Amur branch exhibit modifier-headed possessor agreement. Oroch examples for this type of adjective attribution marking have already been discussed in §**??** but will be repeated here.

(27) Oroch (**avrorin-etal1967 malchukov2000**)
a. *nia **aja-ni***
man good-POSS:3SG
'a GOOD man'

[9] Similar modifier-headed constructions are found in Even where modifier-headed possessor agreement is in fact attested, cf. *Asatkan **nood-do-n** haaram.* [girl beautiful-ACC-POSS:3SG I_-know] (**malchukov1995**). But unlike similar modifier-headed participles (in possessor agreement constructions) in Even (**malchukov1995**) and similar modifier-headed adjectives in Oroch (**malchukov2000** cf. also example **??**) Malchukov translates this example as a true possessive construction with a nominal attribute: 'I know the girl's beauty' (but not: 'I know the beautiful girl').

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

Whereas juxtaposition is the default type of adjective attribution marking in Oroch, modifier-headed possessor agreement occurs only in a special noun phrase type where the adjective is marked for contrastive focus. The special function marked by this construction is to focus on the property denoted by the adjective: 'a man, a property of whom is "to be good"' (**malchukov2000**). This noun phrase type thus resembles the function of relative clause formation.[10]

7.10.2.2 Nanay-Ulcha-Orok

According to the few grammatical sketches available, the Tungusic languages of the Nanay-Ulcha-Orok branch exhibit juxtaposition as the default device for adjective attribution marking, except Orok.

Head-driven agreement in Orok Attributive adjectives in Orok (aka Ulta) show agreement in number but not in case (or other categories) with the modified noun.

(28) Orok (**petrova1967**)

a. ***dāi*** *dalu(n)*
big store
'large store (i.e., warehouse, storehouse)'

b. ***dāi-l*** *dalu-l*
big-PL store-PL
'large stores'

c. ***dāi-l*** *dalu-l-tai*
big-PL store-PL-LOC
'in large stores'

[10] Note also that a similar construction is found in Even from the Northern Tungusic branch where it is only attested with participles: *Beji-l-bu* ***hör-če-wut-ten*** *emu-re-m.* [man-PL-ACC go-PFCT.PTCP-ACC-POSS:3PL bring-NONFUT-1SG] 'I brought back the men who had left' (**malchukov1995**).

Attributive nominalization in Ulcha According to **sunik1985** adjectives do not "normally" agree with the modified noun in Ulcha. The language is thus characterized by simple juxtaposition of attributive adjectives.[11]

Another adjective attribution marking device mentioned in Sunik's grammar is attributive nominalization by means of the suffix *-dŭma ~-dumE* (**sunik1985**).

(29) Ulcha (**sunik1985**)

a. *n'ūči-dumE* 'a/the little one (among other people)'

b. *ulEn-dumE* 'a/the good one (among other people)'

7.10.3 Manchu

The two Manchu languages Manchu proper and Sibe exhibit juxtaposition as the default adjective attribution marking device, similarly to the languages from the Nanay-Ulcha-Orok branch.

7.11 Yukaghir

Yukaghir (aka Yukagir) is a small family consisting of the two individual languages Tundra Yukaghir and Kolyma Yukaghir (aka Forest Yukaghir) (**salminen2007 maslova2003a maslova2003b**).

Noun phrases show strictly head-final constituent order in both Yukaghir languages. True adjective attribution scarcely exists because modifying "property words" in noun phrases are best coded as relative clauses.

The following relevant attribution marking types are attested in Yukaghir languages:

- incorporation
- anti-construct state marking

  of "verbal adjectives"

  of "nominal adjectives".

[11] **sunik1985** mentions, however, that a few adjectives sometimes show agreement with the modified noun in case and number (according to the simple or the possessive declension (sic!), i.e., are equipped with a possessive suffix) if they are "derived into nouns". Unfortunately, he does not provide examples.

Juxtaposition in Kolyma Yukaghir There is no large class of lexical adjectives in Yukaghir. The only true adjectives in both Yukaghir languages belong to two semantic pairs: 'small' versus 'big' and 'old, ancient' versus 'new, fresh; (an)other'. The use of adjectives from the first pair is even restricted to a few lexicalized expressions **maslova2003b** It is hard to categorize these adjectives according to their morpho-syntax. **maslova2003b** glosses the lexicalized expressions with the adjectives 'small' and 'big' as compounds, like in *čom+parnā* [big+crow] 'raven'. The adjective 'new', on the other hand can not only be used in such compounds but can even be marked additionally by the noun attribution suffix *-d* or by the action nominal suffix *-l* (**maslova2003b**).

Anti-construct state in Kolyma Yukaghir With the exception of the very small closed class described in the previous section, there are no adjectives in Kolyma Yukaghir (**krejnovic1982 maslova2003b**). All other words denoting qualities constitute a subclass of verbs. Used as attributes, these stative verbs take the 3rd person singular intransitive suffix *-j(e)*.[12] **maslova2003b** describes the inflected finite verbs, as in (**??**) as "special attributive forms". Syntactically, they have to be analyzed as juxtaposed relative clauses.

(30) Kolyma Yukaghir (**maslova2003b**)

a. Attribution

i. ***kellugī-je*** *šoromo*
lazy-ATTR:INTR.3SG person
'lazy man (lit. 'man who is lazy')' (146)

ii. ***kie-s'e*** *šoromo*
come-ATTR:INTR.3SG person
'man who comes' (67)

b. Predication

i. *id'ī pen* ***omo-s'***
here it good-PRED:INTR.3SG
'this is a nice place (lit. 'here, it is good')' (68)

Since verbs take different inflectional suffixes depending on their use as predicates or attributes (i.e., relative clauses, cf. **??**, **??**) the suffix *-j(e)* glossed as ATTR:INTR.3SG can only be analyzed as an anti-construct state marker, i.e., it

[12] Note that this morpheme takes different phonological shapes as the result of allomorphic alternations.

constitutes a dependent marking attribution device which is not connected to noun phrase internal agreement. Even though the marker belongs to the verbal inflection paradigm it is a true licenser of the attributive relationship between a modifying verb phrase (relative clause) and a noun.

Anti-construct state marking in Kolyma Yukaghir does not, however, belong to the domain of true adjective attribution marking but is a relative clause marking strategy.[13]

Anti-construct state in Tundra Yukaghir Tundra Yukaghir exhibits an anti-construct state marking device of verbs using a relative clause marking strategy similar to Kolyma Yukaghir (**maslova2003a**). In her short grammar, **maslova2003a** mentions the occurrence of a second anti-construct state marking device and gives the following example:

(31) Tundra Yukaghir (**maslova2003a**)
*lugu-je(-**d**)* *apanalā*
very_old-ATTR:INTR.3SG-ATTR woman
'very old woman'

The use of the marker *-d* is not obligatory and is even restricted to head nouns with vowel-initial stems (**maslova2003a**).

Interestingly, the second attribution marking device in Tundra Yukaghir is polyfunctional and regularly serves the licensing of single nouns (**??**) as well as complex noun phrases (**??**) as attributes.

(32) Tundra Yukaghir (**maslova2003a**)
a. *iŋli-**d*** *igije*
breast-ATTR ropes
'breast ropes' (49)
b. *tude kerewe-**d*** *ugurt'e*
3SG cow-ATTR legs
'the legs of his cow'[14] (44)

[13] In order to use a verb as modifier inside a noun phrase, the verb can also be nominalized, for example by means of an action nominal marker: *kel-u-l* [come-0-NMLZ '(a situation of) coming' (**maslova2003b**), *kel-u-l šoromo* [come-0-NMLZ person] '(a/the) man who came (i.e., (a/the) already arrived man)' (**maslova2003b**). This derivational nominalization of verbs to nominals is not considered to constitute an adjective attribution marking device either.

[14] The regular use of the cognate attribution marker *-d* (~*-n*) with nouns and noun phrases as attributes is described for Kolyma Yukaghir as well. The use of the marker as a licenser of adjective attribution, however, seems to be restricted to one adjective, 'new' (**maslova2003b**).

7.12 Yeniseian

Three branches are posited for the Yeniseian family, but only the Ket language from the northern branch still exists today (**werner1997a salminen2007**).

The following adjective attribution marking devices are attested in Ket:

- juxtaposition
- head-driven agreement
- attributive nominalization.

Juxtaposition and head-driven agreement in Ket Attributive adjectives in Ket are normally juxtaposed to the left of the noun they modify (**vajda2004**). Only a few simple adjective stems describing visible shapes or sizes may optionally take the plural suffix *-ŋ*, as shown in (**??**). The other morphological features assigned to the noun phrase, i.e., gender (or class) and case, are not sensitive to syntax in Ket.

(33) Ket (**vajda2004**)

a. ***qà*** *quˀŋ*
big tent:PL

b. ***qēŋ*** *quˀŋ*
big:PL tent:PL
'large tents'

vajda2004 notes that the optional number agreement marking is "a stylistic device used to emphasize the visual impression created by the quality being described". This emphasizing construction probably marks contrastive focus of the adjective: 'large tents' versus 'LARGE tents'.

Attributive nominalization in Ket **vajda2004** also mentions the nominalizing suffix -s which marks lexical and derived adjectives (**??**), noun phrases (**??**), and adposition phrases (**??**) as adnominal modifiers in headless noun phrases.[15]

[15] Note that the examples (**??** and **??**) seem to represent phonological compounds. This is evidenced by the phonological reduction in syllable-mediate vowels. The non-nominalized phrases, according to **vajda2005** are *úgda ślin* 'a long nose' and *qō-t-hwtɯ-ɣa* 'under the ice [ice-GEN-under]'. It is not clear from the description, however, if incorporation is relevant to morpho-syntax as well. But this phenomenon deserves further attention since adjective incorporation is scarcely attested in the world's languages but occurs in a few other non-related branches of the northern Eurasia.

(34) Ket (**vajda2005**)

a. Nominalized adjective

i. *sîn-**s***
old-NMLZ
'the old one'

ii. *súl-tu-**s***
blood-DERIV-NMLZ
'the bloody one'

b. Nominalized noun phrase

i. *úgd-ɔ́lin-**s***
long-nose-NMLZ
'the long-nosed one'

c. Nominalized adposition phrase

i. *qó-t-hɯtɯ-ɣa-**s***
ice-GEN-under-NMLZ
'the one under the ice'

Grammatical descriptions of Ket (**vajda2004** cf. also **krukova2007**) only give examples where these nominalized (headless) noun phrases are used in apposition, as in the contrastive focus construction (**??**).

(35) Ket (**vajda2005**)

a. Adjective predication
bū ***sîn-du*** / *bū* ***sîn-dʌ***
3SG old-M.COP 3SG old-F.COP
's/he is old'

b. Contrastive focus construction
bū ***sîn-s***
3SG old-NMLZ
's/he is OLD (i.e., 'an old one')'

The available data does not provide enough evidence for a detailed description and analyses of attributive nominalization by means of the suffix *-s* as a regular attribution marking device in Ket. It it possible that these nominalizations cannot be used as true modifiers of nouns but are restricted to headless noun phrases and are used only in special contrastive focus constructions.

There is even evidence against the analysis of nominalization as attributive marking in Ket. Vajda's examples of nominalized adverbials suggest that this contrastive focus marking is used predominantly in copular constructions (as predicates). Since the otherwise regular predicative agreement marking never occurs on these nominalizations (**vajda2004**) it could also be argued that the nominalizer *-s* constitutes a strategy for secondary predication marking rather than attribution marking.

Attributive nominalization in Ket definitely deserves more attention. The construction might constitute an example of the development of attributive nominalization independent of definiteness marking.

7.13 Turkic

Languages from the Turkic language family are spoken across all of northern Eurasia, including northeastern and southeastern Europe, and beyond. The family is divided into two major branches: Bulgar and Common Turkic. Whereas Bulgar Turkic is represented only by one language, the Common Turkic branch can be further divided into nine groups. Seven of these groups have members spoken in northern Eurasia: Oguz, Karluk, Kipchak, Altay Turkic, Yenisey Turkic (Khakas), Sayan Turkic, and Lena Turkic (**salminen2007**).

All Turkic languages are characterized by strict head-finality in their noun phrase structure. The prototypical adjective attribution marking device in Turkic languages is juxtaposition. This type occurs as the unmarked construction in all Turkic languages. In some Turkic languages, however, an attributive nominalizer marks an attributive adjective in contrastive focus constructions. This construction is systematically described (more or less) only for Chuvash from the Bulgar Turkic branch.

The following types of adjective attribution marking are attested:

- juxtaposition
- attributive nominalization.

7.13.1 Bulgar Turkic

The Bulgar (aka Oghur) subbranch of the Turkic language family is represented only by a single language, Chuvash.

Juxtaposition and attributive nominalization in Chuvash Similar to all other Turkic languages, Chuvash exhibits juxtaposition as the default and general adjective attribution marking device (**??**). Besides juxtaposition, an attributive nominalizer is used in contrastive focus constructions (**??**).

(36) Chuvash (**clark1998a**)

a. Juxtaposition
χura *χut*
black paper
'black paper'

b. Attributive nominalization
χur-i *χut*
black-ATTR paper
'BLACK paper (not of another color)'

The attributive article *-i* is similar to the possessive suffix 3rd singular. As in other Turkic languages, this article is also obligatorily used in headless noun phrases marked as direct (accusative) objects in Chuvash.

(37) Attributive nominalization in Chuvash (**benzing1993b**)
χur-i-ne */ **χĕrl-i-ne*** *ildem*
black-ATTR-ACC red-ATTR-ACC I_bought
(Which pen did you buy?) 'I bought a/the black / red one.'

Besides *-i*, a second nominalizer *-sker* is attested in Chuvash. Both formatives are used with similar classes of adjectival and other attributes.

(38) Attributive nominalization in Chuvash (**krueger1961**)

a. Article #1 *-i* (⇐ POSS:3SG)

i. Attributive adjective
lajăχχ-i
good-ATTR
'which is good / (a/the) good one'

ii. Attributive participle
vulan-i
read.PRF-ATTR
'which is read'

iii. Attributive noun
vărman-t-i
forest-LOC-ATTR
'which is in the forest'

b. Article #2 *-sker* (< Mari *ÿsker*)

i. Attributive adjective
lajăχ-sker
good-ATTR
'which is good / (a/the) good one'

ii. Attributive participle
vulană-sker
read.PRF-ATTR
'which is read'

iii. Attributive noun
vărman-ta-sker
forest-LOC-ATTR
'which is in the forest'

7.13.2 Common Turkic

7.13.2.1 Oguz

Juxtaposition in Azerbaijani Similar to all other Turkic languages, attributive adjectives are simply juxtaposed to the modified noun in Azerbaijani.

(39) Azerbaijani (**siraliev-etal1971**)

a. ***učа*** *daɣ* 'high mountain' [**high** mountain(NOM)]
b. ***učа*** *daɣ-ın* [**high** mountain-GEN]
c. ***učа*** *daɣ-da* [**high** mountain-LOC]
d. ***učа*** *daɣ-lar* [**high** mountain-PL]
e. ***učа*** *daɣ-lar-da* [**high** mountain-PL-LOC]
f. …

Attributive nominalization in Turkish Similar to other Turkic languages, the attributive nominalization device is used obligatorily in headless noun phrases marked as direct (accusative) objects in Turkish.

(40) Attributive nominalization in Turkish (**benzing1993b**)
kara-sını */ **kızıl-ını*** *aldım*
black-ATTR:ACC red-ATTR:ACC I_bought
(Which pen did you buy?) 'I bought a/the black / red one.'

7.13.2.2 Karluk

The default and general adjective attribution marking device in the languages of the Karluk subbranch of Common Turkic is juxtaposition and is similar to example (**??**) from Azerbaijani. Besides juxtaposition, attributive nominalization is also attested.

Attributive nominalization in Uigur The possessive suffix 3rd person singular occurs as an attributive nominalizer in contrastive focus constructions in Uigur. This construction is thus similar to example (**??**) from Chuvash from the Bulgar branch of Turkic.

(41) Uigur (**benzing1993b**)
uluy-ï *qatun*
biggest-ATTR woman
'the FIRST wife'

Attributive nominalization in Uzbek Similar to other Turkic languages, the article is also used obligatorily in headless noun phrases marked as direct (accusative) objects in Uzbek.

(42) Attributive nominalization in Uzbek (**boeschoten1998**)
(mėŋå qaysisi yarašadi,) ***qizilim-i,*** ***åqim-i****?*
red-ATTR:ACC white-ATTR:ACC
'(Which one suits me,) the red one, or the white one?'

7.13.2.3 Kipchak, Altay, Yenisey, Sayan, Lena

The default and general adjective attribution marking device in the languages of the Kipchak, Altay, Yenisey (aka Khakas), Sayan and Lena subbranchs of Common Turkic is **juxtaposition** and is similar to example (**??**) from Azerbaijani.

7.14 Nakh-Daghestanian

Nakh-Daghestanian is a language family of the Caucasus. It is named after its two main branches: Nakh and Daghestanian. Whereas Nakh comprises only a few single languages, the Daghestanian branch can be further divided into several subbranches **salminen2007**

The predominant order of noun phrase constituent in Nakh-Daghestanian languages is adjective-noun. Regarding the morpho-syntactic licensing of adjective attribution, the Nakh-Daghestanian family is characterized by a relatively high diversity of noun phrase types.

The following adjective attribution marking devices are attested:

- juxtaposition
- head-driven agreement marking
- anti-construct state agreement marking
- anti-construct state marking
- attributive nominalization.

7.14.1 Daghestanian

7.14.1.1 Avar-Andi-Tsezic

The Avar-Andi-Tsezic group of Daghestanian is named after three groups of closely related languages: Andi (comprising the languages Akhvakh, Andi, Bagvalal, Botlikh, Chamalal, Godoberi, Karata and Tindi), Tsezic (comprising the languages Tsez (aka Dido), Hinuq, Khwarshi, Inkhokvari, Bezhta (aka Kapucha) and Hunzib. The single language Avar forms the third group of Avar-Andi-Tsezic (**salminen2007**).

The prototype of adjective attribution marking in the Avar-Andi-Tsezic languages seems to be head-driven agreement, which occurs in all languages of this group.

Head-driven agreement in Godoberi The unmarked constituent order in Godoberi is adjective-noun.[16] Adjectives agree with the head noun in the features GENDER (if a position for the class-marker is available) and NUMBER.

[16] The reversed order marks contrastive focus on the adjective: *hac'a χ°aji* [white dog] 'white dog', *χ°aji hac'a* [dog white] 'that very dog (of several others) which is white' (**kazenin1996a**).

(43) Godoberi (**tatevosov1996a**)

a. Adjectives taking a gender class prefix

i. ***w-oχar*** *ima* ‘old father’ [M]

ii. ***j-aχar*** *ila* ‘old mother’ [F]

iii. ***b-aχar*** *hamaχi* ‘old donkey’ [N]

iv. ***r-aχar*** *hamaχi-be* ‘old donkeys’ [N.PL]

b. Adjectives taking a gender class suffix

i. ***q’arúma-w*** *ima* ‘greedy father’ [M]

ii. ***q’aruma-j*** *ila* ‘greedy mother’ [F]

iii. ***q’arúma-b*** *hamaχi* ‘greedy donkey’ [N]

iv. ***q’arúma-r*** *hamaχi-be* ‘greedy donkeys’ [N.PL]

Attributive nominalization in Tsez In Tsez, two lexical classes of adjectives have to be distinguished. The members of the first class take gender agreement prefixes. The (few) members of the second class are simply juxtaposed to the modified noun (**alekseev-etal2004**).

There is an additional attributive marker: the attributive nominalizing suffix *-ni* which marks attributive adjectives in headless noun phrases and also “restrictive” forms of the adjective.

(44) Tsez (**alekseev-etal2004**)

a. Nominalized headless adjective

i. *igu-**n**-a:*
good-ATTR-ERG
‘a good one’

ii. *igu-**ni**-r*
good-ATTR-DAT
‘to a good one’

b. “Restrictive” attributive adjective

i. *(eyda) **eġe-ni*** *uži* *dey* *esiy* *yoł*
this little-ATTR boy 1:GEN brother:NOM be:PRS
‘(this) little boy (and not one of the others) is my brother’

The content of the “restrictive” (aka “definite”) form remains somewhat uncertain. The translation of (**??**) in the description of **alekseev-etal2004** clearly resembles contrastive focus marking (‘the LITTLE boy’).

7.14.1.2 Lak

The Lak subbranch of Daghestanian is formed by one single language: Lak proper.

Head-driven agreement in Lak Constituent order in Lak is adjective-noun. The language exhibits two adjective attribution marking devices. The unmarked and default attribution marking device is head-driven agreement which characterizes adjectives derived by means of the adjectivizer *-ssa*, as in (??). These derived adjectives only agree in gender class. Other morpho-syntactic marking is not applied.

(45) Lak (**zirkov1955**)

a. ***uč-ssa*** *adimina*
fat.I-ADJZ person(I)
'fat man'

b. ***b-uč-ssa*** *nic*
III-fat-ADJZ bull(III)
'fat bull'

c. ***b-uč-ssa*** *nic-ru*
III-fat-ADJZ bull(III)-PL
'fat bulls'

Note that the suffix *-ssa* is a derivational formative rather than a marker of attribution since it occurs on adjectives in attributive and predicative position alike. Predicative adjectives even show similar gender agreement inflection (**zirkov1955**).

Anti-construct state agreement in Lak While head-driven agreement marking, as in (??), constitutes the basic and unmarked adjective attribution marking device in Lak, anti-construct state agreement marking is restricted to contrastive focus constructions.

(46) Lak (**zirkov1955**)

a. ***uč-ma*** *adimina*
fat.I-ATTR:I person(I)
'FAT man'

b. ***b-uč-mur*** *nic*
III-fat-ATTR:III bull(III)
'FAT bull'

c. ***buč-mi*** *nic-ru*
III-fat-ATTR:PL bull(III)-PL
'FAT bulls'

Note that the occurrence of the anti-construct state agreement marking suffixes *-ma, -mur, -mi* is restricted to attributive adjectives. Unlike adjectives with the derivational formative *-ssa* with head-driven agreement marking in number only, adjectives in contrastive focus (occurring in the anti-construct state agreement noun phrase type) show agreement in number as well (**zirkov1955**).

7.14.1.3 Dargwa

The Dargwa subbranch of Daghestanian has traditionally been described as consisting of one single language (i.e., Dargwa proper) with several sub-varieties (**salminen2007**). According to **korjakov2006a** Dargwa varieties exhibit fairly diverse grammatical structures and can therefore be described as separate languages.

Anti-construct state agreement and juxtaposition in Dargwa In Dargwa, two adjective attribution marking devices occur. Whereas anti-construct state (number) agreement marking (**??**) is the default type, juxtaposition (**??**) is restricted to "poetic language" (**isaev2004**).

(47) Dargwa (**isaev2004**)
a. Anti-construct state agreement
i. ***aɢ-si*** *ɢali*
high-ATTR:SG house(SG)
'lofty house'
ii. ***aɢ-ti*** *ɢulri*
high-ATTR:PL house:PL
'lofty houses'
b. Juxtaposition
i. ***aɢ*** *dubura*
high mountain
'high mountain'

7.14.1.4 Lezgic

The Lezgic subbranch of Daghestanian comprises the languages Agul, Archi, Badukh, Kryz (aka Kryts), Lezgian, Rutul, Tabasaran, Tsakhur and Udi.

Adjective-noun is the basic constituent order in the noun phrase of all Lezgic languages. Regarding their adjective attribution marking, the Lezgic languages exhibit the highest degree of diversity. All types found in Nakh-Daghestanian are attested: juxtaposition, head-driven agreement marking, anti-construct state agreement marking, anti-construct state marking and attributive nominalization.

Juxtaposition in Udi The default adjective attribution marking device in Udi is juxtaposition, like in the following (incomplete) paradigm.

(48) Udi (**schulze-furhoff1994**)

a. ***kala*** *ĝara-Ø* 'the old son' [ABS]

b. ***kala*** *ĝara-en* [ERG]

c. ***kala*** *ĝara-i* [GEN]

d. …

Juxtaposition and head-driven agreement in Tabasaran The default adjective attribution marking device in Tabasaran is juxtaposition, as in Udi. Only a minor lexical subclass of two adjectives in this language deviate in this respect and show gender and number agreement.

(49) Tabasaran (**kurbanov1986**)

a. ***učču-r*** *adaš*
beautiful-I father(I)
'beautiful father'

b. ***učču-b*** *gjajvan*
beautiful-II horse(II)
'beautiful horse'

c. ***učču-dar*** *gjunšjir*
beautiful-PL horse:PL
'beautiful horses'

Head-driven agreement in Archi Attributive adjectives in Archi show agreement in gender and number with the modified noun; see the complete agreement paradigm for the adjective 'good'.

(50) Archi (**kibrik1994a**)

a.	*hibàɬu* 'good'	[I SG]
b.	*hibàɬu-r*	[II SG]
c.	*hibàɬu-b*	[III SG]
d.	*hibàɬu-t*	[IV SG]
e.	*hibàɬ-ib*	[PL]

Anti-construct state agreement in Tsakhur Adjectives in Tsakhur can be divided into three subclasses according to their choice of attribution marking devices. The first, minor lexical class of adjectives in Tsakhur is characterized by missing inflection. Adjectives belonging to this class are simply juxtaposed to the modified noun (**talibov2004**). Members of the two other adjective classes exhibit anti-construct state agreement marking.

(51) Tsakhur (**talibov2004**)

a. i. Gender class I–III
bat'raj-na *jis* */ diǩ* */ balk^{h}an*
beautiful-ATTR:I–III girl(I) son(II) horse(III)
'beautiful girl / son / horse'

ii. Gender class IV
bat'raj-n *č'alag*
beautiful-ATTR:IV forest(IV)
'beautiful forest'

b. i. Gender class I
x̲arna *jis*
big:ATTR:I–III mother(I)
'old mother (viz. grandmother)'

ii. Gender class IV
ǩadɨn *balag*
big:ATTR:IV sack(IV)
'big sack'

Whereas the anti-construct agreement marker of adjectives from the first group (**??**) is formally identical with the genitive case suffixes of nouns, adjectives from the second group (**??**) are equipped with a morphologically complex formative including the genitive suffix and a phonological stem alternation (**talibov2004**).

Nominalization in headless noun phrases in Udi The default adjective attribution marking device in Udi is head-driven agreement. In headless noun phrases, however, attributive adjectives are obligatorily nominalized by means of the stem augment *-o-* ABS / *-t'-* OBL.

(52) Udi (**schulze-furhoff1994**)

a.	*kala-o* 'the big/old one'	[NMLZ.ABS]
b.	*kala-o-r*	[NMLZ.ABS-PL]
c.	*kala-t'-in*	[NMLZ:OBL-ERG]
d.	*kala-t'-ĝ-on*	[NMLZ:OBL-PL-ERG]
e.	*kala-t'-ay*	[NMLZ:OBL-GEN]
f.	…	

Nominalization in headless noun phrases in Lezgian Attributive adjectives in headless noun phrases are nominalized in Lezgian as well. The nominalizing suffix exhibits different forms in the absolute singular case (*-di*), in the oblique cases (*-da*) and in plural (*-bur*).

(53) Headless adjectives in Lezgian (**haspelmath1993**)

a.	*q̃acu-di* 'green one'	[ATTR:SG]
b.	*q̃acu-da*	[ATTR:ERG.SG]
c.	*q̃acu-da-n*	[ATTR-GEN]
d.	*q̃acu-bur*	[ATTR:PL]
e.	*q̃acu-bur-u*	[ATTR:PL-ERG]
f.	…	

The same attribution marker is also used for the nominalization of noun phrases.

(54) Nominalized noun phrases in Lezgian (**haspelmath1993**)

a. Pronoun

i.	*zi* 'my'	[POSS:1SG]
ii.	*zi-**di*** 'mine'	[POSS:1SG-ATTR]

b. Lexical noun

i. *dide.di-n* 'mother's' [mother-GEN]

ii. *dide.di-n-**di*** 'mother's' [mother-GEN-ATTR]

Even though adjectives without a lexical head in Udi and Lezgian are nominalized there is no evidence that these nominalizations serve as attribution marking devices.

Anti-construct state in Rutul In Rutul, attributive and predicative adjectives are differentiated by means of two different derivations. Whereas attributive adjectives take an anti-construct suffix *-d ~-di̯*,[17] predicative adjectives take a suffix *-i̯ ~-i̯*[18] or are not marked at all (**alekseev1994a**).

Attributive adjectives do not inflect other than by means of anti-construct state marking.

(55) Rutul (**alekseev1994a**)

a. ***äkkà-d*** *dahàr*
big-ATTR stone
'big stone'

b. ***äkkà-d*** *dahàr-bi̯r*
big-ATTR stone-PL
'big stones'

Note that the anti-construct state marker *-d ~-di̯* is identical to the genitive case of nouns and thus constitutes a polyfunctional marker (**alekseev1994a**).

7.14.2 Nakh

The Nakh branch of Nakh-Daghestanian comprises only three languages: Bats, Ingush and Chechen. The latter two form a common subbranch (**salminen2007**).

The noun phrase structure in all three languages is basically similar. Attributive adjectives precede the modified noun and show head-driven agreement. Adjectives in headless noun phrases are additionally marked with an attributive nominalizer.

[17] The allomorph *-di̯* occurs after consonants (**alekseev1994a**).

[18] The allomorph *-i̯* occurs after dorsal consonants (**alekseev1994a**).

7.14.3 Chechen-Ingush

Head-driven agreement in Ingush Attributive adjectives in Ingush agree in case with the modified noun. The adjective agreement paradigm, however, exhibits only a single case distinction of nominative versus oblique.

(56) Case agreement paradigm in Ingush (**nichols1994b**)

a. ***joqqa*** *jurt* 'large village' [NOM]
b. ***joqqa-ča*** *jurt-a* [GEN]
c. ***joqqa-ča*** *jurt-aa* [DAT]
d. ***joqqa-ča*** *jurt-uo* [ERG]
e. ***joqqa-ča*** *jurt-aca* [INS]
f. …

Some adjectives also show agreement in gender; but only very few adjectives additionally agree in number with the modified noun (**nichols1994b**).

Nominalization in headless noun phrases in Chechen Beside head-driven agreement, Chechen (similar to the other Nakh languages) exhibits attributive nominalization as the regular adjective attribution marking device in headless noun phrases. The formative is a thematic stem extension merged with the case inflection.

(57) Chechen (**nichols1994a**)

a. *leqa kert*
high fence
'high fence'
b. *leqa-**nig***
high-ATTR:NOM.SG
'the high one'

Even though adjectives without a lexical head in Chechen are nominalized, there is no evidence that these nominalizations serve as attribution marking devices.

7.14.4 Bats

The noun phrase structure in Bats (aka Tsova-Tush or Batsbi) is similar to the structure found in closely related Chechen and Ingush. Attributive adjectives show head-driven agreement. Adjectives in headless noun phrases are additionally marked by means of nominalization (**holisky-etal1994**).

7.15 Abkhaz-Adyghe

The Abkhaz-Adyghe (aka Northwest Caucasian) family consists of the two branches Abkhaz and Circassian, each of which comprises two languages. A third branch, Ubykh, is now extinct (**salminen2007**). All languages are spoken in the northwestern Caucasus region.

Whereas the adjective-noun constituent order is similar in all Abkhaz-Adyghe languages, the adjective attribution marking devices

- head-driven agreement (Abkhaz)
- incorporation (Circassian)

occurring in the two branches of this family diverge considerably.

7.15.1 Abkhaz

The Abkhaz branch of Abkhaz-Adyghe comprises the two very closely related varieties Abkhaz proper and Abaza. The constituent order inside the noun phrase of both languages is normally noun-adjective. Only adjectives denoting nationality deviate from this rule and precede the modified noun (**comrie1981**).

Head-driven agreement in Abkhaz Attributive adjectives in Abkhaz show number agreement.[19] Note, however, that a plural noun modified by an adjective may remain unmarked (**hewitt1989a**). Even though the plural marker may attach only once at the right phrase edge, it is best analyzed as an agreement marker and not a clitic. This is evidenced by the fact that the adjective may take the non-human pluralizer even if it modifies a human noun.[20]

(58) Abkhaz (**hewitt1989a**)

a. *a-là(-k˚à)* ***bzə̀ya-k˚a***
DEF-dog-PL:NONHUM good-PL:NONHUM
'the good dogs'

[19] Noun phrases with an attributive adjective following a non-inflected noun in Abkhaz have alternatively been analyzed as polysynthetic constructions (hence adjective incorporation), e.g., by **rijkhoff2002** and **gil2005**

[20] Note that in the closely related language Abaza, plural marking occurs twice but the non-human pluralizer constitutes the obligatory plural agreement marker on adjectives modifying nouns of any gender class (**lomtatidze-etal1989**).

b. *à-ʒġab(-ċ°a)* ***bzə̀ya-k°a*** */ **bzə̀ya-ċ°a***
DEF-girl-PL:HUM good-PL:NONHUM good-PL:HUM
'the good girls'

7.15.2 Circassian

The Circassian (aka Adyghe) branch of Abkhaz-Adyghe comprises the two languages Adyghe and Karbardian. Both languages exhibit similar noun phrase structures. The constituent order inside the noun phrase is normally noun-adjective. Noun phrases with modifying adjectives in Adyghe and Karbardian are often described as single compound words (**comrie1981**).

Adjective incorporation in Karbardian Attributive adjectives in Karbardian (aka Eastern Circassian) occur in a polysynthetic structure to the right of the modified noun. Number and case inflection of the noun phrase is suffixed to the adjective.

(59) Karbardian (**colarusso1989**)

a. *pṡaaṡa-**daax̂a**-r*
girl-beautiful-ABS
'the beautiful girl'

b. *pṡaaṡa-**daax̂a**-ha-r*
girl-beautiful-PL-ABS
'the beautiful girls'

c. *pṡaaṡa-**daax̂a-c'ək'°**-ər*
girl-beautiful-little-ABS
'the small beautiful girl'

7.16 Kartvelian

Kartvelian is a language family comprising the four languages Georgian, Svan, Laz and Mingrelian (aka Megrelian or Iverian). The latter two languages constitute the Zan subbranch inside the family (**salminen2007**). Kartvelian languages are all spoken in the southern Caucasus, mainly in Georgia but also in adjacent countries.

In the modern Kartvelian languages, the unmarked constituent order of adjectival modifiers and head is noun-final, although the opposite order is also possible (**harris1991a**).

Three adjective attribution marking types are attested:

- juxtaposition
- head-driven agreement
- appositional head-driven agreement.

The inherited Common Kartvelian agreement marking, however, is more or less preserved only in the marked (but inherited) head-initial noun phrase type. In the head-final noun phrase type, on the other hand, the modern Kartvelian languages display a strong tendency to lose head-driven agreement. Preposed attributive adjectives in Mingrelian and Laz are juxtaposed to the head noun as a rule. In Modern Georgian and Svan, the agreement paradigm of preposed attributive adjectives shows a high degree of syncretism.

7.16.1 Georgian

Head-driven agreement in Georgian The only agreement feature in Modern Georgian is CASE. Note, however, that the adjective agreement paradigm exhibits only three differentiated forms.[21]

(60) Georgian (**aronson1991**)

a. ***ʒvel-i*** *c'ign-i* 'old book' [NOM]
b. ***ʒvel-ma*** *c'ign-ma* [ERG]
c. ***ʒvel-Ø*** *c'ign-s* [DAT]
d. ***ʒvel-i*** *c'ign-is* [GEN]
e. ***ʒvel-i*** *c'ign-it* [INS]
f. ***ʒvel-Ø*** *c'ign-ad* [ADV]
g. …

[21] In the marked head-initial constituent order of noun and adjective, which is used in archaic style or for emphasis, case agreement is complete (**tuite1998**).

Juxtaposition in Georgian Whereas the so-called consonantal-stem adjectives like 'old' in (**??**) show head-driven agreement, there is another lexical class of adjectives (characterized by a stem-final vowel, hence "vocalic-stem adjectives"), the members of which are simply juxtaposed to the modified noun.

(61) Georgian (**aronson1991**)

a. ***parto*** *gza* 'wide road' [NOM]
b. ***parto*** *gza-m* [ERG]
c. ***parto*** *gza-s* [DAT]
d. ***parto*** *gz-is* [GEN]
e. ***parto*** *gz-it* [INS]
f. ***parto*** *gz-ad* [ADV]
g. …

Appositional head-driven agreement in Georgian Appositional modification seems to occur as a secondary type of adjective attribution marking in Georgian. Attributive adjectives are normally preposed and show only limited agreement (**??**). In postposition (marking emphasis), however, the adjective inflects for the full set of cases and numbers (**??**). This construction thus resembles an independent (headless) noun phrase in apposition to the semantic head (**testelec1998**). The construction probably marks contrastive focus of the adjective.

(62) Georgian (**testelec1998**)

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

7.16.2 Svan

Head-driven agreement in Svan Attributive adjectives in Svan show limited agreement in case. The paradigm of the agreement marker exhibits only two members: one for nominative and one for the oblique cases.

(63) Svan (**tuite1997**)

a. ***luwzer-e*** *ma:r-e*
diligent-NOM man-NOM
'a diligent man'

b. ***luwzer-a*** *ma:re:m-i(s&) (nas&dabw)*
diligent-OBL man-GEN work
'(the work) of a diligent man'

!!strange symbols in the Svan example, check the source!!

schmidt1991 however, describes the tendency in Svan to abolish agreement completely and use an uninflected variant of the attributive adjective in the oblique cases instead.

7.16.3 Zan

Zan is a subbranch of Kartvelian formed by the two languages Mingrelian and Laz. The default type of adjective attribution marking in both languages is juxtaposition which occurs obligatorily in the unmarked head-final noun phrase. In the marked head-initial noun phrase, however, attributive adjectives normally agree in number and case with the head noun.

Juxtaposition and head-driven agreement in Mingrelian The two adjective attribution marking devices occurring in Zan languages are illustrated with Mingrelian examples.

(64) Mingrelian (**harris1991b**)

a. Juxtaposition
skvam *cira-en-k*
beautiful girl-PL-NAR
'beautiful girl'

b. Head-driven agreement
cira-en-k ***skvam-en-k***
girl-PL-NAR beautiful-PL-NAR
'BEAUTIFUL girl'[22]

[22] Note that the case marking formative does not obligatorily occur on both constituents in the marked head-initial noun phrase in Mingrelian (**harris1991b**).

7.17 Semitic

Semitic languages are only marginally represented in northern Eurasia. The few languages considered here belong either to the Arabic subbranch of Central Semitic or to Northwest Semitic.

Only one single type of adjective attribution marking is attested in these two branches:

- head-driven agreement.

7.17.1 Northwest Semitic

Neo-Aramaic (aka Modern Aramaic) is the only language of the northwestern branch of Semitic considered in the present survey. It is spoken in the Middle-East in north-western Iran, Iraq and south-eastern Turkey, but also in adjacent areas of the Caucasus in Azerbaijan, and therefore falls into the geographic area of investigation.

Head-driven agreement in Neo-Aramaic Constituent order inside the noun phrase of Neo-Aramaic is noun-adjective. Attributive adjectives agree with the modified noun in gender and number.

(65) Neo-Aramaic (Kurdistan) (**krotkoff1982**)
   a. *ya:la* ***zu:ra*** ‘small boy’
   b. *bra:ta* ***zurta*** ‘small girl’
   c. *bnu:ne* ***zu:re*** ‘small kids’

7.17.2 Central Semitic

7.17.2.1 Arabic

Cypriot Arabic (aka Kormakiti) and Maltese are two Arabic languages of the Central Semitic branch spoken on the Mediterranean islands Cyprus and Malta, and thus belong to Europe geographically.

Head-driven agreement in Maltese The basic and unmarked constituent order in Maltese is noun-adjective. A few adjectives, however, can precede the noun in an emphatic construction (**borg-etal1996**).

Adjectives show distinct forms for gender and number in accordance with the morphological features of the modified noun.

(66) Maltese (**aquilina1959**)

a. *ra:jel* ***sabi:ħ***
man beautiful:M:SG
'beautiful man'

b. *mara* ***sabi:ħa***
woman beautiful:F:SG
'beautiful woman'

c. *nies* ***sbieħ***
people beautiful:PL
'beautiful people'

Optionally, the attributive adjective can additionally be marked for definiteness.

(67) Maltese (**aquilina1959**)
il-ktieb ***(il-)qadi:m***
DEF-book (DEF-)old
'the old book'

Even though the construction with a repeated definite marker resembles attributive nominalization, it is best analyzed as agreement in the DEFINITE value of the feature SPECIES (**himmelmann1997**). Himmelmann compares the construction in Maltese to Standard Arabic, where similar definite (and indefinite) agreement occurs.

7.18 Uralic

The Uralic language family comprises the branches (roughly from West to East) Hungarian, Saamic, Finnic, Permic, Mari, Mordvin, Khanty, Mansi, and Samoyedic (**salminen2007**). Except for most languages from the Samoyedic subbranch of the family, Uralic languages are all spoken in Europe. Uralic is thus one of the major families on the European linguistic map.

The constituent order inside the noun phrase is strictly adjective-initial in all Uralic languages. Similar to Mongolic, Turkic and many other languages of North Asia, the prototypical adjective attribution marking device in Uralic languages is juxtaposition. This type occurs as the unmarked construction in all Uralic languages with the exception of the two western branches Saamic and Finnic, which have abandoned juxtaposition and developed new types.

Secondary adjective attribution marking devices are also attested in languages of the Permic and Mari (and probably also other) branches of Uralic, even though

juxtaposition is used in these languages as the default strategy for adjective attribution marking.

The following five adjective attribution marking devices occur in Uralic:

- juxtaposition
- head-driven agreement
- anti-construct state marking
- appositional head-driven agreement
- attributive nominalization.

7.18.1 Samoyedic

7.18.1.1 Enets

The two languages Forest Enets and Tundra Enets constitute the Enets branch of Uralic.

Juxtaposition in Forest Enets In both Enets languages, attributive adjectives are juxtaposed to the modified noun by default.

(68) Forest Enets (**siegl2013a**)

a. ***aga*** *to*
large lake(NOM:SG)
'a/the large lake'

b. ***aga*** *to-ʔ*
large lake-NOM:PL
'large lakes'

c. ***aga*** *to-xiʔ*
large lake-NOM:DU
'two large lakes'

7.18.1.2 Nenets, Selkup, Nganasan

The two languages Forest Nenets and Tundra Nenets constitute the Nenets subbranch of Samoyedic. The Selkup branch consists of the three very closely related languages Northern Selkup, Central Selkup and Southern Selkup. The Nganasan branch consists only of one language: Nganasan proper.

Attributive adjectives in Nganasan and the Nenets and Selkup languages are juxtaposed to the modified noun by default, similar to examples (**??**) from Forest Enets and (**??**) from Hungarian.

7.18.2 Hungarian

The Hungarian branch of Uralic consists only of one language, i.e., Hungarian proper.[23]

Juxtaposition in Hungarian In Hungarian, attributive adjectives are juxtaposed to the modified noun by default.

(69) Hungarian (**hall1938**)

a. *a* ***fekete*** *szem*
DEF black eye
'the black eye'

b. *a* ***fekete*** *szem-ek*
DEF black eye-PL
'the black eyes'

c. *a* ***fekete*** *szem-ek-nek*
DEF black eye-PL-DAT
'to the black eyes'

d. *a* ***fekete*** *szem-eid*
DEF black eye-PL:POSS:2SG
'your black eyes'

7.18.3 Khanty, Mansi, Mari, Mordvin

The two languages Northern Khanty and Eastern Khanty constitute the Khanty branch of Uralic. A third language, Southern Khanty, is extinct (**salminen2007**).

[23] The outlying dialect Csángó Hungarian spoken in Romania is not considered as a distinct language here.

The Mansi branch of Uralic consists of the two very closely related languages Northern Mansi and Eastern Mansi. Two other Mansi languages, Western Mansi and Southern Mansi, are extinct (**salminen2007**). The Mari branch of Uralic is formed by Western Mari (aka Hill Mari) and Eastern Mari (aka Meadow Mari) (**salminen2007**). The Mordvin branch of Uralic is formed by the two closely related languages Erzya and Moksha (**salminen2007**).

Attributive adjectives in all Khanty, Mansi, Mari and Mordvin languages are juxtaposed to the modified noun by default, similar to examples (**??**) from Forest Enets and (**??**) from Hungarian.

LATEX Verweise zu Bsp.

7.18.4 Permic

All three Permic languages, Komi-Permyak, Komi-Zyrian and Udmurt exhibit two distinct types of adjective attribution marking. The default type is juxtaposition, which is the inherited Proto-Uralic type (**decsy1990**). However, an attributive nominalization device is used in contrastive focus constructions as a second type.

Juxtaposition in Komi-Zyrian The unmarked sequence of adjective and noun, i.e., juxtaposition, is illustrated by an example from Komi-Zyrian.

(70) Komi-Zyrian (**lytkin1966a**)

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

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

Attributive nominalization + appositional head-driven agreement in Udmurt In Udmurt, an attributive nominalizer homophonous with the 3rd person possessive inflection marker is regularly used as an adjective attribution marking device in contrastive focus constructions. Historically, both formatives are similar (cf. §**??** in Part IV Diachrony).

(71) Udmurt (**winkler2001**)

a. Juxtaposition (default)

i. ***badʒ̑ym*** *gurt*
big house
'large house'

ii. ***badʒ̑ym*** *gurt-jos-y*
big house-PL-ILL
'to large house/s'

b. Attributive nominalization (contrastive focus)

i. ***badʒ̑ym-ėz*** *gurt*
big-ATTR house
'LARGE house'

ii. ***badʒ̑ym-jos-a-z*** *gurt-jos-y*
big-PL-ILL-ATTR house-PL-ILL
'to LARGE house/s'

An adjective equipped with the nominalizer is also marked with (agreeing) case and number suffixes indicating that the nominalized adjective occurs in an attributive appositional construction. Note that the nominalizer also serves as the licenser of adjectival (and other) modification in headless noun phrases.

(72) Nominalization in Udmurt (**winkler2001**)

a. Adjective
*badʒ̑ym-**ėz***
big-ATTR
'the big one'

b. Demonstrative
*taiz-**ėz***
DEM:DIST-ATTR
'that one over there'

c. Possessor noun phrase
*Ivan-len-**ėz***
Ivan-GEN-ATTR
'that one of Ivan's'

(73) Contrastive focused attribute

a. Demonstrative
*taiz-**ėz*** *gurt*
DEM:DIST-ATTR house
'THAT (particular) house over there'

b. Possessor noun phrase
*Ivan-len-**ėz*** *gurt*
Ivan-GEN-ATTR house
'IVAN's house (and not someone else's)'

Examples (??–??) show that attributive nominalization in Udmurt is a true attribution marking device which is polyfunctional and not restricted to headless noun phrases.

Note that the attributive article is normally labeled "determinative suffix" (or in similar terms) in the Udmurt (and Uralic) grammatical tradition. This label probably originates from the formative's function as a quasi-definite marker. But "determinative" inflection is obligatory only in the case of differential object marking with the marked versus the unmarked accusative. Note also that the definite-marked accusative suffix, again, is historically identical with the 3rd person possessive suffix.

(74) Differential object marking in Udmurt (**winkler2001**)

a. *mon kniga* *lid̑ʒ́-i*
1SG book(ACC) read-1SG.PST
'I have read a book.'

b. *mon (ta) kniga-**jez** lid̑ʒ́-i*
1SG this book-ACC read-1SG.PST
'I have read the (i.e., 'this certain') book.'

Note also that in these and similar examples, the concept of definiteness does not always coincide with the use of the differential "in/definite accusative" marking. According to **winkler2001** "the marked accusative is used if the object itself is focused, whereas the unmarked is employed if the action itself bears the logical accent." Accordingly, even such occurrences of the "determinative suffix" thus resemble focus marking rather than definiteness marking.

Even though contrastive focus inflection of nouns (or noun phrases) would be the result of purely morphological (morpho-semantic) assignment, contrastive focus inflection of adnominal adjectives can only be analyzed as a morpho-syntactic feature assigned noun phrase internally. This is evidenced by the agreement pattern: whereas adjectives in non-contrasted (unmarked) constructions are simply juxtaposed to the head noun, contrastive focused adjectives normally show head-driven number agreement.[24] Agreement marking on the adjective is clearly as-

[24] The different order of morphemes in certain members of contrastive focus inflection paradigms

signed by syntax, the head noun being the agreement trigger and the attributive adjective (in contrastive focus) being the agreement target.

Attributive marking in contrastive focus constructions in Udmurt (and the other Permic languages) is similar in theory to prototypical anti-construct state agreement marking in languages like Russian, with regard to both synchrony and diachrony. The construction is still analyzed as attributive nominalization because the agreement marking on the nominalized attribute is the indirect result of the attributive appositional construction and the nominalizing and agreement formatives are not fused synchronically.

Appositional head-driven agreement in Udmurt Note, however, that in Udmurt, number agreement also sometimes occurs without the contrastive focus marker.

(75) Head-driven plural agreement in Udmurt

a. *badʒ̑ym-eś̑ gurt-jos*
big-PL house-PL
'LARGE houses' (**winkler2001**)

b. *paśkɨt-eś uram-jos*
wide-PL street-PL
'WIDE streets' (**csucs1990**)

According to **csucs1990** head-driven agreement marking in constructions without the "determinative suffix" is the result of analogy. The fact that their use is still restricted to contrastive focus constructions, and is therefore an appositional attribution marking device, is crucial for the analysis as appositional head-driven agreement (as opposed to true head-driven agreement).

7.18.5 Finnic

The Finnic (aka Fennic or Baltic Finnic) branch of Uralic comprises the following languages: Livonian, Estonian, Votic, Finnish, Ingrian, Karelian, Lude and Veps.[25]

(i.e., number-, case-, and (former) possessive suffix) as compared to the historically similar "regular" possessive inflection (**winkler2001**) is not of concern here. This phenomenon does, however, provide evidence for the analysis of the contrastive focus marker of adjectives and the possessive marker of nouns as two different formatives from a synchronic point of view.

[25] The Võro variety of Estonian, the Meänkieli and Kveeni varieties of Finnish, and the Olonets variety of Karelian are not considered distinct languages here.

The Finnic branch is exceptional among Uralic in that all of its member languages regularly exhibit head-driven agreement as the regular type of adjective attribution marking.

Head-driven agreement in Finnish The morphological features assigned to the head noun in Finnish are passed on to its adjectival (and other) modifiers. Finnish adjectives thus show a prototypical instance of head-driven agreement.

(76) Finnish (personal knowledge)

a. ***iso*** *talo*
big house
'large house'

b. ***iso-t*** *talo-t*
big-PL house-PL
'large houses'

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

d. ***iso*** *(*iso-ni)* *talo-ni*
big big-POSS:1SG house-POSS:1SG
'my large house'

Note, however, that not all morphological features assign their values to the attributive adjective in Finnish. Whereas number (**??**) and case marking (**??**) are assigned to the adjective, possessive marking (**??**) is not (as noted earlier in §**??**).

7.18.6 Saamic

Saamic languages are spoken on the Scandinavian peninsula in north-central Norway and Sweden as well as in northern Finland and on the Kola peninsula in northwesternmost Russia. Saamic branches further into an eastern and a western subgroup.

The Saamic languages are exceptional among Uralic and the languages of most other families of Europe in that they exhibit special attributive marking of adjectives, prototypically expressed by an invariable attributive suffix. In §?? of Part II (Typology), this noun phrase type was characterized as *dependent marked attributive state*; the corresponding formative is labeled *anti-construct state marker*.

Note, however that the regular use of this inflectional category of adjectives and the relevant formatives vary considerably across the different Saamic languages.

7.18.6.1 East Saamic

The four living East Saamic languages Ter, Kildin, Skolt and Inari Saami are spoken on the Kola peninsula in northwesternmost Russia and in the adjacent parts of northern Finland.

Anti-construct state in Skolt Saami Prototypically, the anti-construct state marking suffix in Saamic languages has the shape *-(V)s ~-(V)s'*.[26] The suffix is found in all Saamic languages (**riesler2006b** see also §**??** where the origin of attributive state marking in Saamic is dealt with in detail).

In Skolt Saami, the prototypical pairs of predicative and attributive adjective forms are equipped with the suffixes *-(V)d* PRED and *-(V)s* ATTR respectively, although other suffix pairs occur as well (**feist2015a**). Whereas the suffix *-(V)d* in (**??**) marks the predicative state of the adjective, the suffix *-(V)s* is an attributive state marker. The examples (**??**) show that the formative is invariable and does not alter its form in a plural or case marked noun phrase.

(77) Skolt Saami (personal knowledge)

a. Predicative

tät nijdd lij ***moočč-âd***
this girl is beautiful-PRED
'this girl is beautiful'

b. Attributive

i. *tät lij* ***moo´čč-es*** *nijdd*
this is beautiful-ATTR girl
'this is a beautiful girl'

ii. *täk liâ* ***moo´čč-es*** *niõđ*
this are beautiful-ATTR girl\PL
'these are beautiful girls'

iii. ***moo´čč-es*** *niõđ-i* *põrtt*
beautiful-ATTR girl-GEN.PL house
'the house of the beautiful girls'

[26] The palatalized variant occurs in Ter Saami and Kildin Saami.

In all Saamic languages, attributive (and predicative) state marking of adjectives is complex and determined by certain lexically defined classes and subclasses of adjectives. Many adjectives are marked only for attributive state but show the unmarked stem form in the predicative form. Consider for instance *neu´rr* [PRED] versus *neeu´r-es* [ATTR] 'bad', in Skolt Saami. In addition, in the predicative forms of several adjectives, suffixes other than *-(V)d* also occur. Finally, there are a few adjectives which also use the attributive suffix in their predicative forms (**feist2015a**).

In fact, a general tendency is noticeable in all Saamic languages: the differentiated morphological marking of predicative and attributive adjectives is being abolished in favor of using the pure or extended stem forms in both syntactic positions. As a result, attributive state marking seems to be in dissolution (**riesler2006b**). Several classes of adjectives, however, do not seem to be as affected by the functional spread of the juxtapositional type. In Skolt Saami, the anti-construct state marker is even used productively in several derived adjective classes, such as with the abessive adjectivizer.

(78) Derived adjectives in Skolt Saami (Notozero) (**senkevic-g1968**)

a. Attributive
päärn-t´em-es *neezzan*
child-ABESS.ADJZ-ATTR woman
'(a) woman without children'

b. Predicative
tät neezzan lij ***päärn-t´em***
this woman is child-ABESS.ADJZ
'this woman is without children'

Juxtaposition in Skolt Saami Whereas dependent marked attributive state is the prototypical type of adjective attribution marking in Skolt (as well as in the other Saamic languages), certain adjectives are never inflected in their attributive form, one instance being *nuõrr* 'young' (**feist2015a**).

(79) Skolt Saami (personal knowledge)

a. Attributive
tät lij ***nuõrr*** *nijdd*
this is young girl
'this is a young girl'

i. *täk liâ* ***nuõrr*** *niõđ*
this are young girl\PL
'these are young girls'

b. Predicative
täk niõđ liâ ***nuõr***
this girl is young\PRED.PL
'these girls are young'

The noun phrase type in which 'young' and other members of this adjectival class occur must be characterized as juxtaposition. Hence, Skolt Saami exhibits a second, minor adjective attribution marking device in addition to attributive state marking.

7.18.6.2 West Saamic

The five West Saamic languages are Northern, Lule, Pite, Ume and Southern Saami. They are spoken in northern Norway and Sweden and in the adjacent parts of northern Finland.

The default adjective attribution marking device in all West Saamic languages is anti-construct state marking, just as in East Saamic. Only the few members of a marginal subclass of adjectives are attributed by means of other devices. In general, West Saamic languages are similar to East Saamic in their high degree of irregularity in the morphological marking of attributive adjectives, although grammars of Northern Saami, usually taking a rather normative-descriptive approach (e.g., **nickel1990**; **sammallahti1998b**; **svonni2009a**), stress the systemic character of attributive versus predicative marking with the suffix *-(V)s* being the prototypical formative for attributive morphology.

For another West Saamic language, Pite Saami, and using exclusively corpus data **wilbur2014a** argues that the formative *-(V)s* is used much too irregularly to be considered a productive attributive suffix. Because of the considerable inconsistencies in morphological patterns between corresponding attributive and predicative adjectives, **wilbur2014a** generally prefers to analyze these two sets of adjectives simply as semantically and etymologically related, rather than morphologically derivable adjectives. However, even if a large part of these adjectives consists of suppletive pairs, the morpho-syntax of adjectives in Pite Saami shares one important characteristic with the other Saamic languages: whereas attributive adjectives never show morphological agreement, predicative adjectives agree (in NUMBER) with the subject noun phrase.

Head-driven agreement in Northern Saami For Northern Saami, the default attribution device is anti-construct state marking, like in all Saamic languages. A few adjectives, however, regularly show agreement with the head noun in number and case. In Northern Saami, the adjective 'good' and sometimes also the adjective 'bad' follow this type.

(80) Northern Saami (**nickel1990**)

a.	*buorre niibi* 'good knife'	[good(NOM:SG) knife(NOM:SG)]
b.	*buori niibbi*	[good\GEN:SG knife\GEN:SG]
c.	*buori niibá-i*	[good\GEN:SG knife-ILL:SG]
d.	*buori niibi-s*	[good\GEN:SG knife-LOC:SG]
e.	*buri-in niibbi-in*	[good-COM:SG knife-COM:SG]
f.	*buori-t niibbi-t*	[good-NOM:PL knife-NOM:PL]
g.	*buori-id niibbi-id*	[good-GEN:PL knife-GEN:PL]
h.	*buori-id(~ide) niibbi-ide*	[good-GEN:PL(~ILL:PL) knife-ILL:PL]
i.	*buri-in niibbiin*	[good\LOC:PL knife-LOC:PL]
j.	*buori-id(~iguin) niibbi-iguin*	[good-GEN:PL(~COM:PL) knife-COM:PL]
k.	*buorri-n niibi-n*	[good-ESS knife-ESS]

Note that the agreement inflection of the adjective can be characterized as defective because it does not distinguish all single case forms in the paradigm.

7.19 Indo-European

Indo-European is among the world's language families with the greatest geographic distribution. Most of the European languages belong to this family. But Indo-European languages are spoken as far East as on the South Asian subcontinent. The family can be divided into nine branches (**salminen2007**), all of which are represented in the present investigation.

The prototypical adjective attribution marking type in Indo-European is head-driven agreement. This type is also reconstructed for the Proto-Indo-European language (**decsy1991**; **watkins1998**). Due to the development of certain secondary types of adjective attribution marking devices, however, divergence is relatively high inside the Indo-European family. Furthermore, in several branches of Indo-European, head-driven agreement has been lost in favor of various other types of attribution marking (as will be shown in Part IV Diachrony).

Among the languages of northern Eurasia, the Indo-European family exhibits the highest diversity with regard to the number of possible adjective attribution marking devices. The following types are attested in different Indo-European languages:

- juxtaposition
- head-driven agreement
- construct-state marking
- anti-construct state marking
- anti-construct state agreement marking
- attributive nominalization
- incorporation.

7.19.1 Albanian

The Albanian branch of Indo-European is represented by the two languages Standard Albanian and Arvanitika.

Attributive nominalization + head-driven agreement in Albanian In both Albanian languages, adjectives normally follow the head noun and are marked with an article which links its host to the modified noun. Additionally, adjectives are equipped with agreement inflection suffixes co-referencing the NUMBER-, GENDER-, CASE- and SPECIES values of the head noun. The language thus exhibits an attributive marking device which is a combination of a phonologically free article (historically an attributive nominalizer) and agreement suffixes.

(81) Standard Albanian (**himmelmann1997**)

a. *një* *shok* ***i*** ***mirë***
one:M friend:INDEF:M ATTR:NOM.SG.M good:NOM.SG.M
'one good friend'

b. *shok=u* ***i*** ***mirë***
friend=DEF:NOM.SG.M ATTR:NOM.SG.M good:NOM.SG.M
'the good friend'

c. *shok=un* ***e*** ***mirë***
friend=DEF:ACC.SG.M ATTR:ACC.SG.M good:ACC.SG.M
'the good friend (acc.)'

Note that the circum-positioned agreement marker also occurs with predicative adjectives.

(82) Standard Albanian (**demiraj1998**)
shok=u *është* ***i*** ***bukur***
friend-DEF:NOM.SG.M be.3SG.PRS ATTR:NOM.SG.M pretty:NOM.SG.M
'the friend is pretty'

Since adjectives in attributive and predicative position are both equipped with the circumfixed agreement marker the language seems to belong simply to the head-driven agreement type. However, true predicative adjectives are not found in Albanian. Instead, attributive adjectives in headless noun phrases are used in predicative position. This is evidenced by case agreement of predicates.

(83) Standard Albanian (**demiraj1998**)
a. *Agimi* *u kthye* ***i*** ***dëshpëruar***
Agimi(NOM.SG.M) returned ATTR:NOM.SG.M sorrowful:NOM.SG.M
'Agim returned sorrowfully'
b. *Agimi(ACC.SG.M) e pashtë* ***të*** ***dëshpëruar***
Agimi I saw ATTR:ACC.SG.M sorrowful:ACC.SG.M
'I saw Agimi sorrowful'

On the other hand, the similar agreement behavior of attributive and predicative adjectives seems to indicate the absence of specific attributive morpho-syntactic marking. However, the attributive article is polyfunctional and can also link other adnominal attributes in addition to adjectives to the modified noun. The analysis of adjective attribution marking in Albanian as belonging to the attributive nominalization type (in combination with head-driven agreement) thus seems justified.

(84) Standard Albanian (**demiraj1998**)
a. *roman-i* ***i*** *tretë*
novel-DEF:NOM.SG.M ATTR:NOM.SG.M third
'the third novel'
b. *libr-i* ***i*** *nxënës-it*
book(M)-DEF:NOM.SG.M ATTR:NOM.SG.M pupil-DEF:GEN/DAT.SG
'the pupil's book'

Head-driven agreement in Albanian Note, however, that the occurrence of the attributive article is restricted to a lexically defined subclass of adjectives in Albanian: only the so-called "article adjectives" are regularly marked with the article. Other adjectives are marked with head-driven agreement affixes alone.

(85) Standard Albanian (**himmelmann1997**)

a. *shok=u* ***besnik***
friend-DEF:NOM.SG.M true:NOM.SG.M
'the faithful friend'

b. *një* *shok* ***besnik***
one:M friend:INDEF:M true:NOM.SG.M
'one faithful friend'

Again, predicative adjectives behave similar to attributive adjectives.

(86) Standard Albanian (**demiraj1998**)

a. Predicative agreement of "article adjectives"
shok=u *është* ***i*** ***bukur***
friend-DEF:NOM.SG.M be.3SG.PRS ATTR:NOM.SG.M pretty:NOM.SG.M
'the friend is pretty'

b. Predicative agreement of "simple" adjectives
shok=u *është* ***besnik***
friend-DEF:NOM.SG.M be.3SG.PRS true:NOM.SG.M
'the friend is faithful'

Attributive nominalization + head-driven agreement in Arvanitika Adjective attribution marking in Arvanitika is very similar to Standard Albanian. One adjective class shows head-driven agreement marking by means of suffixes. The second adjective class is cognate with the so-called "article adjectives" in Albanian and exhibits attributive nominalization.

(87) Arvanitika (**sasse1991**)

a. *ɲə́* *djáʎə* ***i-mírə***
one:M boy:INDEF.M M-good:M
'one good boy'

b. *djáʎi* ***i-mírə***
boy:DEF.M M-good:M
'the good boy'

Unlike in Standard Albanian, however, the preposed attributive nominalizer in Arvanitika is a phonologically bound formative. This is evidenced by its phonological behavior in adjective compounds, where the marker remains in its position bound to the adjective stem.

(88) Arvanitika (**sasse1991**)

a. ***miso-i**-ngránə* / ***miso-tə**-ngránə*
half-M-mounted:M half-ACC.M-mounted:M
'half-mounted'

b. * *i-miso-ngránə* / *tə-miso-ngránə*

Example (**??**) shows that the compound degree word *miso-* does not move between the adjective stem and the attributive nominalizer. Consequently, the nominalizer can be characterized as a clitic (because it is phonologically bound but morpho-syntactically free) which always attaches on a fixed position, i.e., on the left of the adjective stem.[27]

7.19.2 Armenian

Armenian is a branch consisting only of two closely related varieties, of which only the Eastern Armenian standard language is considered here.

Juxtaposition in Eastern Armenian In the unmarked construction, attributive adjectives are unmarked and precede the modified noun.

(89) Eastern Armenian (**ajello1998**)

a. ***bari*** *gorc*
good work(NOM.SG)
'good work'

b. ***bari*** *gorc-s*
good work-ACC.PL
'good work (acc.)'

Head-driven agreement in Armenian A few monosyllabic adjectives show head-driven agreement marking in Armenian.

[27] Note, however, that the agreement categories CASE/NUMBER/GENDER are merged into several differentiated morphemes in the suffixed part of the circumfix **sasse1991**

In theory, however, all adjectives in an emphatic construction can occur in a noun phrase with reversed constituent order. In "emphatic position" **ajello1998** i.e., in contrastive focus attributive adjectives show agreement in case and number as a rule.

(90) Eastern Armenian (**ajello1998**)
bazum gorc-s ***bari-s***
much work-ACC.PL good-ACC.PL
'much GOOD work (acc.)'

7.19.3 Indo-Iranian

Indo-Iranian (aka Aryan) is a major branch within Indo-European. But only a few Indo-Iranian languages belonging to the Iranian and Indo-Aryan subbranches are spoken in northern Eurasia and thus considered here. Most other Indo-Iranian languages are spoken in the Middle East and in South Asia and hence outside the investigated geographic area.

7.19.3.1 Indo-Aryan

Indo-Aryan (aka Indic) is a large subbranch of Indo-Iranian, most member languages of which are spoken on the South Asian subcontinent. Outlier languages, spoken in northern Eurasia include Parya, a language which was recently discovered in Tajikistan in Inner Asia (**masica1991**), and the group of Romani languages. Several varieties of Romani are spoken all over Europe. Some of them are not mutually intelligible. Rather than being one single language, Romani is thus a group of languages which comprise at least the four subbranches Vlax Romani, Balkan Romani, Central Romani and North Romani with several sub-varieties in each of them (**halwachs-etal2002**).

The default type of adjective attribution marking in Indo-Aryan languages is head-driven agreement in noun phrases with head-final constituent order (**masica1991**). Agreement features in the Romani languages are GENDER and NUMBER, and in most varieties also CASE. The unmarked constituent order in all varieties of Romani is adjective-noun.

Head-driven agreement in Burgenland Romani In the Burgenland variety of Romani, adjectives normally show agreement in gender, number and also case with the head noun. Case agreement, however, can be characterized as defective,

since all attributive adjectives preceding oblique cases have one similar oblique form.

(91) Burgenland Romani (**halwachs-etal2002**)

a. ***bar-o*** *phral*
big-NOM:M.SG brother(M)
'big brother'

b. ***bar-i*** *phen*
big-NOM:F.SG sister(F)
'big sister'

Juxtaposition in Burgenland Romani A minor lexically defined subclass of adjectives in Burgenland Romani is indeclinable and juxtaposed to the head noun.

(92) Burgenland Romani (**halwachs-etal2002**)

a. ***schukar*** *phral*
beautiful brother(M)
'beautiful brother'

b. ***schukar*** *phen*
beautiful sister(F)
'beautiful sister'

Attributive nominalization in Vlax Romani **hancock1995** describes the use of a "repeated definite article" in contrastive focus constructions in Vlax Romani.

(93) Vlax Romani (**hancock1995**)

a. Head-driven agreement (unmarked construction)
o ***baro*** *raklo*
DEF big boy
'the big boy'

b. Attributive nominalization (emphatic construction)
o *raklo* ***o*** ***baro***
DEF boy ATTR big
'the BIG boy'

7.19.3.2 Iranian

The second subbranch of Indo-Iranian is formed by Iranian languages, only a few of which are spoken in northern Eurasia.

A well-known characteristic of noun phrase structure in Iranian languages is the occurrence of the Ezafe construct marking which licenses the attribution of adjectives (and other syntactic classes of modifiers). The Iranian languages surveyed in the present investigation, however, exhibit some diversity in this respect. Attributive construct state marking occurs regularly only in the western Iranian languages Northern Kurdish (aka Kurmanji, Kirmancî) and Tajik.

Attributive construct state in Tajik Tajik follows the Iranian prototype and exhibits a head-marking construct state marking suffix.

(94) Tajik (**rastorgueva1963**)

a. *duxtar-**i** **xušrūj***
girl-ATTR beautiful
'a pretty girl'

b. *duxtar-on-**i** **xušrūj***
girl-PL-ATTR beautiful
'pretty girls'

Anti-construct in Northern Talysh The constituent order in noun phrases in Northern Talysh is adjective-noun. The language is exceptional among the Iranian (and Indo-European) languages considered here in exhibiting dependent marking anti-construct state instead of head-marking construct state as the default type of adjective attribution marking.

(95) Northern Talysh (**schulze2000**)

a. ***āğəlmānd-a** odam-on*
clever-ATTR man-PL
'clever people'

b. ***yol-a** di*
big-ATTR tree
'(a) big tree'

Juxtaposition in Ossetic Ossetic is another exceptional language among Iranian, because the language exhibits juxtaposition as the default type of adjective attribution marking.

(96) Ossetic (**abaev1964**)

a. Simple noun
færǽt / fǽræt
ax ax\DEF
'axe' / 'the axe'

b. Noun phrase with adjectival modifier
***cyrg'**-fǽræt / **cýrg'**-færæt*
sharp-axe sharp-ax\DEF
'sharp axe' / 'the sharp axe'

Stress patterns provide evidence for the analysis of Ossetic noun phrase structure as phonological compounds. According to **abaev1964** "syntactically connected word groups" (such as noun phrases) are marked by single stress. Note that stress, moving from the second to the first syllable marks definiteness in Ossetic (**abaev1964**). There is, however, no evidence that the compounded adjectives are syntactically incorporated.

Note that attributive construct state marking which is cognate with the Ezafe in other Iranian languages occurs in Ossetic as well, but its use is restricted to certain "emphatic", i.e., contrastive focus constructions **thodarson1989**

7.19.4 Baltic

7.19.4.1 East Baltic

The Baltic languages form a small branch among Indo-European and are represented in the present survey only by the two languages Lithuanian and Latvian. Both belong to the eastern subbranch of Baltic. All languages from the former western branch of Baltic are extinct.

Two types of adjective attribution marking occur in modern Baltic languages: head-driven agreement and anti-construct state agreement. In the descriptive literature on Baltic languages, however, these two noun phrase types are normally not ascribed to syntax, but are described as different agreement declension types determined by the definite or indefinite semantics of the noun phrase.

In §?? of Part II (Typology) I have already argued extensively in favor of a syntactic differentiation of these two agreement marking devices in Baltic (as well

as in various Slavic) languages. Consequently and for the sake of completeness, examples of head-driven agreement marking (the so-called indefinite declension) and anti-construct state agreement marking (the so-called definite declension) in Latvian and Lithuanian will be repeated in the following paragraphs.

Head-driven agreement in Latvian and Lithuanian Adjectives modifying indefinite nouns show head-driven agreement in Latvian and Lithuanian.

(97) a. Latvian (**dahl2015a**)
liel-a *māja*
big-F.NOM.SG house(F)
'a large house'

b. Lithuanian (**bechert1993**)
gẽr-as *profèsorius*
good-NOM.SG.M professor(M)
'a good professor'

Anti-construct state agreement in Latvian and Lithuanian Adjectives modifying definite nouns show anti-construct state agreement marking in Latvian and Lithuanian.

(98) a. Latvian (**dahl2015a**)
liel-ā *māja*
big-ATTR:F.NOM.SG house(F)
'the large house'

b. Lithuanian (**bechert1993**)
ger-àsis *profèsorius*
good-ATTR:NOM.SG.M professor(M)
'the good professor'

7.19.5 Celtic

The modern Celtic languages belong to two main branches: Gaelic and Brittonic. By and large, all Celtic languages have preserved the Proto-Celtic noun phrase structure, including head-driven agreement marking on attributive adjectives and noun-adjective constituent order.

7.19.5.1 Gaelic

Head-driven agreement in Scots Gaelic In Scots Gaelic (aka Scottish Gaelic) adjectives (as well as other modifiers) show agreement in GENDER, NUMBER, and CASE.

(99) Scots Gaelic (**macauley1992**)

a. *an* *cù* ***dubh***
DEF:M dog(M) black\M
'the black dog'

b. *a'* *chaora* ***dhubh***
DEF:F sheep(F) black\F
'the black sheep'

Similar agreement patterns as in Scots Gaelic, with non-linear marking by means of word-initial permutation, are found in Irish **odochartaigh1992** In the third Gaelic language Manx, however, most adjectives are used in an invariable form. Only a certain subclass of monosyllabic adjectives have preserved number agreement in Manx (**thomsen1992**).

7.19.5.2 Brittonic

The tendency towards a loss of agreement inflection of adjectives is also noticeable in the languages of the Brittonic branch of Celtic. Adjective inflection seems to be most intact in Welsh with preserved gender and number agreement (**thomas1992a**). Breton and Cornish exhibit only agreement in gender (**ternes1992 thomas1992b**).

7.19.6 Germanic

The modern Germanic languages belong to two branches: North and West Germanic. The third Germanic subbranch, East Germanic, is extinct and is not considered here.

The constituent order of adjective and noun is relatively strictly head-final in all modern Germanic languages.[28] Most Germanic languages have also preserved the inherited agreement marking on attributive adjectives. But several secondary

[28] The exclusive adjective-initial constituent order in modern Germanic languages is clearly innovative. In documents of all Old Germanic languages, the order of adjective and noun was still relatively free (**heinrichs1954**).

attributive marking devices have evolved at different stages in the history of Germanic.

The following noun phrase types occur inside the Germanic branch of Indo-European:

- Anti-construct state agreement
- Anti-construct state agreement + head-driven agreement
- Attributive article + head-driven agreement
- Head-driven agreement
- Incorporation.

Whereas head-driven agreement and attributive nominalization are attested for the earliest stages of Germanic, adjective incorporation is a rather recent innovation (cf. §??).

7.19.6.1 West Germanic

The most common type of adjective attribution marking in West Germanic languages is head-driven agreement. In most languages of this group, this is the only existing type.

Anti-construct state agreement in German Attributive adjectives in German show head-driven agreement according to the features GENDER, NUMBER, CASE and SPECIES. The complete agreement paradigm was illustrated in Part II (Typology) (Figure ?? on page ??). Note that the adjective agreement paradigm of German exhibits a high degree of syncretism due to merger of originally differentiated formatives. The whole paradigm distinguishes only the four suffixes *-e, -em, -en, -er, -es*.

(100) Attributive adjectives in German (personal knowledge)

a. *ein* ***hoh-es*** *Haus*
INDEF high-INDEF.N house(N)
'a high house'

b. *das* ***hoh-e*** *Haus*
DEF high-DEF.N house(N)
'the high house'

c. ***hoh-e*** *Häus-er*
high-PL house-PL
'high houses'

d. *der* ***hoh-en*** *Häus-er*
DEF:PL.GEN high-DEF.PL.GEN house-PL.GEN
'of the high houses'

Attributive and predicative adjectives are morpho-syntactically differentiated in German (and the other West Germanic languages, except English): whereas attributive adjectives show head-driven agreement, predicative adjectives are used in an invariable form. Given the definition of dependent marking attributive state which is applied here (see §??), German thus exhibits anti-construct state agreement marking of attributive adjectives.

(101) Predicative adjectives in German (personal knowledge)

a. *das / ein Haus is* ***hoch***
DEF INDEF house(N) is high
'a / the house is high'

b. *(die) Häus-er sind* ***hoch***
DEF house-PL are high
'(the) houses are high'

Attributive nominalization + head-driven agreement in Yiddish The default noun phrase structure in Yiddish is similar to the other West Germanic languages. Head-driven agreement occurs as the default type of attribution marking of adjectives. In contrastive focus constructionc, however, adjectives and other modifiers follow the modified noun in an attributive nominalization construction.

(102) Yiddish (Eastern) (**jacobs-etal1994**)

a. Head-driven agreement (unmarked)

i. *a* ***sheyn*** *meydl*
INDEF:F pretty:INDEF.F girl(F)
'a pretty girl'

ii. *di* ***grine*** *oygn*
DEF:PL green:DEF.PL eye:PL
'the green eyes'

b. Attributive nominalization (contrastive focus)

i. *a* *meydl* ***a*** ***sheyne***
INDEF:F girl(F) ATTR:INDEF.F pretty:ATTR:INDEF.F
'a PRETTY girl'

ii. *di* *oygn* ***di*** ***grine***
DEF:PL eye:PL ATTR:DEF.PL green:DEF.PL
'the GREEN eyes'

Incorporation in English English is the only West Germanic language where head-driven agreement is missing completely because the original Germanic agreement inflection on adjectives was lost.

(103) English (personal knowledge)

a. *a* ***pretty*** *girl*
INDEF pretty girl

b. *the* ***pretty*** *girl*
DEF pretty girl

c. ***pretty*** *girl-s*
pretty girl-PL

Attributive adjectives cannot, however, occur in headless noun phrases in English but are obligatorily marked with an article used as dummy head.

(104) English (personal knowledge)

a. *a* */ the* ***smart one***
INDEF DEF smart ART

b. ***smart one-s***
smart ART-PL

The marker *one* in English (originating from the homophonous numeral *one*) is a prototypical instance of an article: it constitutes a phonologically free grammatical word which is the target of agreement.

Given that attributive adjectives cannot occur other than syntactically bound to a head noun, the regular noun phrase type in English is best analyzed as incorporation. Note that the article is not an attribution marking device in the proper sense. Even though the marker projects a noun phrase by syntactic nominalization, this noun phrase does not modify a higher noun. The nominalization strategy can only be used in noun phrases with an empty lexical head.

(105) English (personal knowledge)

a. *a smart girl*
[NP INDEF A smart N girl]

b. *a smart **one***
[NP INDEF A smart HEAD]

c. * *a smart **one** girl*
[NP [NP INDEF A smart HEAD] N girl]

Because attributive adjectives in English are obligatorily bound to a syntactic head and because the nominalizer ("dummy head") cannot occur in noun phrases modifying a higher head, English exhibits neither true juxtaposition nor attributive nominalization.

7.19.6.2 North Germanic

With regard to existing attribution marking devices, the North Germanic languages exhibit even a higher degree of diversity than West Germanic. This is especially true if major sub-varieties are considered as well. Practically all types attested in West Germanic occur here as well, including adjective incorporation which is otherwise scarcely attested in the languages of northern Eurasia.

Head-driven agreement in North Germanic Although head-driven agreement marking constitutes the prototypical adjective attribution marking device in North Germanic, the adjective agreement paradigms across the different languages reflect the ongoing decline in differentiated categories.

In **Icelandic**, adjectives inflect for the agreement features GENDER, NUMBER, CASE and SPECIES. The adjective agreement paradigm of Modern Icelandic (Table **??** in §**??**) is thus relatively similar to Old Icelandic even though the different case endings are already merged in the definite paradigm.

In **Danish**, there is no agreement feature CASE, while GENDER is marked on the attributive adjective only in indefinite noun phrases. In definite noun phrases, the attributive adjective is marked with an invariable definite agreement suffix (Table **??**).

The **Western Jutlandic** dialect of Danish is most innovative with regard to the decline of agreement features because it has almost completely lost its agreement features and thus resembles English (Table **??**).

Table 7.1: Agreement paradigm for the adjective 'yellow' in Danish (personal knowledge)

	UTR.SG	N.SG	PL
INDEF	gul	gul-t	gul-e
DEF	gul-e	gul-e	gul-e

Table 7.2: Agreement paradigm for the adjective 'yellow' in Western Jutlandic (in phonemic transcription) (**ringgaard1960**)

	SG	PL
INDEF	gul?	gul
DEF	gul	gul

Anti-construct state + head-driven agreement in Swedish Swedish, Norwegian,[29] and Faroese exhibit two adjective attribution marking morphemes simultaneously: an inflectional suffix expressing the agreement features GENDER, NUMBER and SPECIES (but the indefinite utrum gender form of the adjective is always unmarked) plus an article (which again is not found in the indefinite plural form).

In the (North-)Germanic and typological linguistic tradition, the definite noun phrases with adjectives have most often been characterized as "double definite" (cf. **kotcheva1996a borjars1994 julien2003 plank2003**). This makes sense from a historical perspective because the articles (Swedish *den, det, de*) are cognate with the Old Germanic demonstratives which developed into definite markers (cf. German *der, die, das* or English *the*). Synchronically, however, the articles in the North Germanic languages with so-called double definiteness (Swedish, both Norwegian languages, Faroese) are not definiteness markers. Unlike in West Germanic, definiteness is exclusively expressed by an inflectional suffix (Swedish *-(e)n* UTR, *-(e)t* N, *-n* PL.)

Unlike in West Germanic languages, where the definite markers are noun phrase markers always attach at the left edge of the phrase, the presence or absence of the cognate articles *den* UTR, *det* N, *de(m)* PL in Swedish is determined by

[29] The two Norwegian standard languages Dano-Norwegian (Norwegian *bokmål*) and New Norwegian (Norwegian *nynorsk*) do not differ in their marking of adjective attribution and they will simply be referred to as Norwegian

the availability of an adjective and not the referential status of the noun phrase.

(106) Swedish (personal knowledge)

a. *(*det)* *hus-**et***
ATTR:DEF.N house-DEF:N
'the house'

b. **(**det**)* ***hög-a*** *hus-et*
ATTR:DEF.N high-DEF.N house-DEF.N
'the high house'

c. **(**det**)* ***hög-a***
ATTR:DEF.N high-DEF.N
'the high one' (about a house)

Example (??) shows how the article can neither attach to a noun nor can an adjectival modifier in a definite noun phrase occur without being marked by the article.[30] Since the definite value of the feature SPECIES is always marked by the respective definite inflectional noun suffixes and since the article only attaches to adjectives, the latter cannot be analyzed as anything but a morpho-syntactic device, i.e., as an adjective attribution marker.

In definite noun phrases, Swedish thus exhibits a circumfixed adjective attribution marking device combined by head-driven agreement inflection plus the article. It is plausible that the article developed from an attributive nominalizer. Its use with adjectives in headless noun phrases, as in (??) resembles attributive nominalization. There is, however, no evidence that the adjective marked by the article is part of a complex constituent (i.e., a headless noun phrase) modifying a noun. According to the definition of attributive nominalization presented in §?? of Part II (Typology), the article in Swedish is thus not a syntactic nominalizer. Its function is the licensing of the attributive state of the adjective along with marking of head-driven agreement. Since head-driven agreement is additionally marked by inflectional suffixes, the Swedish noun phrase exhibits circumpositioned (i.e., phonologically free and phonologically bound) agreement marking.

Note that the circum-positioned agreement marker only occurs with attributive adjectives. Predicative adjectives, on the other hand, exhibit "pure" gender

[30] The expression *det hus* is grammatical only with the homophonous demonstrative *det*, similarly (but restricted to certain regiolects) *det hus-et*. Even the expressions *höga hus-et* is possible for some expression similar to English *White house*. Note also that possessive pronouns replace the article: *min hög-a hus* [POSS:1SG high-DEF.N house(N] 'my high house'.

and number agreement (**??**). The analysis of adjective attribution marking in Swedish as belonging to anti-construct state agreement marking is thus justified.

(107) Predicative adjectives in Swedish (personal knowledge)

a. *kåken är **hög*** ‘the (bad) house is high’ [UTR]

b. * *kåken är **en hög / den hög-a***

c. *huset är **hög-t*** ‘the house is high’ [N]

d. * *huset är **ett hög-t / det hög-a***

e. *husen är **hög-a*** ‘the houses are high’ [PL]

f. * *husen är **de hög-a***

Table 7.3: Agreement paradigm for the adjective *hög* ‘high’ in Swedish (personal knowledge); *stuga* (UTR) ‘cabin’, *hus* (N) ‘house’

	INDEF			DEF		
UTR.SG	en	**hög-Ø**	stuga	**den**	**hög-a**	stuga-n
N.SG	ett	**hög-t**	hus	**det**	**hög-a**	hus-et
PL		**hög-a**	stug-or	**de**	**hög-a**	stug-or:na

Adjective incorporation in Västerbotten Swedish The dialect spoken in the Västerbotten province in northern Sweden exhibits adjective incorporation as a regular type of adjective attribution marking.

(108) Västerbotten Swedish (**holmberg-etal2003**)

a. ***grann**-kweinn-a*
pretty-woman-DEF
‘the pretty woman’

b. *en* ***grann**-kweinn*
INDEF pretty-woman
‘a pretty woman’

Adjective incorporation also occurs in several other northern North Germanic dialects of Sweden, Finland and Norway. Whereas adjective incorporation is the

default type in Västerbotten Swedish,[31] its occurrence is restricted to definite noun phrases in most other dialects where this type it attested.

Attributive adjectives cannot occur in indefinite headless noun phrases in Västerbotten Swedish but are obligatorily bound to an article used as dummy head.

(109) Västerbotten Swedish (**holmberg-etal2003**; **delsing1996b**)

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'

7.19.7 Hellenic

The Hellenic branch of Indo-European is represented by a single language: Modern Greek.

Head-driven agreement and attributive nominalization + head-driven agreement in Greek Attributive adjectives in Greek show agreement in GENDER, NUMBER and CASE.[32]

The unmarked constituent order in Greek is adjective-noun, as in (**??**). The reverse constituent order (noun-adjective), however, is commonly used as well and marks contrastive focus on the attribute, as in (**??**).

(110) Greek (**ruge1986**)

a. Head-driven agreement

i. *to* ***kokino*** *aftokinito*
DEF:M red:M car(M)
'the red car'

b. Attributive nominalization

[31] In indefinite noun phrases, however, adjective incorporation is often restricted to monosyllabic adjective stems: *en* ***grann****-kweinn* but **en* ***vacker****-kweinn* 'a **pretty** woman'. Furthermore, a certain semantic relation between noun and adjective seem to be obligatory: (incorporation) *n* ***ny****-bil* 'a **new** car (straight from the factory)', *n* ***ny*** *bil* 'a **new** car (new for me)', (incorporation) **n* ***ny****-hunn* 'a **new** dog', *n* ***ny*** *hunn* 'a **new** dog (new for me)' (**holmberg-etal2003**).

[32] A minor class of loan adjectives in Greek belong to a different noun phrase type, juxtaposition, because they do not inflect at all (**ruge1986**).

i. Contrastive focus on the attribute
   to *aftokinito* ***to*** ***kokino***
   DEF:M car(M) ATTR:M red:M
   'the RED car (not the blue one)'

ii. Contrastive focus on the noun
   to ***kokino*** *to* *aftokinito*
   ATTR:M red:M DEF:M car(M)
   'the red CAR (not the buss)'

Note that the noun can move to the contrastive focus position as well, as in (??).

Example (??) illustrates the use of the article *to* in two different syntactic functions: whereas *to* DEF is a determiner marking the noun phrase as definite, *to* ATTR is an attributive marker (i.e., a true article) attaching to the adjective noun phrase internally. Attribution of the adjective (in contrastive focus) in (??) is marked by means of attributive nominalization. The article marks the adjective as phrasal constituent, i.e., as a syntactic complement to the noun.

7.19.8 Romance

All Romance languages exhibit head-driven agreement marking as the main and default adjective attribution marking device. The prototypical agreement features characteristic of most modern Romance languages are NUMBER and GENDER. A third agreement feature, CASE, was present in earlier stages of Romance but has disappeared in the modern languages.

Three noun phrase types have existed in the Romance branch from its earliest stages:[33]

- head-driven agreement
  - noun-adjective order
  - adjective-noun order
- attributive nominalization.

The unmarked and prototypical noun phrase type in Romance is head-driven agreement with the adjective following the noun. Besides the basic head-initial constituent order, most Romance languages exhibit a small subgroup of very common adjectives, such as 'good–bad, young–old, small–large', which normally

[33] A minor class of adjectives belong to a different noun phrase type, juxtaposition, because they do not inflect at all.

precede the head noun (**posner1996** cf. also **silvestri1998**). However, most other adjectives can also precede the noun in the modern Romance languages. This reversed constituent order is regularly determined by semantics-pragmatics in Rumanian and is used to give these adjectives a certain emphasis or contrastive focus, as in the following examples from Rumanian (**??**) and Italian (**??**).

(111) a. Rumanian (**beyer-etal1987**)

i. *băiat=ul* ***bun***
boy=DEF good
'the good boy'

ii. ***bun****=ul* *băiat*
good=DEF boy
'the GOOD (i.e., different) boy'

b. Italian (**posner1996**)

i. *un* *vestito* ***nuovo***
INDEF dress new
'a (brand-)new dress'

ii. *un* ***nuovo*** *vestito*
INDEF new dress
'a new (i.e., different) dress'

Note that the definite marker in Rumanian is not connected with attribution marking on adjectives. Even though the marker can occur on the attributive adjective which precedes the noun in contrastive use (**??**), definiteness is a purely morpho-semantic feature in Rumanian and is not assigned by syntax (see also §**??** of Part I Preliminaries).

The common distinction between an "emphatic" adjective preceding a noun and a "descriptive" adjective following a noun goes probably back to the earliest stages of Romance, although it is first attested in Classical Latin (**posner1996**).

Head-driven agreement in Italian In Italian, as in the other Romance languages, the agreement features GENDER and NUMBER are marked on adjectives and on other modifiers within the noun phrase.

(112) Italian (personal knowledge)

a. *la casa **alt-a***
DEF:F house(F) high-F
'the high house'

b. *le cas-e **alt-e***
DEF:PL house-PL high-PL
'the high houses'

Attributive nominalization in Rumanian Beside the default type of head-driven agreement (with either noun-adjective or adjective-noun constituent order), Standard Rumanian (aka Daco-Rumanian) exhibits attributive nominalization as a differentiated third type of adjective attribution marking. The agreement paradigm of the attributive nominalizer (traditionally labeled "adjective article" in the grammatical descriptions of Rumanian) is shown in Table **??**. The use of the non-

Table 7.4: Agreement paradigm of the attributive article in Rumanian (**beyer-etal1987**).

	F	N	M
SG	cea	cel	
PL	cele		cei

obligatory attributive marker emphasizes the adjective following a noun (**beyer-etal1987 posner1996**). But it is also regularly used to mark definite headless noun phrases, as in the following example.

(113) Rumanian (**beyer-etal1987**)
*Punct-e=le **cele** **negr-e** se disting mai bine decît*
dot-PL=DEF.M.PL ATT:M.PL black-PL are distinguishing COMPAR better than
***cele** **cenuşi-i**.*
ATT:M.PL grey-PL
'The black dots are more distinguishable than the grey ones.'

The content of this marker, besides licensing of the attributive relation, is not clearly defined in descriptions of Rumanian. The article seems to regularly mark definite headless adjectives and superlative adjectives. **kramsky1972** compares the function of the article with that of the definite marker and describes the function of the attributive article in Rumanian as a "deictic reactualizer" because it

has a referential function but can co-occur with the definite marker (**??**). Note, however, that the definite marker is absent in a noun phrase with reversed constituent order marking contrastive focus (**??**).

(114) Rumanian (**beyer-etal1987**)

a. *poet=ul* ***cel*** ***mai*** ***mare***
poet(M)=DEF.M ATT:M.SG SUPER great
'the greatest poet'

b. ***cel*** ***mai*** ***mare*** *poet*
ATT:M.SG SUPER great poet(M)
'the GREATEST poet'

7.19.9 Slavic

Slavic (aka Slavonic) forms a branch inside the Indo-European family. All Slavic languages are spoken in Europe, except Russian, which is also spoken in North Asia.

The prototypical type of adjective attribution marking is head-driven agreement. The prototypical agreement features characteristic of Slavic languages are NUMBER, GENDER and CASE. In the closely related South Slavic languages Bulgarian and Macedonian however, case inflection of nouns and adjectives has been lost.

Beside head-driven agreement, anti-construct state agreement arose in Slavic languages as a secondary type of adjective attribution marking. The opposition between head-driven and anti-construct state agreement can be traced back to all Old Slavic languages and already existed in the oldest Slavic manuscripts, the best documented of which are from Old Bulgarian (aka Old Church Slavonic). To a certain extent, this state of development is still reflected in South Slavic. In most other modern Slavic languages, however the opposition between the two types was lost by abolishing one or the other type.

Basically, the modern Slavic languages belong to three types and exhibit the following three attribution marking devices:

- exclusively head-driven agreement
- exclusively anti-construct state agreement
- simultaneously head-driven agreement and anti-construct state agreement
- attributive nominalization.

Constituent order in Slavic can be described as basically adjective-noun, although there is much variation across the single languages. The reversed order of constituents is often possible but in some languages it is restricted to "emphasized" constructions or poetic language.

7.19.9.1 West Slavic

All West Slavic languages exhibit head-driven agreement as the exclusive type of adjective attribution marking.

Head-driven agreement in Lower Sorbian Lower Sorbian exemplifies a Slavic language with head-driven agreement as the exclusive type of adjective attribution marking. Attributive adjectives in Lower Sorbian show agreement in gender, number and case.

(115) Lower Sorbian **(janas1976)**

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

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

c. ***dobr-e*** *człowjek-y*
good-NOM.PL person-NOM:PL
'good people'

7.19.9.2 East Slavic

All three East Slavic languages Belorussian, Russian and Ukrainian exhibit anti-construct state agreement marking. There is, however, a tendency to merge attributive ("long") and predicative ("short") adjective agreement declension classes, yielding pure head-driven agreement as in West Slavic.

Anti-construct state agreement in Russian 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 forms.[34]

[34] This is true for the stylistically marked "short form adjectives", see in more detail §??.

(116) Russian (personal knowledge)

a. Attribution

i. ***krasiv-yj*** *mal'čik*
beautiful-ATTR:M.NOM boy(F)
'a handsome boy'

ii. ***krasiv-ogo*** *mal'čik-a*
beautiful-ATTR:M.GEN boy-M.GEN
'of a handsome boy'

iii. ***krasiv-aja*** *devuška*
beautiful-ATTR:F.NOM girl(F)
'a pretty girl'

b. Predication

i. *Etot mal'čik* ***krasiv***
DEM:M boy(M) beautiful:M
'this boy is handsome'

ii. *Eta devuška* ***krasiv-a***
DEM:F tower(F) high-F
'this girl is pretty'

The agreement suffixes of attributive and predicative adjectives clearly belong to different paradigms (cf. Table ??). The so-called long agreement suffixes (??) mark the values of the morphological agreement features. Simultaneously, they license the (morpho-syntactic) attributive relation inside the noun phrase (cf. also the discussion in §??).

Table 7.5: Attributive and predicative adjective declension in Russian (personal knowledge) for nominative case

	M	F	N	PL
ATTR	–yj/–ój	–aja/–ája	–oje/–óje	–yje/–ýje
PRED	Ø	–a	–o	–y/–i

7.19.9.3 South Slavic

All South Slavic languages exhibit head-driven agreement marking as the default type of adjective attribution marking. In Serbo-Croatian (aka Bosnian-Croatian-

Montenegrin-Serbian) and Slovenian, anti-construct state agreement marking occurs as a secondary type. Even attributive nominalization is attested in Slovenian.

Head-driven agreement in Bulgarian Attributive adjectives in Bulgarian show agreement in the features GENDER and NUMBER.

(117) Bulgarian (personal knowledge)[35]

a. Indefinite noun phrase

i. ***dobăr*** *i* ***vesel*** *măž*
good:M and cheerful.M man(M)
'good and cheerful man'

ii. ***dobr-a*** *i* ***vesel-a*** *žena*
good-F and cheerful-F woman(F)
'good and cheerful woman'

iii. ***dobr-i*** *i* ***vesel-i*** *žen-i*
good-PL and cheerful-PL woman-F.PL
'good and cheerful women'

b. Definite noun phrase

i. ***dobr-ij***=*ăt* *i* ***vesel-ij***=*ăt* *măž*
good:M=DEF.M and cheerful:M=DEF.M man(M)
'the good and cheerful man'

ii. ***dobr-a***=*ta* *i* ***vesel-a***=*ta* *žena*
good-F=DEF.F and cheerful-F=DEF.F woman(F)
'the good and cheerful woman'

iii. ***dobr-i***=*te* *i* ***vesel-i***=*te* *žen-i*
good-PL=DEF.PL and cheerful-PL=DEF.PL woman-PL
'the good and cheerful women'

Anti-construct state agreement in Serbo-Croatian Serbian (similar to the other varieties of Serbo-Croatian) exemplifies a Slavic language which exhibits both head-driven agreement and anti-construct state agreement in different functions.

[35] The stem allomorph with inserted *-ăˑ-* in M.SG is the result of a phonological process. The stem allomorph with the extension *-ij-* is morpho-phonological and triggered by the definite marker. Note that *-ij-* is a reflex of the Old Bulgarian anti-construct state agreement marker.

head-driven agreement constitutes the basic type of adjective attribution marking in Serbian. Most adjectives, however, have "double forms" (**kramsky1972**). Consider the following example.

(118) Serbian (**zlatic1997**)

a. Indefinite noun phrase ("pure" head-driven agreement)
dobar, ***veseo*** *čovek*
good:M cheerful:M person(M)
'a good, cheerful person'

b. Definite noun phrase (anti-construct state agreement)
dobr-i, ***vesel-i*** *čovek*
good-ATTR:M cheerful-ATTR:M man(M)
'the GOOD, CHEERFUL person'

Anti-construct state agreement marking ("long form agreement") in Serbo-Croatian is sometimes described as a definite marker on the adjective (e.g., by **kordic1997**). However, the short-form adjective can also be used in a noun phrase marked as definite, for instance by a demonstrative pronoun (**??**). And the "long form" adjective can also be used in a noun phrase marked as indefinite, for instance by the indefinite article (**??**).

(119) Serbian (**marusic-etal2007**)

a. Definite noun phrase with "pure" head-driven agreement
ovaj ***dobar***, ***veseo*** *čovek*
DEM:M good:M cheerful:M person(M)
'this good, cheerful man'

b. Indefinite noun phrase with anti-construct state agreement
Treba mi ***jedan*** ***crven-i*** *kaput.*
need.3SG 1SG.DAT INDEF:M red-ATTR:M coat(M)
(in a store with red coats on display)
'I need a RED coat (viz. one of those red coats).'

The examples with "short form" adjectives in definite contexts and "long form" adjectives in indefinite contexts provides the best evidence against the analyses of the two different adjective agreement suffixes as markers of the category SPECIES of the head noun.

Rather than as a definite marker, the long-form adjective agreement suffixes in Serbian are best analyzed as anti-construct state agreement markers used in

special contrastive focus constructions.[36]

Anti-construct state agreement in Slovenian In theory, Slovenian (aka Slovene) is identical to Serbo-Croatian in exhibiting head-driven agreement marking and anti-construct state agreement marking as two separate devices for adjective attribution.

(120) Slovenian (**priestly1993**)

a. "Short form" adjective (head-driven agreement)

i. ***nȍv*** *pə̑s*
new:NOM.M.SG dog(M)
'new dog'

ii. *en* ***nȍv*** *pə̑s*
INDEF:M.SG new:NOM.M.SG dog(M)
'a new dog'

b. "Long form" adjective (anti-construct state agreement)

i. ***nóvi*** *pə̑s*
new:ATTR:NOM.M.SG dog(M)
'NEW dog'

ii. *ta* ***nóvi*** *pə̑s*
ATTR new:ATTR:NOM.M.SG dog(M)
'the NEW dog'

Note, however, that the use of morphologically differentiated adjectives for head-driven agreement versus anti-construct state agreement in Slovenian is very restricted and is found more or less only with masculine adjectives in nominative singular (**priestly1993**).

Similar to Serbo-Croatian, anti-construct state agreement marking in Slovenian is sometimes described as a definite marker on the adjective (e.g., by **priestly1993**). Semantic definiteness in Slovenian, however, is not marked obligatorily (cf. example **??**). Furthermore, the analysis of the anti-construct state agreement as a definite marker can be rejected completely because examples are found in which this marker also occurs in overtly marked indefinite noun phrases.

[36] Note even that school grammars of Serbian sometimes explain the rules for the use of the two adjective declensions with the help of the the questions "what sort?" (requires the "short form") and "which one?" (requires the "long form") (**browne1993**).

(121) Slovenian (**marusic-etal2007**)
rabi *mi* ***en*** ***rdeči*** *plašč*
need.3SG 1SG.DAT INDEF:M red:ATTR:M coat(M)
(in a store with red coats on display)
'I need a RED coat (viz. one of those red coats).'[37]

Anti-construct agreement marking are thus analyzed as attribution marking device with the additional content of contrastive focus rather than as a detached definite marker.

Attributive nominalization + head-driven agreement in Slovenian Besides head-driven agreement and anti-construct state agreement, adjectives in (colloquial) Slovenian can also be marked by means of an attributive article.

(122) Slovenian (**marusic-etal2007**)
a. Indefinite noun phrase
Lihkar *je* *mim prdirkal en* ***ta*** ***hiter*** *avto.*
just_now AUX by sped INDEF:N ATTR fast:n car(N)
'Some FAST car has just sped by (viz. one of the fast type of cars has just sped by).'
b. Definite noun phrase
ta ***ta*** ***zelen*** ***ta*** ***debel*** *svinčnik*
DEM ATTR green:M ATTR thick:M pencil
'this GREEN, THICK pencil'

The attributive article *ta* in Slovenian is homophonous with the demonstrative determiner (from which it originates historically), but (**??**) with the double use of *ta* on stacked adjectives and after the determiner clearly shows that these markers serve two different functions: whereas *ta* DEM is a determiner marking the noun phrase for special local deictic species *ta* ATTR is an attributive marker (i.e., a true article) attaching to the adnominal adjective. Attribution of the adjective in contrastive focus in (**??**) is marked by means of attributive nominalization (in combination with head-driven agreement).

According to **marusic-etal2007**; **marusic-etal2007b** the article *ta* gives the adjective a classifying reading and the construction *ta*+A:ATTR can be compared to a "reduced relative clause", hence a syntactic complement to the noun.

[37] Cf. the similar construction with concatenative anti-construct state agreement marking in Serbian in (**??**).

7.20 Basque

Basque is a language isolate spoken in the Basque country in northeastern Spain and in adjacent parts of France in southwestern Europe.

Juxtaposition in Basque Attributive adjectives are juxtaposed to the right of the noun they modify.

(123) Basque (**saltarelli1988**)
gona ***gorri estu****-ak*
skirt red tight-DEF.PL.ABS
'the tight red skirts'

Note that the features SPECIES, NUMBER, and CASE in (**??**) are not assigned to the adjective through agreement. The corresponding portmanteau suffixes marking the values of these morphological features always attach to right edge of the phrase in Basque. Consequently, they always attaches to the attributive adjective if one is present **hualde-etal2003**

8 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 within both the whole area and each genealogical unit.

8.1 Attested attribution marking devices

Altogether 13 (simple and combined) types of adjective attribution marking devices are attested in the languages of northern Eurasia:

1. Anti-construct state
   as in Kildin Saami
2. Anti-construct state + head-driven agreement ("double agreement")
   as in Swedish
3. Anti-construct state + construct state ("double-construct state")
   as in Northern Saami
4. Anti-construct state agreement
   as in Russian
5. Appositional head-driven agreement
   as in Georgian
6. Attributive article
   as in Yiddish
7. Attributive article + head-driven agreement ("double agreement")
   as in Albanian

8. Attributive nominalization
   as in Udmurt
9. Construct state
   as in Northern Kurdish
10. Incorporation
    as in Chukchi
11. Juxtaposition
    as in Komi-Zyrian
12. Head-driven agreement
    as in Finnish
13. Modifier-headed possessor agreement
    as in Oroch

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

The Indo-European family has the largest absolute number of attested adjective attribution marking devices (nine). It is 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 combined with head-driven agreement is also very rare. This type occurs as the primary device only in the Albanian languages (Indo-European), but it is also attested as a secondary or tertiary device in a few other languages. Head-marking construct state is also relatively uncommon in the northern Eurasian area as it is attested only in Iranian languages (Indo-European).

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

8.2 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 juxta-

position 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 Indo-European family. Even though the attested deviation from the prototype is much higher in Indo-European than in Mongolic, Turkic and Uralic, head-driven agreement marking can be shown to occur prototypically in most Indo-European genera.

For the Abkhaz-Adyghe, 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, Yukaghir, 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 representing a strikingly high diversity in regard to the attested absolute number of adjective attribution marking devices are Indo-European, Nakh-Daghestanian, Uralic and Tungusic. A strikingly high degree of uniformity is found in Mongolic and Turkic.

8.3 Diachronic implications of uniformity and diversity inside and across genera

Measuring the degree of diversity (or uniformity) from a synchronic point of view may help identify diachronic processes. A very high degree of diversity inside a given taxon 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 uniformity inside a given taxon indicates the inheritance of original types without significant innovations.

A taxon is defined as a group of related languages which go back to a common reconstructed (or documented) language, i.e., a subbranch of a language family or, ultimately, the proto-form of a whole language family. 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 sam-

ple in the Appendix). 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-Baltic/Slavic daughter languages of Proto-Indo-European inherited the original Indo-European 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-Mongolic, 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 **??** illustrates 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 Indo-European 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) is counted.

The second column in Table **??** ("Languages (abs.)") gives the number of coded languages from each taxon. 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 taxon by the number of types attested for the higher branch:

$$\text{Diversity}_{\text{taxon}} = \frac{\text{Types}_{\text{taxon}}}{\text{Types}_{\text{family}}}.$$

For instance, West Saamic has a ratio of 1.00 because it exhibits all four types attested in the whole Saamic branch. The Saamic branch as such has a ratio of 1.25 because four types are found in Saamic compared to five types attested for the whole Uralic family. Similarly, South Slavic also has a ratio of 1.00 because it exhibits all three 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 three types are attested in this branch out of nine types for the whole Indo-European family.

The last ratio figures ("Ratio (lgs.)") result from dividing the overall number of

languages by the number of attested types in the given taxon:

$$\text{Diversity}_{\text{languages}} = \frac{\text{Languages}_{\text{taxon}}}{\text{Types}_{\text{taxon}}}.$$

For instance, five West Saamic languages are coded for four different types, resulting in a ratio of 1.25. For the whole Saamic branch altogether nine languages are coded for five types, resulting in a somewhat higher ratio figure of 1.80. South Slavic has the ratio of 1.33 because the four South Slavic languages are coded for three types; Slavic, however, has 4.33 because 13 Slavic languages are coded for only three different types.

1	2	3	4	5	6
Family Main branch	Languages (abs.)	Types (abs.)	Ratio (gen.)	Ratio (lgs.)	Diversity value
Abkhaz-Adyghe	*4*	*2*	–	*2.00*	low
Abkhaz	2	1	2.00	2.00	–
Circassian	2	1	2.00	2.00	–
Chukotkan	*3*	*2*	–	*1.50*	–
Chukchi	1	2	1.00	0.50	–
Koryak-Alutor	2	2	1.00	1.00	–
Indo-European	*65*	*9*	–	*7.22*	low
Albanian	2	2	4.50	1.00	–
Armenian	1	2	4.50	0.50	–
Baltic	2	2	4.50	1.00	–
Celtic	6	2	4.50	3.00	low
Brittonic	3	1	2.00	3.00	–
Gaelic	3	2	1.00	1.50	–
Germanic	14	5	1.29	2.00	high
N-Germanic	6	4	1.25	2.25	mid
W-Germanic	8	3	1.67	2.67	mid
Hellenic	1	2	4.50	0.50	–
Indo-Iranian	14	7	1.23	2.00	high
Indo-Aryan	6	3	2.33	2.00	mid
Iranian	8	6	1.67	1.33	high
Romance	10	2	4.50	5.00	low
E-Romance	1	2	1.00	0.50	–
Italo-W-Romance	7	1	2.00	7.00	low

Family Main branch	Languages (abs.)	Types (abs.)	Ratio (gen.)	Ratio (lgs.)	Diversity value
S-Romance	2	1	2.00	2.00	–
Slavic	13	3	3.00	4.33	low
E-Slavic	3	2	1.50	1.50	–
S-Slavic	4	3	1.00	1.33	mid
W-Slavic	6	1	3.00	6.00	very low
Kartvelian	*4*	*3*	–	*1.33*	mid
Georgian	2	3	1.00	0.67	–
Svan	1	2	2.00	0.50	–
Zan	2	2	2.00	1.00	–
Mongolic	*6*	*1*	–	*6.00*	very low
Dagur	1	1	1.00	1.00	–
Moghol	1	1	1.00	1.00	–
Mongolian	5	1	1.00	5.00	very low
Nakh-Daghestanian	*28*	*5*	–	*5.60*	low
Daghestanian	25	5	1.00	5.00	mid
Avar-Andi-Tsezic	13	4	1.25	3.25	mid
Dargwa	1	2	2.50	0.50	–
Lak	1	2	2.50	0.50	–
Lezgic	10	4	1.25	2.25	mid
Nakh	3	3	1.67	1.00	–
Bats	1	2	1.50	0.50	–
Chechen-Ingush	2	2	1.50	1.00	–
Tungusic	*10*	*4*	–	*2.25*	mid
Amur Tungusic	5	4	1.00	1.25	high
Nanay-Ulcha-Orok	3	3	1.33	1.00	–
Oroch-Udege	2	3	1.33	0.67	–
Manchu	1	1	4.00	1.00	–
N-Tungusic	4	4	1.00	1.00	high
Turkic	*22*	*2*	–	*11.00*	very low
Bulgar	1	2	1.00	0.50	–
Common Turkic	21	2	1.00	10.50	low
Altay	2	1	2.00	2.00	–

Family Main branch	Languages (abs.)	Types (abs.)	Ratio (gen.)	Ratio (lgs.)	Diversity value
Karluk	2	2	1.00	1.00	–
Kipchak	8	1	2.00	8.00	very low
Lena	2	1	2.00	2.00	–
Oguz	4	2	1.00	2.00	low
Yenisey	2	1	2.00	2.00	–
Uralic	*32*	*5*	–	*6.40*	low
Finnic	7	1	5.00	7.00	low
Hungarian	1	1	4.00	1.00	–
Khanty	1	1	4.00	1.00	–
Mansi	1	1	4.00	1.00	–
Mari	2	2	2.00	1.00	–
Mordvin	2	1	4.00	2.00	–
Permic	3	3	1.33	1.00	–
Saamic	9	4	1.25	1.80	high
E-Saamic	4	3	1.33	1.33	high
W-Saamic	5	4	1.00	1.25	high
Samoyedic	4	1	5.00	4.00	low
Enets	1	1	2.00	1.00	–
Nenets	1	1	2.00	1.00	–
Nganasan	1	1	2.00	1.00	–
Selkup	1	1	2.00	1.00	–
Yukaghir	*2*	*2*	–	*1.00*	–

Table 8.1: Number and ratio of attested types per genealogical unit: absolute number of types (column 3), ratio against the generally attested number of types in the respective higher branch or family (column 4, higher numbers mean less diversity), ratio against the number of coded languages (column 5, higher numbers mean less diversity) and a diversity value tested for statistical significance (column 6, only for genera with more than three languages).

LATEX: caption über der Tabelle?

The absolute number of types shows directly which families or branches in-

side families exhibit more types than other comparable genera. The first ratio in column 4 (against the number of types in the taxon) indicates where the more diverse or the more uniform branches are located inside a primary taxon (i.e., inside a family or a higher branch). These ratio figures can be used for a comparison of languages inside families or between comparable genera across families because the figures result from dividing the absolute number of attested adjective attribution marking devices in a given family by the number of devices attested in a given subfamily (i.e., branch or subbranch). East Saamic (Uralic) with a ratio of 1.33, for instance, seems just as diverse as South Slavic (Indo-European). The proto-stages of both genera have comparable time depth (approximately 1000 AD), both genera have four members and they both exhibit three attested types of adjective attribution marking devices. The number of three attested types in the two branches can then be checked against the overall number of types attested in the respective families: four types are attested in Uralic, nine types are attested in Indo-European. As compared to Uralic, the Saamic branch with a ratio of 2.25 is thus much more diverse (exhibiting almost all types attested for the whole family) than South Slavic languages within Indo-European with a ratio of 4.33 (exhibiting less than half of the generally attested types in the whole family).

The second ratio in column 5 (against the number of coded languages) relativizes the first two figures statistically. It seems much more likely that a higher number of coded languages results in a higher number of detected devices. The second ratio can thus serve to test the degree of diversity (in column 3 and 4) for statistical significance.

The simple statistics presented in Table **??** can perhaps illustrate the degree of diversity, at least in those cases where the two ratio figures (against the number of coded types and the number of coded languages) and the degree of diversity in absolute numbers coincide to a certain degree. The significant values in column 6 ("Diversity") are labeled impressionistically as *very low, low, mid-low, mid-high, high.* A hyphen marks those cases where a significant value cannot be found because the taxon in question has too few members (less than four). Note that a value *very high* is not found. This classification, however, does not mark diversity in absolute terms but the deviation from the average value of the whole sample. The Turkic family, for instance, can be shown to have a very low diversity and several of its branches clearly have a low diversity level as well. For the Mongolic family, a very low value has been calculated. Whereas a low value has even been calculated for the whole Nakh-Daghestanian family, the Daghestanian branch as well as two of its subbranches have a relatively high diversity value. The same

is true for Uralic, which has a low diversity value as a family and in several of its branches. One Uralic branch, Saamic, has a high value. Tungusic has a middle diversity value but two of its branches are clearly more highly diverse. For Indo-European,, a significant value has not been found, although inside Indo-European, high values are calculated for Indo-Iranian and Germanic.

Thus, the general picture partly coincides with what is known about areal distribution and spread of other linguistic features (**nichols1992**): less diversity (higher numbers) is found in the inner parts of North Asia (Mongolic, Turkic), whereas languages in the northern Eurasian periphery, especially in south-easternmost Europe (Caucasus) but also in north-easternmost Europe (Circum-Baltic) and in north-easternmost Asia (Pacific Rim), exhibit a higher degree of diversity (lower numbers) with respect to the morpho-syntax of adjective attribution.

Even though the figures in Table **??** summarize exclusively synchronic findings and the applied statistics is rather impressionistic, it stands to reason that they reflect historical developments (i.e., language changes) in certain parts of the area. Note that the underlying sample is not balanced and thus perhaps not easily applicable for statistical analyses. However, this is an exploratory study; detailed statistical investigations are left for future research.

The massive innovations in several neighboring genera or in larger geographic sub-areas attested synchronically may even point to contact-induced changes in areal hotbeds of innovation. In Part IV (Diachrony), some light will be shed on diachronic variation and on the evolution of highly diverse adjective attribution marking inside language families of northern Eurasia.

Part IV

Diachrony

9 The evolution of attribution marking in northern Eurasian languages

Attribution marking devices were typologized in Part II (Typology) and their geographic distribution across the genealogical entities of northern Eurasia was presented in Part III (Synchrony). The present, diachronic part focuses on linguistic changes which led to the emergence of the attested synchronic diversity within the northern Eurasian area.

Not all attested changes are investigated in equal depth in each genealogical unit. Special focus lies on the grammaticalization of attributive markers from attributive nominalizers in the Saamic and Finnic branches of Uralic as well as in the Baltic, Slavic and Germanic branches of Indo-European. Different types of adjective attribution marking have been grammaticalized from attributive nominalizers in different languages of the area and during different periods of time. Up to know, these diachronic patterns have not been systematically investigated from a cross-linguistic perspective.

The parallel evolution of attributive nominalizers and other adjective attribution marking devices is interesting not only from a general typological perspective. The linguistic interference zone between Uralic and Indo-European in northeastern Europe exhibits a relatively high degree of diversity from a synchronic point of view (see §??). Consequently, it appears that the synchronically and diachronically attested developments have to be described in areal linguistic terms and provide further evidence for establishing a Northern European *Sprachbund*.

9.1 The emergence of attributive nominalizers

Attributive nominalization as a special subtype of dependent marking attributive state (see §??) is not synchronically attested as a default licenser of the attributive connection of adjectives in any language of northern Eurasia. However, in several languages of the area, attributive constructions with nominalizers constitute a special type of noun phrases characterized earlier as attributive apposition. A typical example is Udmurt (Uralic) where an adjectival attribute equipped with

an article is marked for contrastive focus (see §??).

The only two Northern Eurasian languages exhibiting attributive nominalization as a default attribution marking device synchronically are Albanian proper and Arvanitika from the Albanian branch (Indo-European). The marker, however, is used only in a circumfixed construction together with the inherited head-driven agreement.

Attributive nominalizers are also documented in historical stages of several Indo-European branches, such as Baltic, Slavic and Germanic. But even here, these markers are not the default devices. Instead, attributive articles compete with other attributive markers and are restricted to emphatically marked noun phrases. In several of these Indo-European languages, however, the articles have evolved into new default types of attribution marking. A prototypical example of attribution marking originating from an attributive article is anti-construct state agreement marking in Russian (see §??). In other languages, the former attributive article is still traceable as secondary type of attribution marking, as in the modern Baltic languages. Here, the attributive article also evolved into an anti-construct state agreement marker but it is still restricted to a semantically defined subset of noun phrases (see §??).

The synchrony and diachrony of attributive articles have also been dealt with in a cross-linguistic investigation of grammaticalized adnominal D(eictic) elements by **himmelmann1997** Himmelmann assumes that attributive articles ("linking articles" in his terminology) originally occurred in appositional nominal expressions. These "linking constructions" are characterized as complex noun phrases in which the attribute occurs as a syntactically independent nominal expression. The "linking article" (i.e., *attributive article* in terms of the present typology) serves as a nominalizer and licenses the attribute as a syntagma of its own **himmelmann1997**

The diachronic data from several Indo-European, Uralic and Turkic languages presented in the following sections support Himmelmann's conclusions about a common source of attributive marking originating from pronouns or other deictic elements used as attributive nominalizers.

9.1.1 Attributive nominalizers in Uralic and Turkic

Juxtaposition has been the prototype of adjective attributive marking in all Uralic and Turkic languages since the proto-stages of these languages (cf. **decsy1990** for Uralic and **decsy1998** for Turkic). As the result of a secondary development, however, in some branches of Uralic and Turkic, an attributive nominalizer grammaticalized. Synchronically, it occurs as minor attribution marking device in

specially marked noun phrase types in several languages of these two families.

In the Saamic and Finnic branches of Uralic, juxtaposition has been replaced completely by new adjective attribution marking devices. In Proto-Saamic the prototypical attributive connector of adjectives was probably anti-construct state marking. A comparison of synchronic evidence across modern Saamic languages makes this reconstruction very likely (**riesler2006b**). However, the modern Saamic languages show a strong tendency to abandon the anti-construct state marker and re-introduce the morphologically unmarked adjective attribution marking device juxtaposition. In Proto-Finnic, the original Uralic type has also been lost and is now replaced by head-driven agreement marking of attributive adjectives. In §**??** and **??**, the emergence of agreement in Finnic and anti-construct state marking in Saamic will be explored and describes as a possible result of the grammaticalization of attributive nominalizers.

Since the emergence of attributive nominalizers in Udmurt (and other modern Uralic languages) probably reflect structurally similar stages of development as those assumed for Proto-Saamic and Proto-Finnic, the Udmurt case will be described in-depth in the following sections.

9.1.1.1 The contrastive focus marker in Udmurt

Synchronic data from Udmurt illustrates the emergence of an attributive article and might even indicate how this attribution marker has been generalized as an anti-construct state marker.

The use of the 3rd person possessive suffix as a contrastive focus marker in Udmurt was exemplified in §**??** on the synchrony of attribution marking in Permic. In the following sections, the etymological source and the evolution of this contrastive focus construction will be illustrated with the help of further examples.

As in several other Uralic languages, the possessive suffix 3rd person singular in Udmurt is often used as a definite-like marker. Grammatical descriptions of Udmurt use different terms to define the function of this formative, for example as "determinative" (**kelmakov-etal1999**), "contrastive-deictic" (**alatyrev1970**), "anaphorical-emphasizing" (**kiekbaev1965**), or simply "definite" (**winkler2001**). The suffix is characterized in the following as "quasi-definite" since Udmurt (as most other Uralic languages) has no morphologized feature SPECIES. The use of the marker is obviously determined by the referential status of the noun phrase, but it does not occur obligatorily in definite noun phrases. Since the rules for definiteness marking are not the subject of the present investigation, the formative in definite-like constructions will simply be referred to as *determinative suffix*, which is also consistent with some of the grammatical descriptions mentioned

above (e.g., **kelmakov-etal1999**).

Besides its function as a possessive marker, the 3rd person singular possessive suffix occurs not only in quasi-definite noun phrases but is even used as an (attributive) nominalizer and as a marker of contrastive focus on adjectives. From a synchronic point of view, the functions of POSS:3SG in the different non-possessive uses are probably better analyzed as belonging to different grammatical categories. Consequently, different glosses (such as POSS, DEF, NMLZ, CONTR) should be applied. In order to illustrate the similar historical source of the synchronically differentiated grammatical meanings, however, one and the same gloss (i.e., POSS:3SG) is used in the following examples.

(1) Possessive and non-possessive functions of (historical) POSS:3SG

a. Possessive marking

i. *gurt-**ėz***
house-POSS:3SG
'her/his/its house'

ii. *gurt-jos-a-**z***
house-PL-ILL-POSS:3SG
'into her/his/its houses'

b. "Determinative" marking

i. *gurt-**ėz***
house-POSS:3SG
'this house'

ii. *gurt-jos-a-**z***
house-PL-ILL-POSS:3SG
'into these houses'

c. Attributive nominalization

i. Demonstrative
*ta-**iz*** */ so-**iz***
DEM:PROX-POSS:3SG DEM:DIST-POSS:3SG
'this one over here' / 'that one over there'

ii. Possessor noun phrase
*Ivan-len-**ėz***
Ivan-GEN-POSS:3SG
'the one of Ivan (Ivan's)'

iii. Adjective
*badʒ́ym-**ėz***
big-POSS:3SG
'the big one'

d. Contrastive focus marking
i. *badʒ́ym-**ėz** gurt*
big-POSS:3SG house
'a/the LARGE house'
ii. *badʒ́ym-jos-a-**z** gurt-jos-y*
big-PL-ILL-POSS:3SG house-PL-ILL
'into (the) LARGE houses'

The use of the suffix *-ėz* as marker of contrastive focus is obviously connected to its other non-possessive functions. The order of examples (??–??) probably reflects the functional expansion of the original possessive marker to a "determinative" marker on noun phrases and a contrastive focus marker on adjectives. The clue for understanding this development is the use of the suffix *-ėz* as an attributive nominalizer in headless noun phrases, as shown in (??). Here, the determinative suffix is used as a true attributive nominalizer to mark a demonstrative (??), a possessor noun (??) or an adjective (??) as modifiers by projecting a full (headless) noun phrase. Note however that headless adjectives, demonstratives, and noun possessor (in genitive) are not obligatorily marked by means of attributive nominalization in Udmurt. The marker is used in order to emphasize the property denoted by the attribute and to contrast it to other properties of the same set.

The emphasizing function of the determinative suffix, finally, is the link to its use as contrastive focus marker on adjectives. It seems clear that these contrastive focus constructions originate from appositional constructions of nouns with emphasized headless attributes.[1]

(2) $[_{NP}\ [_{NP'}\ _{A}big\ _{HEAD}Ø\text{-}NMLZ]\ _{N}house]]$

(3) $[_{NP}\ [_{NP'}\ _{A}big\ _{HEAD}Ø\text{-}NMLZ]_{N}house]]$

The agreement patterns in noun phrases with attributes in contrastive focus provide the best evidence for this assumption. In their default use, attributive adjectives (as well as other modifiers) do not show agreement with the head noun.

[1] The zero-morpheme (equipped with the nominalizer Ø-NMLZ) in (??) is only presented for a better illustration of the empty head position to which the (nominalized) adjective moves in this appositional noun phrase.

However, when the attribute is marked for contrastive focus (by means of the attributive nominalizer ATTR ⇐ POSS:3SG), case and number marking spread to the adjective.

(4) Juxtaposition versus anti-construct state agreement marking (i.e., in contrastive focus) (**kelmakov-etal1999**; **winkler2001**)

a. Adjective attribute

i. *badžym / badžym-**ėz** gurt*
big big-ATTR house
'large house' : 'LARGE house'

ii. *badžym / badžym-**jos-a-z** gurt-jos-y*
big big-PL-ILL-ATTR house-PL-ILL
'to (the) large houses' : 'to (the) LARGE houses'

b. Possessor noun attribute[2]

i. *Ivan-len / Ivan-len-**ėz** gurt-ėz*
Ivan-GEN Ivan-GEN-ATTR house-POSS:3SG
'Ivan's house' : 'IVAN's house'

ii. *Ivan-len / Ivan-**jos-a-z-len** gurt-jos-a-z*
Ivan-GEN Ivan-PL-ILL-ATTR-GEN house-PL-ILL-POSS:3SG
'to Ivan's houses' : 'to IVAN's houses'

c. Demonstrative attribute

i. *so / so-**iz** gurt*
DEM:DIST DEM:DIST-ATTR house
'that house' : 'THAT house'

ii. *ta / ta-**os-a-z** gurt-jos-y*
DEM:PROX DEM:PROX-PL-ILL-ATTR house-PL-ILL
'to these houses' : 'to THESE houses'

Following the intuition of the authors of grammatical descriptions of Udmurt, however, one could also analyze these constructions as true noun phrases with a syntactic structure as in (**??**) (as opposed to **??**) where the original nominalizer of the attribute in the headless noun phrase became a dependent marking attributive construct device linking the attribute in contrastive focus to the semantic head 'house' in the noun phrase.

[2] Note that the cross-referencing possessive agreement marker does not occur with a genitive construction in contrastive focus (**kelmakov-etal1999**).

(5) *? [$_{NP}$ $_{A}$big-CONTR $_{HEAD}$house]*

Even if head-driven number and case agreement is involved in attribution marking of adjectives in contrastive focus, Udmurt is better analyzed as a language exhibiting an attributive appositional construction rather than an anti-construct state agreement marking. The agreement and anti-construct state marking formatives are not fused and agreement marking occurs only indirectly as the result of the nominalization of the appositional headless adjective.

9.1.1.2 Possessive suffixes as attributive nominalizers in other Uralic and in Turkic languages

Non-possessive uses of 3rd person singular possessive suffixes similar to Udmurt are well attested in several Uralic and Turkic languages.[3] In descriptions of these languages, the marker is often characterized as "emphatic-definite" or simply "definite" (cf. **tauli1966 kunnap2004**). But obviously this is greatly oversimplified. It is especially unclear what it would mean to mark an adjectival modifier as "definite".

Besides in Udmurt, the use of the (historical) 3rd person singular possessive suffix as a marker of contrastive focus is similarly regular (though less systematically described) in the other Permic languages (cf. **serebrennikov1963**).

In the Mari languages, which belong to the Volgaic branch of Uralic, the possessive suffix is also commonly used as a determinative suffix for nouns (cf. **alhoniemi1993**). The regular use of the formative to derive a certain set of "determinative" or contrastive focused demonstratives and quantifiers in Mari (similar to the Udmurt example (**??**) on page **??**) gives at least some evidence that the Mari languages have (or had) an attributive nominalizer in contrastive focus constructions as well.[4]

(6) Eastern Mari (Uralic; **alhoniemi1993**)
   a. "Short" demonstratives (i.e., unmarked)
      i. *tide* 'this' / tudo 'that' (82)
   b. "Long" demonstratives (i.e., in contrastive focus)
      i. *tide-**že*** 'this one' / tudo-**že** 'that one' (82)

[3] In several languages, even 2rd person singular possessive occurs in the same function.

[4] The homophonous focus clitic *=že* in Eastern Mari (*təi=že kuze ilaš tüŋalat?* 'And how are you going to live?' **alhoniemi1993**) is most likely not cognate with the 3rd person singular possessive suffix but borrowed from the formally and functionally similar marker focus marker in Russian.

c. Quantifiers in contrastive focus
Tə̂nar-žə̂-m mə̂j nalam, Tə̂nar-žə̂-m tə̂j.
so.much-POSS:3SG-ACC I take, so.much-POSS:3SG-ACC you
'So much I will take, so much you.' (76)

A similar use of the (historical) 3rd person singular possessive suffix as a marker of contrastive focus in the Turkic language Chuvash has been shown in §**??**. Interestingly, the Turkic language Chuvash and the Uralic languages Eastern and Western Mari and Udmurt are among the core members of the Volga-Kama area.[5] The languages of this linguistic area show linguistic convergence on several levels of their grammars. In all Uralic and Turkic languages of that area, at least the "emphatic-definite" use of the 3rd person singular possessive suffix is attested. Thus, it cannot be ruled out that the evolving attributive nominalizer in Chuvash, Udmurt and the Mari languages has been borrowed in either direction.

The phenomenon might even reflect a much older and more widespread feature of a larger subarea of northern Eurasia including at least Tungusic. As demonstrated in the synchronic §**??** on Tungusic, similar constructions with the 3rd person singular possessive suffix also seem to regularly occur in this family. Even in other languages of the area, examples of the use of the 3rd person singular possessive suffix as an attributive nominalizer (though not on adjectives) are attested. Example (**??**) illustrates the use of the 3rd person singular possessive suffix as an attributive nominalizer of pronouns in Khalkha Mongolian.

(7) Attributive nominalization in Khalkha (Mongolic; **pavlov1985**)
a. *olan 'much' – olan-**ki** 'what is in majority; the largest part'*
b. *numaj 'much' – numajj-**i** 'what is in majority; the largest part'*

Not also that the (historical) 3rd person singular possessive suffix occurs in practically all Turkic languages in lexicalized local and temporal attributes.

(8) Attributive nominalization in Chuvash (Turkic; **benzing1963**)
a. *śul-χi*
year-LOC:POSS:3SG
'yearly, annual' (originally 'what is in a year')
b. *yal-t-i*
village-LOC-POSS:3SG
'local' (originally 'what is in a village')

[5] Other core members of the Volga-Kama Sprachbund area are the Turkic languages Tatar and Bashkir. The Uralic languages Mordvin and Komi-Permyak are considered peripheral members (**helimski2005**).

c. *kil-t-i*
home-LOC-POSS:3SG
'domestic' (originally 'what is in the home')

It remains unclear whether the evolution of attributive nominalization and contrastive focus marking of attributive adjectives occurs independently in certain branches or areal groupings across Indo-European, Uralic, Turkic and Tungusic or goes back to a general northern Eurasian areal tendency.

9.1.2 Attributive nominalizers in Indo-European

9.1.2.1 Attributive articles and the emergence of anti-construct state agreement marking in Baltic and Slavic

Russian is the only Slavic language exhibiting anti-construct state agreement marking as the default and only type of attributive connection of adjectives (*xorošij* ATTR:NOM.M.SG 'good' versus *xoroš* PRED:NOM.M.SG, see also §??). The Russian construction where attributive adjectives are obligatorily equipped with special anti-construct state agreement suffixes resembles a construction in the closely related Baltic languages. In the latter, however, the occurrence of anti-construct state agreement marking is usually described as being restricted to definite noun phrases. The competition between complex attributive agreement and "pure" agreement marking was already characteristic of Old Baltic languages (cf. Lithuanian *geràsis* versus *gẽras*, Latvian *labais* versus *labs* 'good') and Old Slavic languages (cf. Old Bulgarian *dobrъjь* versus *dobrъ* 'good'). Old Slavic and Old Baltic languages are thus similar to modern Lithuanian and modern Latvian in exhibiting two types of adjective attribution marking suffixes in different functions.

In the Slavic and Indo-European linguistic traditions, adjectives equipped with anti-construct state agreement marking are normally referred to as "long-form adjectives" (contrasted to "short-form adjectives"). Other commonly used terms for the anti-construct state agreement markers are "pronominal, complex" or "compound" agreement suffixes. Analogically, the two inflectional paradigms of long- versus short-form adjectives equipped with number, gender, and case agreement values are normally labeled in a similar way as "long-form, pronominal, complex, or compound" versus "short-form" adjective declension. Obviously, these terms describe the form or the origin of the formative rather than its function and are rather useless for a typological comparison.

Similar to the modern Baltic languages, the markers are sometimes also labeled "definite" agreement suffixes in Old Slavic. As will be shown below, the notion

of "definiteness" does not exactly cover the functionality of the marker in Old Slavic either.

The corresponding attributive constructions in modern Slavic and Baltic languages have already been dealt with in the synchronic part of this investigation (especially §§??, ??). In the present chapter, the origin and development of anti-construct state agreement marking in Baltic and Slavic along two possible grammaticalization paths (see **??** below) will be discussed. It will be argued that these constructions have arisen from attributive articles which originally marked contrastive focus of the attribute rather than from nominal relative constructions. Before dealing with the syntactic evolution of the attributive constructions in Slavic and Baltic, however, the etymology of the formative (which is similar for both scenarios) will be sketched in the following short section.

9.1.2.2 Etymology of the formative

Whereas the "pure" agreement declension (of the so-called short-forms) of adjectives continues the Proto-Indo-European default type of adjective attribution marking, the anti-construct (long-form) agreement suffixes, as in Lithuanian *geràs-is žmõgus*, Latvian *laba-is cilvēks*, or Old Bulgarian *dobrъ-jь človekъ* 'the good person', arose as a result of a phonological merger between the short-form agreement suffixes of the adjective and a pronominal stem reconstructed as Proto-Baltic/Slavic **-jĭ/jь-*.

This pronominal part of the long-form agreement suffix likely goes back to a pronominal stem reconstructed as Proto-Indo-European **i̯o-* (**wissemann1958**). The anti-construct state agreement marker in Baltic/Slavic could thus be cognate with relative markers in other Indo-European languages, such as Old Indo-Aryan *yá-h*, Old Iranian *yō*, or Ancient Greek *hós* **heinrichs1954**

An alternative etymology has been suggested by Mikkola (**mikkola1950** see also **leskien1871 leskien1919 wijk1935**). Mikkola believes that Proto-Baltic/Slavic **-jь-* was an anaphoric marker which goes back to the 3rd person singular pronoun (cf. Lithuanian *jìs*, *jõ* 3SG:GEN or Old Bulgarian *jь*, *jego* 3SG:GEN). The phonological merger of Indo-European **is* 3SG.M with **i̯os* M 'which' in Baltic/Slavic **schmidt1959** makes this explanation possible from the point of view of sound correspondence.

The terminus post quem of the innovative attribution marking in Baltic and Slavic can be determined relatively easily. Different phonological and morphological developments of the long-form agreement suffixes in Baltic and Slavic imply that the phonological merger of adjective and the formative **-jь-* took place independently in Old Slavic and Old Baltic (**koch1992**).

It is not certain whether the Baltic and Slavic branches of Indo-European go back to a common proto-form or Proto-Baltic/Slavic have to be reconstructed as independent Indo-European daughter languages. If the latter case proves to be right, the rise of anti-construct state agreement marking could be parallel, but due to contact in Proto-Baltic/Slavic (as stated, for example, by **pohl1980**). Since the reconstruction of proto-languages is not an aim of this investigation and since the developments in Baltic and Slavic are similar from a chronological, functional and (Indo-European) etymological point of view, discussing the rise of anti-construct state agreement marking in Baltic and Slavic together in the same section makes perfect sense.

9.1.2.3 Evolution of the construction

It is commonly assumed that the function of the long-form suffix on the adjective in Old Baltic and Old Slavic was to mark the noun phrase as definite. This opinion is repeated by practically all authors of comparative grammars and reference books of the Baltic/Slavic languages as well as in works dealing specifically with adjectives and noun phrase syntax of these languages (cf. **mendoza2004** with references).

Definite nouns, however, are not obligatorily modified by long-form adjectives in Old Slavic. Furthermore, nominalized (headless) adjectives are normally equipped with long-form suffixes, regardless of the referential status of the noun phrase as definite or indefinite. The analysis of the long-form adjective suffix as definite marker might thus not be as straightforward as it appears in the reference books.

mendoza2004 connects the original distribution of long- versus short-forms to contrastive focus marking, i.e., the restrictive versus non-restrictive semantics of the attribute, instead of the referential status of the modified noun. In a similar way argues **tolstoj1957** who sees the main function of the long-form adjectives likewise in setting a certain property of a referent apart from properties of the rest of similar referents.

The later re-interpretation of such "restrictive" (i.e., contrastive focus) expressions as definite and even the generalization of the original restrictive adjective marker to a marker of anaphoric reference of the modified noun seems functionally plausible. There is no indication, however, that the long-form agreement suffixes morphologized to true definite markers in the Old Slavic languages. Even in the modern stages of the South Slavic languages Slovenian and Serbo-Croatian, where remnants of the two different adjective inflections still occur, the so-called definite (long-form) declension of adjectives is semantically restricted to certain

adjectival subclasses (see §??).

Furthermore, in Bulgarian and Macedonian, which are the only modern Slavic languages exhibiting a fully morphologized category SPECIES, the corresponding definite marking does not originate from the long-form adjectives. This is true despite the fact that the long-form agreement marking in Old Bulgarian (i.e., the ancestor language of Modern Bulgarian and Modern Macedonian) is attested to have almost grammaticalized as a marker of anaphoric reference of the noun phrase.

Note also that even the morphological status of the so-called definite adjectives in the modern Baltic languages has been doubted. It has sometimes been argued that the long-form adjective in Lithuanian might convey emphasis rather than definiteness, at least in certain expressions (cf. **kramsky1972**).

Even though the suffixes marking long-form agreement in Old Baltic and Old Slavic show some functional extension to markers of anaphoric reference or even definiteness of the noun phrase, this development is secondary. The original function of the long-form agreement suffixes was to mark an adjectival attribute in an emphatic or contrastive focus construction. Consequently, the suffix **-jъ-* in Proto-Baltic/Slavic has to be analyzed as an attribution marker on the adjective rather than as a marker of definiteness of the modified noun.

Leaving the question about the further development of the anti-construct state agreement marker **-jъ-* in different Baltic and Slavic languages aside, two opposing theories about its original function and the assumed functional developments of the anti-construct state agreement marker in Baltic and Slavic will be discussed in the following sections:

- **Scenario 1**: The formative ATTR arose from a relative pronoun, hence: DEM ⇒ REL ⇒ ATTR
- **Scenario 2**: The formative ATTR arose from an attributive article, hence: DEM ⇒ NMLZ ⇒ ATTR

9.1.2.4 Scenario 1: Nominal relative constructions in Proto-Baltic/Slavic

According to the first theory, the attributive marker in Baltic and Slavic originates from a relative pronoun. This theory seems to be widely accepted since Delbrück's and Brugmann's (cf. **delbruck1893 brugmann-etal1916**) statements on the question. Their argumentation has been taken up and augmented with new data by **schmidt1959 koch1992**; **koch1999** and others. Koch argues that a reflex of the Proto-Indo-European relative pronoun **(h)i̯o-* is attested as an attributive marker of adjectival, possessive, and adverbial modifiers of nouns in

Proto-Baltic/Slavic. He describes the constructions in which these attributes occur as "nominal relative constructions" **koch1999**

The most substantial part in Koch's argumentation seems to be the similar use of cognate relative pronouns as polyfunctional markers in relative constructions as attested in Old Iranian and Old Indo-Aryan languages.

(9) Ezafe in Old Persian (Indo-European; **meillet1931** here cited after **samvelian2007b**)

a. [*kāra* [***hya*** *manā*]]
'my army' (lit. 'army which is mine')

b. [*kāsaka* [***hya*** *kapautaka*]]
'the blue stone' (lit. 'stone which is blue')

c. *vivānam jatā utā avam* [*kāram [**hya** dārayavahaus xšāyaθiyhyā*]]
'Beat Vivâna and his army which declares itself as a proponent of the king Darius.'

Koch's (**koch1992**) main arguments for the old age of the relative function of **(h)i̯o-* in Proto-Indo-European are found in attested cognate markers. In several Indo-European languages, the historical **(h)i̯o-* pronoun marks similar relative constructions as in the Old Persian examples (**??**). Koch does not disprove, however, the assumption that the relative function of the pronoun derives from the deictic-anaphorical marking by means of a demonstrative. In fact, the Old Persian examples (**??**) clearly show verb-less relative constructions linked to the head noun with an attributive article.

Furthermore, it is not certain whether the old pronoun (or article) **(h)i̯o-* was inherited into Proto-Baltic/Slavic. The pronominal stem is attested in Baltic or Slavic only as the base of some derived connectors **heinrichs1954** Even though the etymological pronoun seems to be preserved in the stem of the Old Bulgarian relative marker *jь-že*, the function of this marker is clearly yielded by the emphatic particle *-že* **heinrichs1954** The old relative pronoun seems to be completely lost in Old Baltic where different relative markers occur (as in Lithuanian *kuřs* ⇐ *kurìs*, Latvian *kuŕš* noted by **schmidt1959**).

koch1999 dates the original relative construction back to an early Pre-Proto-Baltic/Slavic age. According to him, the relative pronoun did not agree in case with the head noun in the inherited Indo-European relative construction (**??**). Such morpho-syntactic behavior would in fact be expected from a true relative pronoun. But according to Koch's reconstruction (**??**), case agreement between a head noun and a relative pronoun was already present in Proto-Baltic/Slavic.

Finally, the long-form agreement inflection arose independently as a result of the phonological merger of the adjective and the original pronoun in Old Baltic and Old Slavic (??). Most crucial in this reconstruction is the fact that the assumed original relative pronoun has obviously never marked a true relative clause construction in Proto-Baltic/Slavic.

(10) a. Nominal relative constructions in Pre-Proto-Baltic/Slavic (**koch1999**)[6]
**dråugås gīvås jås / *dråugåm gīvås jås*
friend:NOM good:NOM REL:NOM / friend:ACC good:NOM REL:NOM
N_{nom} A_{nom} REL_{nom} N_{acc} A_{nom} REL_{nom}

b. Proto-Baltic/Slavic attributive article
**dråugås gīvås-jås / *dråugåm gīvåm-jåm*
friend:NOM good-NMLZ:NOM / friend:ACC good-NMLZ:ACC
N_{nom} A_{nom}-$NMLZ_{nom}$ N_{acc} A_{acc}-$NMLZ_{acc}$

c. Old Baltic/Old Slavic anti-construct state agreement marking
**dråugås gīvå-jås / *dråugåm gīvå-jåm*
friend:NOM good-ATTR:NOM / friend:ACC good-ATTR:NOM
N_{nom} A-$ATTR_{nom}$ N_{acc} A-$ATTR_{acc}$

This assumed development presupposes the transition of original "nominal relative constructions" in Pre-Proto-Baltic/Slavic (step 1) to a construction with an attributive article (NMLZ) in Proto-Baltic/Slavic as an intermediate step (2). The anti-construct ("long-form", i.e., ATTR) agreement marking arose as a last step (3) in Old Baltic and Old Slavic.

- Stage 1 $[_{NP}\ HEAD N\ [_{ATTRIBUTE(CLAUSE)} A_{[+agr]}\ REL_{[-agr]}]]$
- Stage 2 $[_{NP}\ HEAD N\ [_{ATTRIBUTE(NP')} A_{[+agr]}\text{-}NMLZ_{[+agr]}]]$
- Stage 3 $[_{NP}\ HEAD N\ _{ATTRIBUTE(A)} A\text{-}ATTR_{[+agr]}]$

Koch's reconstruction gives no conclusive arguments for the existence of "nominal relative constructions" marked with a relative pronoun **(h)i̯o-* in Pre-Proto-Baltic/Slavic. Theoretically, the attributive nominalization construction (step 2) could be much older and be the primary one in Indo-European. The corresponding "nominal relative constructions" in Indo-Aryan and Iranian might just as well originate from attributive nominalization constructions. The Indo-European relative pronoun **(h)i̯o-* would than go back to a deictic pronoun, probably **i-* (⇒ Latin, Gothic *is* DEM) which was used as attributive article as early as in Proto-Indo-European.

[6] The example is glossed in accordance to Koch; a translation is missing in the source.

9.1.2.5 Scenario 2: Attributive nominalizing constructions in Proto-Baltic/Slavic

According to the second idea about the emergence of the long-form adjectives in Baltic/Slavic, the attributive marker was originally an article. One opponent of the "relative" theory is van Wijk, who believes

> [...] dass wir fürs Slavische vollständig auskommen ohne die Annahme relativer Pronominalformen vom idg. Stamme *i̯e/i̯o-*, und dass dasselbe für das Baltische gilt. (**wijk1935**)

Leaving open whether an attributive article or a relative pronoun constitutes the ultimate origin of the anti-construct state agreement in Pre-Proto-Baltic/Slavic, Koch's reconstruction would in fact be compatible with Wijk's "article theory". The attribute nominalizing construction with the pronominal marker **-jъ-* as attributive article in Proto-Baltic/Slavic is clearly reflected in step 2 of Koch's reconstruction (examples **??** and **??**). The final step 3 in which the attributive nominalizer becomes an anti-construct state marker is completely similar to the development assumed by **wijk1935**

The most plausible functional explanation of the grammaticalization of the pronominal marker **-jъ-* into an attributive article is formulated by Wissemann (**wissemann1958**). He argues that the original function of the anti-construct ("long-form") agreement suffixes was that of a "Gelenkspartikel" (**wissemann1958**), i.e., an *attributive article* or *attributive nominalizer* in terms of the present study. Wissemann also shows that the function as anaphoric ("quasi-definite") noun phrase marker is secondary.

Another argument in favor of the attributive nominalizing function of the Proto-Baltic/Slavic attributive article **-jъ-* can be found in its polyfunctional use with different types of attributes. Besides marking the attributive connection of (emphasized) adjectives and participles, the article also served to mark some non-adjectival (and originally non-agreeing) attributes, such as adverbial phrases and noun phrases marked with genitive.

koch1999 gives a list of lexicalized attributive expressions in which **-jъ-* occurs as an attributive marker. These examples of frozen nominalizations present evidence of the original attributive nominalizing function of the Proto-Baltic/Slavic article.

(11) a. Attribution of adverbial phrases
   i. Old Bulgarian
   utrějь 'tomorrow- (attr.)' ← (j)utrě 'morning'

ii. Old Bulgarian
vьnějь ‘outside (attr.)’ ← vьně ‘(on the) outside’
bezumajь ‘ignorant’ ← bez uma ‘without mind’

iii. Old Bulgarian
nabožijǫjь ‘pleasing to God (attr.)’ ← na božijǫ ‘pleasing to God’

b. Attribution of noun phrases in genitive (attested only in Baltic)

i. Lithuanian
diẽvojis ‘god-like (attr.)’ ← diẽvo GEN.SG ← diẽvas NOM.SG ‘God’

ii. Lithuanian
pačiūjis ‘belonging to (attr.)’ ← pačių GEN.PL ← pàts NOM.PL ‘self’

Against his own suggestion that in Baltic/Slavic anti-construct state agreement marking originates from nominal relative constructions, in other words:

- **Scenario 2:** DEM ⇒ NMLZ ⇒ ATTR

Koch’s examples provide the best arguments for the opposite assumption that attributive nominalizing constructions are the source of that marker.

9.1.2.6 Attributive nominalizers and the emergence of anti-construct state agreement marking in Germanic

As in the Baltic/Slavic languages, the emergence of attributive nominalizers in Germanic is functionally connected to the rise of definiteness marking. In Modern Baltic and some South Slavic languages, the occurrence of anti-construct state agreement marking is restricted to (semantically) definite noun phrases. This functional devision between “true” head-driven agreement and anti-construct state agreement marking was already characteristic of all Old Baltic and Old Slavic languages.

As in the Proto-Baltic/Slavic languages, a secondary inflectional paradigm of adjectives was innovated in Proto-Germanic. This so-called weak adjective declension has often been described as the first definite marking device in Germanic (e.g., by **heinrichs1954** and **ringe2006**) because its use was restricted to (semantically) definite noun phrases. Semantic definiteness, however, was never marked obligatorily in any of the Old Germanic languages. Even though demonstrative pronouns were sometimes used in semantically definite phrases, definite markers had not yet been grammaticalized in Old Germanic varieties. Examples from Old Germanic text sources show that the use of both demonstratives and “weak adjectives” in definite phrases was optional (cf. **philippi1997 heinrichs1954**).

Only the modern Germanic languages exhibit true definite markers and thus a grammaticalized feature SPECIES. But the so-called definite articles of modern Germanic languages originate from etymological sources which were different from the older anti-construct state agreement marking suffixes. Following **riesler2006a** the rise of the Germanic "weak" adjective declension is here explained as a result of attributive nominalization.

(12) "Strong" and "weak" agreement in Proto-Germanic (**ringe2006**)

a. Head-driven ("strong") agreement
*k^wik^wa-
quick:M.SG.NOM-

b. Anti-construct state ("weak") agreement
*k^wik^wa-***n**-
quick:M.SG.NOM-**NMLZ**-
'quick'

The Pre-Proto-Germanic formative marking "weak" agreement is sometimes described as an "individualizing" or "nominalizing" suffix of nominals (i.e., adjectives and, perhaps, nouns as well). These functions are reflected in (nick-) names, such as Ancient Greek *ágáthōn* 'the Good' (← *ágáthós* 'good') or Latin *Catō* 'the Shrewd' (← *catus* 'shrewd') which are also derived from nouns equipped with the cognate suffix **-n-* (**ringe2006**).[7]

Some scholars have reconstructed a pronominal stem extension **-en-/-on-* as the origin of the suffix (for example **mikkola1950** and **heinrichs1954**). Others express their doubt about the pronominal origin of this marker (for example **schmidt1959**). But even without a definitely reconstructed etymology of the formative, the construction clearly shows similarities with the attributive nominalization of adjectives in Proto-Baltic/Slavic. It thus seems relatively safe to follow Mikkola (**mikkola1950**) and Heinrichs (**heinrichs1954**) in assuming that the weak adjective declension in Germanic goes back to a construction with an attributive nominalizer.

ringe2006 finds it "reasonable to hypothesize that the *n*-stem suffix of the weak adjective paradigm was originally a definite article". But this hypothesis must be rejected because the marker was never obligatory in definite contexts. Similar to Baltic and Slavic, it seems much more plausible to assume that the article was never a true definiteness marker. It can rather be assumed that the clue for understanding the origin of the "weak" adjective declension in Germanic is the

[7] Names such as Latin *Marcus Catō, Ovidius Nasō* are interpreted as 'Marcus the cunning' and 'Ovidius the nose' (**nocentini1996**).

nominalizing function of the *article*, which originally marked an (emphatically-contrasted) adjective as an appositional attribute.

The rise of anti-construct state agreement marking of attributive adjectives in Germanic thus followed a similar grammaticalization path as in Baltic and Slavic.[8]

(13) Grammaticalization of anti-construct state agreement in Germanic

a. Stage 1

i. Agreement marking (default)
[NP Abig-AGR Nhouse]

ii. Attributive apposition (emphatic)
[NP [NP' Abig HEADØ-NMLZ] Nhouse]]

b. Stage 2

i. Agreement marking (default)
[NP Abig-AGR Nhouse]

ii. Agreement marking (emphatic)
[NP Abig-AGR:CONTR Nhouse]

c. Stage 3

i. Agreement marking (default)
[NP Abig-AGR:ATTR Nhouse]

During Stage 1 (**??**), the attributive nominalizer (i.e., the pronominal stem extension **-en-/-on-*) competed with the default adjective attribution marking device (i.e., the inherited Indo-European head-driven agreement) but was restricted only to emphatic attributive appositional constructions. This stage can be dated back to Proto-Germanic at the latest. In all Old Germanic languages, the original attributive appositional construction is reanalyzed as a true noun phrase in which the former attributive nominalizer marks an adjective in contrastive focus. The secondary attribution marking device still competed with the default adjective attribution marking device (i.e., head-driven agreement during Stage 2 (**??**). The competition between the two different adjective attribution marking devices was dissolved during Stage 3 (**??**). This stage is reflected by the modern West Germanic languages where only one type of adjective attribution marking occurs. Due to the fact that agreement inflection of adjectives in modern West

[8] The zero-morpheme (equipped with the nominalizer Ø-NMLZ) in (**??**) and following examples is only presented for a better illustration of the empty head position to which the (nominalized) adjective moves in the appositional noun phrase.

Germanic languages (except in English) only marks attributive but not predicative adjectives, this adjective attribution marking device has been characterized as anti-construct state agreement (see §??).

9.1.3 Excursus: Definite noun phrases in Germanic

In the previous section, it was shown that the grammaticalization of the feature SPECIES (definiteness) in Germanic is a relatively recent phenomenon which is not directly connected to the rise of attributive nominalization and anti-construct state agreement marking (so-called "weak" or "definite" agreement). Even though anti-construct state agreement usually occurred in semantically definite noun phrases, true definite markers evolved much later.

The etymological source of the definite markers were local-deictic (demonstrative) pronouns: Proto-Germanic **sa, *sō, *þat*, in North Germanic additionally also *en, enn, et* (**heinrichs1954**). Interestingly, the evolving definite markers from the first set of Proto-Germanic demonstratives were also first used as attribution markers of adjectives (**gamillscheg1937**; **nocentini1996**). Later, the use of the articles was extended from appositional (nominalized) adjectives to whole noun phrases (**philippi1997**). If the grammaticalization path illustrated in (??) is extended with one more stage, the evolution of definiteness marking in Germanic can be included as well. Note that the additional developments in the grammaticalization path (??) are also partly connected to adjective attribution.

(14) Grammaticalization of definiteness marking in West Germanic
   a. Stage 3
      i. Agreement marking (default)
         [$_{NP}$ $_{A}$big-AGR:ATTR $_{N}$house]
      ii. Attributive apposition (emphatic)
         [$_{NP}$ [$_{NP'}$ $_{ART}$the $_{A}$big-AGR:ATTR $_{HEAD}$Ø] $_{N}$house]
   b. Stage 4
      i. Definiteness marking
         [$_{NP}$ $_{DEF}$the $_{A}$big-AGR:ATTR $_{N}$house]

Note that an attributive apposition construction for marking emphasis occurs twice in the illustrated grammaticalization path (??). In Stage 1 (??), the attributive nominalizer is the pronominal stem extension **-en-/-on-* which becomes the anti-constract state agreement marker in the following stage (??). The second attributive nominalizer in Stage 3 (??) is the demonstrative pronoun which becomes

the definite marker in the following stage (**??**). These two attributive nominalizers have different etymological sources and attach to different positions inside the noun phrase but they are functional equivalents.

Stage 4 in example (**??**) did not fully affect North Germanic. Instead, the Old North Germanic languages (Old East and Old West Norse) grammaticalized definite markers from the demonstratives *en, enn, et* (**heinrichs1954**). These markers are the complete morpho-syntactic opposites of West Germanic: Unlike the West Germanic preposed and free form definite marker, all modern North Germanic standard languages exhibit a postposed definite noun inflection. The different morpho-syntactic realization of the general Germanic tendency towards grammaticalization of definiteness is best explained as contact-induced change due to Saamic influence in North Germanic (**kusmenko2008**).

(15) Grammaticalization of definiteness marking in Germanic

a. Stage 4

i. Definiteness marking (West Germanic)
[NP DEF the A big-AGR:ATTR N house]

ii. Definiteness marking (North Germanic)
[NP ATTR:AGR the$_{\text{agr:attr}}$ A big-AGR:ATTR N house-DEF]

Note that in North Germanic Stage 4 (**??**) the former preposed nominalizer (article) did not grammaticalize into a true definite marker like in West Germanic but into an anti-construct state agreement marker. The noun phrase structure is thus different from Stage 3 (**??**) because the attributive apposition of a the nominalized headless adjective is lost and the semantic head of the overall noun phrase is syntactically reunited with its adjectival modifier.

Synchronic data from different North Germanic varieties reflect intermediate stages in the evolution of definite noun phrase structure. This cross-linguistic variation is most likely the result of competing grammaticalization of a preposed article and a postposed definite inflection (**dahl2003**).

As with all modern West Germanic languages,[9] the Western Jutlandic dialect of Danish exhibits phrasal definite marking by means of a phonologically free and preposed definite article.

(16) W-Jutlandic[10]

a. *de korn* [DEF corn]

[9] In English, the noun phrase structure is similar in theory, with the exception of adjectives in headless noun phrases which are obligatorily nominalized: *the good **one***; see also §**??**.

[10] The examples are constructed according to **lund1932** cf. also **delsing1993** and **dahl2003**

b. *de god (et)* [DEF good:AGR (NMLZ:AGR)]
c. *de god korn* [DEF good:AGR corn]

In several of the northernmost North Germanic varieties, definiteness is also marked phrasally but by means of a phonologically bound and postposed formative. Consequently, the phrasal definite marker attaches as suffix to definite nouns and definite headless adjectives alike. Note also that adjectives are incorporated into (or compounded with) the head noun.

(17) Västerbotten Swedish[11]
a. *korn-e* [corn-DEF]
b. *god-e* [good-DEF]
c. *god-korn-e* [good-corn-DEF]

In the North Germanic languages Norwegian[12] and Swedish as well as in Faroese, the definite marker is an inflectional suffix as in the Västerbotten dialect of Swedish, i.e., phonologically bound and postposed. The formative is, however, exclusively a noun marker and does not show up on adjectives in definite headless noun phrases. The latter are not overtly marked as definite but show circumpositioned definite agreement marking by means of a preposed attributive article and definite agreement inflection.

(18) Swedish (personal knowledge)
a. *korn-et* [corn-DEF]
b. *det god-a korn-et* [NMLZ:AGR good-AGR corn-DEF]
c. *det god-a* [NMLZ:AGR good-AGR]
d. * *det korn-et*

In Danish and (colloquial) Icelandic, the definite marker has two allomorphs: an inflectional noun suffix similar to Swedish (i.e., a phonologically bound and postposed) and a definite article similar to the West Germanic languages (i.e., phonologically free and preposed). Interestingly, the allomorphy of the definite marker in Danish and Icelandic is triggered by the part-of-speech membership of the host: whereas the bound allomorph selects for nouns, the free form selects for adjectives.

(19) Danish (personal knowledge)

[11] The examples are constructed according to **astrom1893** cf. also Delsing (**delsing1993**) and **dahl2003**
[12] New- and Dano Norwegian

a. *korn-et* [corn-DEF]
b. *det god-e korn* [DEF good-AGR corn]
c. *det god-e* [DEF good-AGR]
d. * *det god-e korn-et*

Table 9.1: Paradigm of the definite marker in Danish (personal knowledge). Note that the choice whether the suffix or the free from constitute the base morpheme or the allomorph seems arbitrary.

	UTR	N	PL
DEF	-en [den]	-et [det]	-Ø [de]

(20) Icelandic (personal knowledge)
a. *korn-ið* [corn-DEF]
b. *hið goð-a* [DEF good-AGR]
c. *hið goð-a korn* [DEF good-AGR corn]
d. * *hið goð-a korn-ið* [DEF good-AGR corn-DEF]

9.1.3.1 "Double definiteness" and a "buffer zone" in North Germanic

The geographic distribution of different morpho-syntactic types of definiteness marking across North Germanic reveals interesting areal patterns. The occurrence of adjective incorporation coincides with the area of the missing preposed article. Both features are characteristic of the northeastern periphery of North Germanic (**delsing1996b** cf. also **riesler2001a**; **riesler2002a**). The structural connection between adjective incorporation and the missing preposed article is obvious: the construction with the compounded (incorporated) adjective in definite noun phrases substitutes the corresponding construction with the preposed article in those dialects where a preposed article has not (yet) been developed from the former demonstrative. The northeastern North Germanic data thus reflects an early Stage 3 in the illustrated grammaticalization path (**??**).

The northeastern North Germanic dialect area constitutes the innovation center of the grammaticalization of a (suffixed) inflectional category SPECIES (definiteness). The southwestern North Germanic dialects, located geographically at

the very opposite periphery, exhibit a structurally reversed picture of northeastern North Germanic which is in its direction of evolution almost identical to the situation in West Germanic.

Dahl describes the phrasal definite markers in southwestern and northeastern North Germanic dialects as the result of structurally and geographically opposed processes of grammatical changes.

> [T]he variation we can see in the attributive constructions is the result of the competition between them about the same territory. (**dahl2003**)

The "competition" between northeastern and southwestern grammaticalization tendencies in Germanic is not restricted to definite marking. Several grammatical categories which developed as the result of common Germanic (or even Indo-European) tendencies, have grammaticalized into non-fusional (analytic) constructions in West Germanic but into concatenate (synthetic) constructions in North Germanic. Language contact with neighboring Uralic languages would offer the most plausible explanation for the structurally differentiated developments inside the Germanic branch. Consequently, Kusmenko (**kusmenko2008**) proposed a model for explaining the morphological fusion of definiteness and other North Germanic innovative categories as the result of interference features during the language shift of the assimilated Saami of Mediaeval Scandinavia.

A direct connection between language contact and the rise of adjective incorporation and the missing preposed adjective article in northeastern North Germanic varieties was also suggested by **riesler2001a**; **riesler2002a** But even if this idea cannot be proven correct the historical connection between missing preposed adjective articles, adjective incorporation and the morpho-syntactic type of definiteness marking (i.e., morphologically fused and postposed) in the northeastern North Germanic dialect area is obvious. Saamic influence (causing the morphological fusion of postposed definiteness marking) would thus at least be an indirect trigger of these areal grammaticalization phenomena in North Germanic which can be described as a "buffer zone" (**stilo2005**).[13]

[13] Stilo created the term for a similar language area between competing grammaticalization tendencies due to contact induced-changes in the Southern Caucasus. The parallel between Stilo's "buffer zone" and Dahl's (**dahl2003**) "competing" morpho-syntactic types in North Germanic languages was first mentioned to the author by Tania Kuteva (p.c.). But neither Dahl nor Kuteva drew contact linguistic implications in the North Germanic case. The idea about the North Germanic "buffer zone" as an indirect result of contact-induced changes was first mentioned by **riesler2006a**

Table 9.2: Article grammaticalization cycle in Germanic languages (adapted from **riesler2006a**).

		Proto-Germanic		Old-Germanic		Modern Germanic			
DEM1	⇒	ART1	⇒	ATTR	⇒	AGR	⇒	Ø	English, (W-Jutlandic)
DEM1	⇒	ART1	⇒	ATTR	⇒	AGR			W+N-Germanic
				DEM2	⇒	ART2	⇒	DEF1	W(+N)-Germanic
				DEM2	⇒	ART2			N-Germanic
				DEM2					Västerbotten Swedish
				DEM3	⇒	DEF2			N-Germanic

9.1.4 Attributive nominalization and the grammaticalization of anti-construct state (agreement) marking

The previous sections described how anti-construct state agreement marking arose in the Baltic, Slavic and Germanic branches of Indo-European. Structurally similar developments were also described for Udmurt from the Permic branch of Uralic, in Chuvash and other so-called Uralo-Altaic languages in §??.

The emergence of attributive nominalizers such as secondary attribution markers seem to reflect a general tendency in several branches of the Indo-European, Uralic and Turkic language families. The etymological source of the attributive nominalizer in all of these languages is either a local deictic determiner or the 3rd person possessive marker with "determinative" functions.

Synchronic data from several languages of the Lezgic (Daghestanian) branch of Nakh-Daghestanian (see §??) seem to reflect a similar grammaticalization path from deictics to attributive nominalizers. Most Lezgic languages sampled for the present study have juxtaposition as the default adjective attribution marking device. Attributive nominalization also occurs in most languages of this branch but is restricted to headless noun phrases. The attributive nominalizer is a stem augment *-tV- / -dV-* which could be connected historically to the deictic pronouns occurring with similar shapes in these languages. In Budukh, the cognate suffix *-ti* is not used as an attributive nominalizer but to emphasize "a high degree of quality", cf. *godak* 'short' : *godak-ti* 'very short' (**alekseev1994b**). In Rutul, the cognate marker *-d* is used as an anti-construct state marker on attributive adjectives as the default (**alekseev1994a**). A different but nevertheless related function of the cognate marker is attested in Archi where the suffix *-ťu* derives adjectives from nouns, adverbs and postpositions (**kibrik1994b**).

The data from Lezgic deserves further investigation, but it suggests a pattern where the dependend-marking attributive state evolves from attributive nominalization. It is also very obvious that the attributive nominalizers in Uralic and Turkic have evolved along a similar grammaticalization path as the one described for several Indo-European (and other) languages by **himmelmann1997** Important differences between Himmelman's "linking articles" and the attributive nominalizers described here, however, are (1) the origin of the Uralic and Turkic nominalizers from person-deictic rather than from local-deictic markers and (2) the inflectional use of the markers in Uralic and Turkic as compared to their original adnominal use in Indo-European.

The data from Uralic and Turkic is especially interesting, since it contradicts Himmelmann's (**himmelmann1997**) assumption that a functional convergence between attributive nominalizers with a person-deictic or a local-deictic etymo-

logical source is unlikely to occur. Of central importance to Himmelmann's analyses is the "anamnestic" use of the deictic markers from which the articles are grammaticalized. According to Himmelmann, the use of "D(eictic) elements" in order to refer to properties the speaker believes to be well-known for her/his interlocutor is the most relevant precondition for their further grammaticalization into articles and definite markers. Whereas the "anamnestic" use is inherent in (local-deictic) demonstratives, the same is not true for (person-deictic) possessive markers. The further grammaticalization of demonstratives into functional determinative elements (like articles and definiteness markers in several Indo-European languages) is accompanied by a functional extension of an original "anamnestic" to an associative-anaphoric use of the markers. This is in contrast to the further grammaticalization of possessive markers into functional determinative elements (like attributive articles and quasi-definiteness markers in certain Uralic languages) which is accompanied by a functional extension from an original associative-anaphoric to "anamnestic" use.

> D-Elemente breiten sich von pragmatisch-definiten Kontexten auf semantisch-definite aus, während Possessivpronomina sich umgekehrt von einem semantisch-definiten Kontext auf einen bzw. mehrere pragmatisch-definite Kontexte ausdehnen. (**himmelmann1997**)

Himmelmann's thesis regarding the opposite functional extension of person-deictics might still be valid and compatible with the Uralic and Turkic data. In those Uralic and Turkic languages with attested attributive nominalization, the definite (or quasi-definite) function of the possessive marker is also always present. It can therefore be assumed that the definite (or quasi-definite) use of the marker obligatorily occurs as an intermediate step during the grammaticalization of possessive markers to attributive nominalizers.

- **Person-deictic source** (Uralic, Turkic)
  POSS ⇒ DEF ⇒ NMLZ

In the Indo-European languages with attributive articles such an intermediate step is probably not necessary.

- **Local-deictic source** (Indo-European)
  DEM (⇒ DEF) ⇒ NMLZ

In fact, in the West Germanic and South Slavic languages, definite markers evolve from attributive nominalizers but not vice versa.

- **Local-deictic source (West Germanic, South Slavic)**
  DEM ⇒ NMLZ (⇒ DEF)

This observation will be taken up again. If the tentative observation on the languages with "grammaticalized person-deictic elements" (i.e., possessive markers as attributive nominalizers) proves right it would imply the following implicational universal:

(21) **Implicational universal**
Possessive markers develop into attributive nominalizers only in languages in which similar possessive markers are already used as markers of (quasi-) definiteness.

Whereas the etymology and the evolution of attribution markers in Indo-European has been described (more or less systematically) by different authors, much less has been written about the emergence of attribution markers in different Uralic and Turkic languages. The emergence of anti-construct state marking in Saamic, which has not been described at all, appears to be especially interesting in this respect.

9.2 The emergence of anti-construct state marking in Saamic

In §**??**, it was shown that the contrastive focus marker in Udmurt most likely evolved from an attributive article. **riesler2006b** suggested the idea that a similar construction was the ultimate source of anti-construct state marking in the languages of the relatively closely related Saamic branch of Uralic. Since this theory about the rise of attribution marking in Saamic is based on a controversial idea, it calls for a relatively detailed discussion which will be presented in the following sections.

In §**??**, it was shown that adjectives in all Saamic languages are normally marked morpho-syntactically by means of differentiated attributive and predicative state markers. Even though the system of attributive and predicative marking is highly irregular in the Saamic languages, it can be shown that the attributive forms of adjectives are prototypically marked with a suffix (Northern Saami) *-s*. This suffix constitutes a prototypical example of an anti-construct state marker, i.e., a dependent marking attributive morpheme.

The origin of anti-construct state marking in Saamic is controversial. The suffix *-s* is definitely not inherited from Proto-Uralic. It is probably not borrowed

from any of the known current or historical contact languages of Saamic either. Considering this as well as the fact that Saamic is a rare instance among the Northern-Eurasian languages in exhibiting anti-construct state marking on adjectives, relatively little attention has been paid to explaining its origin.

9.2.1 State of research

The different proposed theories which explain the origin of the anti-construct state marker on adjectives in Saamic can be subsumed as follows:

1. Grammatical borrowing from Indo-European
2. Functional extension of an adjective derivational marker
3. Grammaticalization from an attributive nominalizer

The idea about a grammaticalization from an attributive nominalizer presented by Nielsen (**nielsen1933**) and Atányi (**atanyi1942**; **atanyi1943**) is the only contribution to the subject spelled out in certain detail. Interestingly enough, the idea has been rejected as "hardly convincing" (my translation) in a one-sentence-statement in Korhonen's (**korhonen-m1981**) historical grammar of Saami. Korhonen's judgement that the origin of the attributive suffix in Saamic is still unclear **korhonen-m1981** seems to reflect the state of research up to today. Neither of the three hypotheses mentioned above has been discussed seriously in Saami or Uralic historical linguistics.[14]. All proposed hypothesis will be evaluated.

9.2.1.1 Loan adjectives

Trond Trosterud (p.c.) has suggested that the attributive suffix in Saamic origins from an ending typical of Proto-Germanic loan adjectives in Saami. The Saamic suffix *-s* would then reflex the (pre-rhotacism) form of the Proto-Germanic case suffix *-R* for masculine nominative singular which was adopted into Proto-Saamic together with loan adjectives. According to this hypothesis (which is not discussed in any publication so far) the adjective ending *-s* occurred originally on Germanic loan adjectives but was later generalized and used with inherited adjectives as well. In fact, a considerable number of Germanic loan adjectives with the corresponding ending *-s* < Proto-North Germanic *-R* M.NOM.SG is attested in Saamic, for instance:

[14] An exception is a short article by **sarv-m2001** who presents the different ideas but does not come to conclusive results.

- Northern Saami *smáves* 'small' ⇐ Proto-Saamic **smāv̀ę* < Proto-North Germanic; cf. Old Norse *smalr* M (or a more recent North Germanic borrowing; cf. Swedish *små*; **sammallahti1998b**)
- Lule Saami *riukas* 'far-reaching' < Proto-North Germanic, cf. Old Norse *drùgr*, Norwegian *drjug* **qvigstad1893**
- Lule Saami *lines* 'soft, yielding, mild' < Proto-North Germanic, cf. Old Norse *linr*, Norwegian *lin* **qvigstad1893**
- Northern Saami *luovȯs* ~ *luovus* 'loose, not tied' ⇐ Proto-Saamic **luovōs* ~ **luovęs* < Proto-North Germanic **lauss* M (where the suffix *-R* is assimilated into /s/) **sammallahti1998b**
- Northern Saami *suohtas* 'fun, nice' ⇐ Proto-Saamic **suohtęs* < Proto-Germanic **swōtu-* **sammallahti1998b** cf. Old Norse **søtr* M
- Northern Saami *viiddis* 'wide, extensive' ⇐ Proto-Saamic **vijðēs* < Proto-North Germanic **lehtiranta1989** cf. Old Norse *víðr* M

The sound change of Proto-Germanic **-z* ⇒ Proto-North Germanic *-R* (⇒ Common North Germanic *-r*) took place around 500 AD. The hypothesis of the loan origin of the Saamic attributive suffix presupposes that the corresponding suffix in Germanic had a sound value [-z] (or ?[-s]). The exact sound value of *-R*, however, is not at all certain. What is commonly accepted is that the sound was phonologically distinguished from /r/ (**skold1954**).

From the point of view of its etymology, the adjective ending *-s* is identical to the ending *-s* of some borrowed Proto-Germanic nouns, such as Proto-Saamic **vālās*, cf. Northern Saami *fàlis* 'whale' < Proto-North Germanic, cf. Old Norse *hvalr*, cf. Norwegian *hval* (**qvigstad1893 lehtiranta1989**) or Proto-Saamic **kāllēs*, cf. Northern Saami *gállis* 'old man' < Proto-Germanic **karilaz* M **lehtiranta1989** The ending *-s* in bisyllabic nominals is thus an indicator that the word in question might belong to the layer of Proto-North Germanic borrowings in Saamic.

In many instances of Germanic loan adjectives the ending *-s*, however, marks only the predicative and not the attributive form, consider (from the list above):

- Northern Saami *smávva* [small.ATTR] ← *smáves* 'small'
- Lule Saami *riuka* [far-reaching.ATTR] ← *riukas* 'far-reaching'
- Lule Saami *littna* [soft.ATTR] ← *lines* 'soft'

Other loan adjectives have identical forms with the ending *-s* in both predicative and attributive function:

- Northern Saami *luovȯs* ∼ *luovus* 'loose'
- Northern Saami *suohtas* 'fun, nice'
- Northern Saami *viiddis* 'wide, extensive'

It is unclear whether the Germanic loan adjectives ending in *-s* regularly occurred in both attributive and predicative positions already in Proto-Saamic, or the ending *-s* expanded from predicative to attributive forms, or vice versa.

The relatively regular occurrence of the ending *-s* in the predicative forms suggests that the corresponding Germanic loan adjectives also ending in *-s* were originally used to denote predicates rather than attributes. This seems reasonable from the point of view of the morpho-semantics of the borrowed Germanic adjectives as well. The ending *-R* (⇐ **z*) marks masculine nominals only in the so-called strong declension and thus occurred more likely on predicative adjectives which normally denote temporary properties. Attributive adjectives in Germanic, by contrast, could be marked either by means of head-driven agreement ("strong declension") or anti-construct state agreement ("weak declension") depending on the semantic or referential status of the attribute. An adjective denoting a permanent property was normally marked with the anti-construct state agreement suffix (see §**??**).

Consequently, the Saamic ending *-s* could have been borrowed exclusively from "strong" adjectives in masculine nominative singular, the only form which had the ending *-R* (⇐ **z*) in Proto-North Germanic. It is thus doubtful that just the borrowed forms with *-s* have been generalized as attributive forms by bilingual speakers in the assumed Saamic-Germanic language contact situation.[15] It should thus be assumed that the Germanic loan etymology of certain adjectives in Saamic does not provide a clue for the origin of the attributive suffix.

Another problem in the hypothesis of the Germanic origin of the Saamic adjective ending *-s* might be the class of inherited Saamic adjectives which also have the ending *-s* when used predicatively. Consider the following examples:

- Northern Saami *báhkas* 'hot' ← *báhkka* [hot.ATTR] ⇐ Proto-Saamic **pāhkȩs* ⇐ Pre-Proto-Saamic **pakka-* 'hot; cold'; cf. Finnish *pakkanen* 'frost' (**sammallahti1998b**)

[15] There is no doubt that language contact between speakers of Proto-Saamic and Proto-North Germanic took place; cf. **kusmenko2008** It is, however, rather irrelevant to the case described here which contact scenario has to be assumed: borrowing proper or shift-induced interference in the Saamic L2 of original Germanic speakers.

- Northern Saami *garas* 'hard' ← *garra* [hard.ATTR] ⇐ Proto-Saamic **kə̑rə̑-* ⇐ Pre-Proto-Saamic **kiri-*; cf. Finnish *kireä* 'tight, tense' (**sammallahti1998b**)
- Northern Saami *ođas* 'new' ← *ođđa* [new.ATTR] ⇐ Proto-Saamic **oə̑̆ə̑-* (**sammallahti1998b**).

Since the most typical Proto-Saamic root can be reconstructed as an open bisyllabic,[16] the ending *-s* of these predicative adjectives could not have belonged to the root originally. The ending-less attributive forms in the examples above would then reflect the original adjective roots, characterized as bisyllabics with an open second syllable. According to the Proto-Saamic morpho-phonological rules, the stem consonant center exhibits the strong grade before an open second syllable, unlike the predicative forms which have a closed second syllable ending in *-s* and show the weak grade of the consonant center.

The same morpho-phonological rule applies to loan adjectives with ending-less attributive forms (like 'small' in Northern Saami: *smávva* [small:ATTR] ← *smáves*). If one adopts the idea of *-s* originally being a Germanic case suffix, the attributive forms of the loan adjectives in Saamic can only be derived from the strong-declension forms of Germanic predicative adjectives and not from attributive adjectives.

In the case of the inherited Saamic adjectives, however, it is usually assumed that the predicative ending *-s* is derivational (see also the following paragraph). This assumption presupposes the ending-less (attributive) adjective being the base form from which the predicative form is derived by means of the derivational ending *-s*.

9.2.1.2 Locative adjective derivation

According to **bergsland1946** the origin of the attributive suffix *-s* in Saamic is identical with that of the synchronically homophonous adjective derivational suffix *-s* originating from a lative case marker. Cognate formatives deriving adjectives from nouns occur in other Uralic languages, like Hungarian *erős* 'powerful, strong' (← *erő* 'power, strength'), *kékes* 'bluish' (← *kék* 'blue').

The development of local case expressions to adjectives is semantically plausible and could in principle be adopted for Saamic. Probably, the local case suffix was first used as adverbalizer of nominal stems and became a true adjectivizer at a later stage, hence:

- LATIVE CASE ⇒ ADVERBALIZER ⇒ ADJECTIVIZER

[16] Cf. the list of reconstructed Proto-Saamic lexemes in **lehtiranta1989**

The intermediate stage in the assumed development from a local case expression to an adjective is reflected in place adverbs like Northern Saami *guhkás* '(going) far' ⇐ Proto-Saamic **kuhkā-sę* **sammallahti1998b** and probably also in other adverbal derivations, like the collective numbers on *-s*, cf. Northern Saami *golmmas* 'a group of three' ← *golbma* 'three'.

Since predicative adjectives are not subject of this investigation, the the observation is sufficient that both the assumed (inherited) locative derivation and the assumed suffix borrowing are possible scenarios which do not necessarily exclude each other. As a result of these developments, a lexically defined subclass of adjectives with predicative forms on *-s* arose in Common Saamic (or earlier). The marker of this class of adjectives, the ending *-s*, is either:

- borrowed from < Proto-North Germanic *-R* M.NOM.SG
- derived (historically) from ⇐ LATIVE CASE,
- the result of merger of both developments.

The adjective class characterized by predicative forms on *-s* (which has more or less regular ending-less attributive forms) is clearly identifiable in all modern Saamic languages.

Bergsland's (**bergsland1946**) suggestion that the similar ending *-s* in the attributive forms of certain adjectives goes back to the Uralic lative case suffix as well is relevant to the present investigation. Deduced from his statement that the attributive suffix *-s* is "originally a Finno-Volgaic lative suffix" Sammallahti (**sammallahti1998b**) agrees with Bergslands explanation. Also **judakin1997** argues in this direction.

The adjective ending *-s*, which is the basis for Bergsland's and Sammallahti's argumentation, marks the predicative form of some adjectives and the attributive form of others. There are only a few adjectives which have the ending *-s* in both predicative and attributive forms. Neither Bergsland nor Sammallahti discuss the question as to whether the assumed lative derivation originally occurred: a) on predicative adjectives, b) on attributive adjectives, or c) on both forms simultaneously.

A cross-comparison of cognate forms of attributive and predicative adjectives in different Saamic languages suggests that adjectives with similar predicative and attributive forms with *-s* form a minor class which very likely arose as the result of a secondary development.

Cross-comparison can also provide evidence for separate etymologies of two homophonous predicative and attributive endings *-s*. The locative derivational

suffix can only be the source of this suffix *-s* which is homophonous on predicative and attributive adjectives in modern West Saamic languages. The original attributive adjective suffix, however, should be reconstructed as a (phonetically palatalized) suffix *[-sV$^{j}_{[+front]}$] preceding a front vowel. In the easternmost Kola Saami languages, the attributive suffix *-s'* has a palatalized coda and is clearly distinct from the non-palatalized *-s* on predicative adjectives as well as from the (cognate) lative adverbalizer *-s*.

(22) a. Adjective stem 'long (pred.)'

guhkki	Northern Saami
kuhk'	Kildin Saami

b. Adverb '(going) far'
(adverbalizer suffix (non-palatalized) ⇐ *-*s*)

guhkás	Northern Saami
kugkas	Kildin Saami

c. Attributive form 'long (attr.)'
(attributive suffix (palatalized) ⇐ *-*s'*)

guhkes	Northern Saami
kugk'es'	Kildin Saami

9.2.1.3 Attributive nominalization

A different hypothesis about the origin of the attributive forms in Saamic has been proposed by Joszéf Budenz (**budenz1870** according to **atanyi1942**; **atanyi1943**) who believed that the suffix *-s* represents the original possessive suffix 3rd person singular. Budenz does not give any evidence specifically for Saami. He simply assumes that the determinative function of the possessive suffix, a similar use of which he observed in different Uralic and Turkic languages (see §**??**), caused the development in Saami. Budenz' idea was taken up specifically for Saamic by Atányi (**atanyi1942 atanyi1943**). Atányi also refers to Nielsen (**nielsen1933** reprinted in **nielsen1945b**), who had a similar idea (probably independently of Budenz, who he does not refer to).

This hypothesis on the origin of the attributive forms in Saamic perfectly accounts for the different phonological shapes of the (historical) adjectivizer *-*s* and the attributive suffix *-s* (⇒ E-Saamic *-s'*). According to this theory, recently taken up again by **riesler2006b** the attributive suffix *-s/-s'* reflects an old 3rd person singular possessive suffix which was used as an attributive article on contrastive-emphasized adjectives.

The reconstructed Proto-Saamic forms of the possessive marker **-sē* **sammallahti1998b** versus the adjectivizer **-sę* are consistent with the synchronic findings. The different phonological form of the two suffixes (/-s^{j}/ versus /-s/) in the Kola Saami languages and the phonological merger of both suffixes (non-palatalized /-s/) in the western Saamic languages can be accounted for by a regular sound law: in the Kola Saami languages the apocope of etymologically front vowels (**i, *e*) is reflected by the palatalization of the consonant preceding the lost vowel. Apocope of non-front vowels (like **-sę*) did not affect the quality of the consonant. This sound law does not apply to the western Saamic languages which do not exhibit (phonological) palatalization and consequently consonants preceding etymologically front and back vowels are non-palatalized.

(23) a. 'guest' (not possessed)

i.	* *kuasse*	Proto-Saamic
ii.	*kuss'*	Kildin Saami
iii.	*guossi*	Northern Saami

b. 'her/his/its guest' (marked with POSS:3SG suffix)

i.	* *kuasse-sē*	Proto-Saamic
ii.	*kuss'es'*	Kildin Saami
iii.	*guossis*	Northern Saami

Beside the overall irregularity in the attributive marking in all Saamic languages (see §??), the different morpho-phonological behavior of the nominal stems which POSS:3SG and ATTR attach to appears to be an argument against this reconstruction.

(24) Strong (STR) and weak (WK) consonant grade in adjectives and nouns

a. *kugk'(WK)-es' suhk(STR)* Kildin Saami
guhke(WK)-s suohkku(STR) Northern Saami
long-ATTR stocking
'the long stocking'

b. *suhk(STR)-es' lī kuhk'(STR)* Kildin Saami
suohkku(STR)-s lea guhkki(STR) Northern Saami
stocking-POSS:3SG is long.PRED.
'her stocking is long'

c. *kugk'(WK)-es' sugk(WK)-es't* Kildin Saami
guhke(WK)-s suohku(WK)-s Northern Saami
long-ATTR stocking-LOC.SG
'in the long stocking'

A noun marked for possession is in the strong consonant grade. An adjective marked for attribution is always in the weak grade. In the example above, the strong grade of the consonant (orthographically represented as *hk* in Kildin Saami and *hkk* in Northern Saami) occurs in the nominative case of the bare or possessive marked noun (*suhk/suohkku, suhkes'/suohkkus*) as well as in the predicative form of the adjective (*kuhk'/kuhkki*). The attributive form of the adjective (*kugk'/guhkes*) and the noun stem hosting the locative suffix (*sugkes't/suohkus*) are in the strong grade.

Historically, consonant gradation was a purely phonological process where the strong consonant grade always occurred before the open final syllable of a disyllabic word. The stem consonant was phonetically shortened when the final open syllable was closed due to inflectional processes. Consonant gradation was later morphologized due to phonological attrition and the loss of certain inflectional suffixes.

From a synchronic point of view, the consonant gradation rules account for the weak consonant grade in the attributive form of the adjective but not for the strong grade in the noun with possessive marking. The Northern Saami words *suohkku* 'stocking' and *guhkki* 'long (pred.)' have open second syllables hence strong consonant stems (here a consonant cluster, the first part of which is a geminate /C̄C/). The second syllable in both forms is closed: *suohkkus* /suoh:.ku-s/ marked with the possessive suffix and *guhkis* /kuh.ki-s/ marked with the attributive suffix. However, the consonant stem of the noun *suohkkus* remains strong (/C̄C/) even before the syllable closing suffix, whereas the geminate part of the cluster is shortened (/CC/) in the adjective *guhkis*.

It is important to note that the possessive suffix is reconstructed as Proto-Saamic **-sē* **sammallahti1998b** and thus originally had a different syllable structure. The formative obviously did not close the second syllable in Proto-Saamic, as in **/kuh:.ke.-sē/ and **/suoh:.ku.-sē/.[17] From a diachronic point of view, the consonant gradation rules would thus account for the strong consonant grade in the noun marked with a possessive suffix but not for the weak grade in the

[17] Note that these invented examples in simplified transcriptions serve the purpose of illustration (and are hence marked with **). The stem of the adjective 'long' is reconstructed as Proto-Saamic **kuh̓kē* **sammallahti1998b** The noun 'stocking' is a loan word (cf. Swedish (dialectal) *sokk*, Finnish *sukka*) and might not be reconstructable for Proto-Saamic.

attributive adjective.

Two possible explanations could explain the different consonant grades in the noun and the adjective marked by means of *-s* ⇐ **-sē*.

- Following Nielsen (**nielsen1945b**), the possessive marker in its function as attributive nominalizer was originally attached to a genitive (i.e., weak stem) form of the adjective. The weak consonant stem was thus triggered by the genitive suffix, reconstructed as Pre-Proto-Saamic **-n* ⇒ Proto-Saamic **-Ø* **sammallahti1998b** and preceding the attributive marker. The date of the morphologization of stem gradation would not be relevant for this explanation.

- The other possible explanation presupposes a relatively late date for the morphologization of stem gradation, i.e., not earlier than the apocope of the possessive marker's final vowel (*-s* ⇐ **-sē*). If the possessive marker was not a true suffix but a phonological word on its own by the time stem gradation was morphologized in Saamic, the marker would have remained outside the phonological domains of its host word and would not have been able to trigger stem gradation on the latter.

Since genitive (or "possessor case") marking on attributive adjectives is attested in other northern Eurasian languages, as in both Yukaghir (see §??) and in Lezgic languages (see §??), Nielsen's assumption that the 3rd singular possessive marker was originally attached to an attributive form of adjectives (or other nominals) in genitive is principally possible.

Yet there is no evidence that genitive attribution marking on adjectives ever occurred regularly in Saamic or even in other Uralic languages.[18] Furthermore, the functional side of the assumed development, in which an adjective marked by two attributive markers (genitive+attributive nominalizer) simultaneously, would also need some further clarification.

The second hypothesis assuming that the possessive marker never triggered stem gradation, could also account for the weak consonant grade in adjectives (remember that the weak grade seemed to contradict the stem gradation rules from a historical point of view). In certain aspects, the possessive marker behaves like a free pronoun rather than like an affix: the possessive marker shows pronominal agreement (and hosts the agreement suffixes which co-reference the

[18] The "defective" agreement paradigm of pronouns (and even sometimes adjectives) with the genitive singular form in all cases except nominative singular can scarcely be connected to Nielsen's idea. As an anti-construct state marker, the "genitive" should occur through the whole paradigm including in nominative singular.

number of the possessor) but the marker itself is hosted by an inflected noun (marked for number and case of the possessed). Note also that the possessive inflection is morpho-syntactically different from case and number inflection in the closely related Finnic languages. Only the latter features trigger noun phrase internal agreement.

Only the 3rd person singular possessive marker was used as an attributive nominalizer. Since this marker was hosted by uninflected adjectives, it is reasonable to assume that at one point the nominalizing possessive marker behaved differently from true possessive markers. The attributive nominalizer might thus have become a true phonologically bound formative earlier than the homophonous possessive marker. As a result of the apocope of the suffix-final vowel, the second syllable in the attributive form was closed:

(25) ***/kuh:.ke.-sē/* ⇒ ***/kuh.ke-s/*
long-POSS:3SG long-ATTR

Subsequently, the stem gradation rules were applied regularly and yielded the short consonant grade of the adjective stem equipped with the affixal attributive marker. The noun equipped with the possessive marker, however, kept its open second syllable even after the apocope. The non-affixal possessive suffix – as a phonological word of its own – remained outside the phonological domain of stem gradation.

(26) ***/suoh:.ku.=sē/* ⇒ ***/suoh:.ku.=s/*
stocking=POSS:3SG stocking=POSS:3SG

9.2.2 The origin of anti-construct state in Saamic

Synchronic data from related Uralic languages provide good evidence in favor of the assumed grammaticalization path from possessive to anti-construct state marking in Saami.

- POSSESSIVE (3SG) ⇒ ATTRIBUTIVE NOMINALIZATION ⇒ ANTI-CONSTRUCT

The first step of this development, i.e., the use of the possessive marker as an attributive article, is attested in the Permic languages Komi-Zyrian and Udmurt. Note also that the possessive marker in Udmurt shows different morphological behavior depending on its function as a true possessive or as an attributive article. For more detail see the respective sections on the synchrony (§??) and diachrony (§??) of attribution marking in Udmurt.

The Permic languages are closely related to Saamic and, theoretically, the rise of attributive marking in these two branches of Uralic could go back to a common Proto-Uralic construction. True evidence to prove such a common development at a relatively early time is, however, missing. Quite the contrary, it could be objected that the innovation of a new type of attribution marking is currently under way in the Permic languages whereas the innovation in Saamic took place 2000 years ago and is obviously loosing ground today in favor of the re-introduced type juxtaposition.

But the comparison with the related Permic languages makes sense from a purely typological perspective. Assuming that the possessive marker already had a "determinative" function in Proto-Uralic (as stated, for instance, by **janhunen1981 decsy1990 kunnap2004**) and that this function is still present in most of the modern Uralic languages, the existence of an attributive nominalizer in Permic indisputably proves that the proposed origin of the attribution marker in Saamic is functionally plausible (**riesler2006b**).

Furthermore, the nominalizing function of the (person-deictic) marker of possession is attested not only in several Uralic languages but also in Turkic languages. And, finally, a typologically similar grammaticalization path of a (local-deictic) demonstrative to an attributive article is also attested in Indo-European languages of the area.

In all mentioned Turkic, Uralic and Indo-European languages where the development of attributive nominalizers is attested, this innovative type of attribution marking originally co-occurred with another, inherited type. The use of contrastive pairs of attributes marked with or without the anti-construct state marker in modern Saamic languages provides good evidence for a similar development in earlier stages of Saami.

Several grammatical descriptions of Northern Saami give examples of such contrastive pairs of attributes with different meanings. Nielsen describes the difference between forms with and forms without an attributive suffix as a difference in "modality" of the attributive relation **nielsen1945b** Most examples, however, do not display true adjectives but rather attributive forms of present participles. If the property denoted by the participle is stressed or emphasized as belonging permanently to the referent of the modified noun the participles are often equipped with the attributive suffix.

(27) a. Northern Saami (**nielsen1945b**)

i. *juhhki olmmoš*
drinking person
'drinking person'

ii. *juhkke-**s*** *olmmoš*
drinking-ATTR person
'alkoholic (i.e., a person addicted to drinking)'

b. Northern Saami (**bartens1989**)

i. *šaddi* *soahki – soahki lea šaddi*
growing birch – birch is growing
'growing birch' – '(a/the) birch is growing'

ii. *Goađi duohkin lea šaddi-**s*** *soahki.*
hut behind is growing-ATTR birch
'There is a fast growing birch behind the hut.'

Besides participles, there are even contrastive pairs of attributive adjectives or nouns which distinguish temporal versus permanent (or otherwise emphasized) properties.

(28) a. Northern Saami (**bergsland1976**)

i. *arve-dálki*
rain-weather
'rain-weather'

ii. *arvve-**s*** *dálki*
rain-ATTR weather
'wet weather (i.e., weather full of rain)'

It must be emphasized that these adjectives equipped with the attributive suffix are additionally marked as denoting permanent or "definite" properties. This is exactly consistent with the reconstructed meaning of the so-called weak adjective forms in Proto-Germanic or the so-called long adjective forms in Proto-Baltic/Slavic (see §**??**). The semantics of the regular and productive contrastive focus constructions in Chuvash and Udmurt (which are often described as "emphatic" or "definite" as well, see §§**??**, **??**) also show a perfect parallel to Saamic.

It is thus most likely that the Saamic anti-construct state marker originates from a construction in which the possessive marker 3rd person singular was used as attributive nominalizer in appositional noun phrases similar to the contrastive focus construction attested in Modern Udmurt and in several other Uralic and non-Uralic languages of northern Eurasia.

Whereas the unmarked noun phrase type in Proto-Saamic was characterized by juxtaposition, the attributive article was used to mark a construction with an

adjective in contrastive focus. The emphatic construction later became generalized as the default marker of the attributive connection.[19]

(29) Grammaticalization of anti-construct state marking in Saamic
  a. Stage 1: Pre-Proto-Saamic
    i. Juxtaposition
    [NP Along Nstocking]
  b. Stage 2a: Proto-Saamic
    i. Juxtaposition (default)
    [NP Along Nstocking]
    ii. Attributive apposition (emphatic)
    [NP [NP' Along HEADØ-NMLZ] Nstocking]
  c. Stage 3: modern Saamic languages
    i. Anti-construct state marking
    [NP Along-ATTR Nstocking]

The irregularities in the use of attributive forms within and across the modern Saamic languages are the result of recent developments. Originally, the attributive form was generated regularly and productively. A cross-comparison of adjectives in different Saamic languages clearly shows that adjectives with deleted *-s/-s'* in one Saamic language exhibit the suffix in another language. Consider, for example, Northern Saami *uhca* but Lule Saami *ucces* 'small' or Northern Saami *seakka* but Kildin Saami *sieŋŋkes'* 'thin' (for more examples see **riesler2006b**).

It is most likely that neither the predicative forms (ending in *-d* or *-s*) nor the attributive form (ending in *-s/-s'*) reflect inherited stems in Saami. Both are complex forms which are derived from either nominal or verbal stems by means of different suffixes. The predicative forms with *-s* evolved from derivations by means of an old lative case suffix. Germanic loan adjectives with the homophonous (Germanic) ending *-s* (⇐ Proto-Germanic *-R*) where integrated into the class of these predicative "lative-derivations". The attributive suffix *-s/-s'*, on the other hand, originates from the possessive marker 3rd person singular which was originally used as an attributive nominalizer (i.e., attributive article) in contrastive focus constructions. The suffix was later generalized as the default attributive state marker.

[19] The zero-morpheme (equipped with the nominalizer Ø-NMLZ) in (**??**) and following examples is only presented for a better illustration of the empty head position to which the (nominalized) adjective moves in the appositional noun phrase.

The merger of predicative and attributive forms of some adjectives observed in modern Saamic languages does not contradict the proposed reconstruction of the original attributive marking. It does, however, reflect another diachronic path of adjective attribution marking: namely the collapsing of an originally regular and productive construction and the innovation of a new type. Interestingly, this secondary development in modern stages of Saamic will most likely result in the renewed introduction of juxtaposition, i.e., the original Uralic prototype of adjective attribution marking.

9.3 The emergence of agreement in Finnic

The languages of the Finnic branch spoken in the northwestern periphery of Uralic are exceptional within this family because they exhibit head-driven agreement as the default type of attribution marking of adjectives.

(30) Finnish (personal knowledge)

a. *iso talo*
big house
'large house'

b. *iso-t talo-t*
big-PL house-PL
'large houses'

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

There is no doubt that agreement marking replaced juxtaposition at a certain point during the linguistic development from Proto-Uralic to Proto-Finnic.

In several Uralic languages, irregular agreement of pronominal modifiers and even some adjectives and adjective-like modifiers are attested (cf. examples in **honti1997** and **stolz2015a**). This might indicate a connection to the fully developed agreement marking of adjectives in Finnic. It is, however, unclear whether the incomplete and irregular agreement phenomena in Saamic and other closely related Uralic languages reflect a stage of development at which agreement marking was more widespread – in at least the Finnic and Saamic branches – or agreement marking is due to a more recent innovation which became completely enforced only in the Finnic branch.

The rise of agreement marking on attributive adjectives, pronouns, and numerals in Finnic is usually regarded as a result of language contact with Indo-European languages from the Germanic and/or Baltic groups (cf. **tauli1955 hajdu1996** see also **stolz2015a**). Indeed, the high amount of Germanic and Baltic loanwords in Finnic languages indicate intimate contacts between speakers of Uralic and Indo-European languages in that area. In order to prove the hypothesis that agreement marking arose as a result of influence from Indo-European languages, however, one has to reconstruct concrete mechanisms behind this profound contact-induced language change. The idea that agreement marking is a borrowed model might not be as straightforward as it appears. Even though many Uralic languages under strong Russian influence seem to have borrowed many more grammatical features than Finnic did under Germanic and Baltic influence, none of these languages shows any trace of borrowed Russian agreement marking.

In a short article, **mark1979** presents a contact-independent explanation of the innovative head-driven agreement marking in Finnic. His explanation is based on the observation that nominalized adjectives in apposition to nouns in Hungarian (as well as in other Uralic languages) show agreement triggered by the semantic head of the elliptic noun phrase.

(31) Hungarian (**mark1979**)
  a. Juxtaposition (no agreement marking)
    i. *őreg postást* [A $N_{nom.sg}$] ‘the old postman’
    ii. *őreg postások* [A $N_{nom.pl}$] ‘the old postmen’
  b. Apposition (agreement marking)
    i. *postást, őreg**et*** [[$N_{nom.sg}$] [$A_{nom.sg}$]] ‘a postman, an old one’
    ii. *postások, őreg**ek*** [[$N_{nom.sg}$] [$A_{nom.sg}$]] ‘postmen, old ones’

Similar ideas about a possible contact-independent origin of head-driven agreement in Finnic have also been put forward, for example by **ravila1941** and **papp1962** In theory, the rise of agreement marking as a result of generalization of an originally emphasized adjective in apposition seems plausible. Language contact with agreement-marking languages could still have been a catalyst.

In Hungarian, the attributive appositions described by Márk are post-positioned while attributive adjectives in Finnish still precede the noun. A comparison to attributive apposition by means of nominalization in Udmurt seems more promising. In §**??** on the synchrony of attributive marking in Udmurt, it has been demonstrated how case and number agreement marking occurs in the contrastive focus construction with attributive adjectives and pronouns.

(32) Udmurt (**winkler2001**)

a. Juxtaposition (no agreement marking)

i. *badʒ̑ym gurt* [A $N_{nom:sg}$] ‘large house’

ii. *badʒ̑ym gurtjos* [A $N_{nom:pl}$] ‘large houses’

iii. *badʒ̑ym gurtjosy* [A $N_{pl:ill}$] ‘to (the) large houses’

b. Attributive apposition (agreement marking)

i. *badʒ̑ym**ėz** gurt* [[A_{contr}] [N]] ‘LARGE house’

ii. *badʒ̑ym**josyz** gurtjos* [[$A_{contr:pl}$] [N_{pl}]] ‘LARGE houses’

iii. *badʒ̑ym**josaz** gurtjosy* [[$A_{contr:pl:ill}$] [$N_{pl:ill}$]] ‘to LARGE houses’

In both Hungarian and Udmurt examples (**??**) and (**??**), the agreement morphology is syntactically spread from the (semantic) head noun to the adjectival modifier only in appositional noun phrases (with the modifier in contrastive focus). In Udmurt, there is an additional morpheme available, i.e., the attributive nominalizer *-(ė)z* (⇐ POSS:3SG). In the Hungarian example, the emphasized construction is only marked by the duplicated number and case agreement (in combination with changed constituent order).

Attributive apposition in contrastive focus constructions is without a doubt innovative in Udmurt. Since all members of the Permic group show similar constructions, the development could be dated back to Proto-Permic and would thus have a time depth comparable to the innovation of head-driven agreement in Finnic. Since head-driven agreement is also involved in Udmurt anti-construct state marking (namely as a “relict” of the appositional structure in which the attribute in contrastive focus originally occurred), the Permic and Finnic innovations could be structural parallels. Modern Finnic languages, however, do not provide any evidence that an attributive nominalizer was ever used as a marker of appositional attribution. The agreement markings thus seems to be the primary innovation assumedly caused by contact with “agreeing” Indo-European languages. Regardless of contact influence being involved or not, the innovative head-driven agreement marking in Finnic could still have been used in an appositional construction originally. Note also that in Udmurt, number agreement sometimes (irregularly) occurs even in constructions without the contrastive focus marker.

(33) Head-driven plural agreement in Udmurt (**winkler2001**)
badʒ̑ym-jos gurt-jos
big-PL house-PL
‘LARGE houses’

Note even that a similar innovation of head-driven agreement in contrastive focus constructions is attested not only for Permic languages but also occurs irregularly in other Uralic branches (cf. **honti1997** for Mari and Nenets; **siegl2013a** for Tundra Enets).

To conclude these tentative considerations, it cannot be ruled out that the rise of head-driven agreement marking in Finnic and anti-construct state agreement in Udmurt are both results of original attributive apposition constructions. For Finnic, however, this idea remains highly speculative unless one can find evidence for the occurrence of an attributive nominalizer such as the marker in Modern Udmurt or in Proto-Saamic.

Whereas anti-construct state agreement marking in Udmurt (and other Permic languages) only substitutes for the default marker in contrastive focused constructions, Finnic and Saamic have completely lost Uralic juxtaposition as the default adjective attribution marking device and innovated completely new morpho-syntactic devices. It must also be noted that the Finnic and Saamic innovations took place in two closely related and geographically adjacent branches of Uralic. Moreover, the developments are of similar age. And finally, non-related but geographically adjacent languages (Baltic, Germanic, Slavic) show structurally similar developments.

9.4 Other attested scenarios of grammaticalization

The previous sections dealt with the rise of adjective attribution marking devices in a few branches of Indo-European, Uralic and Turkic. The synchronic data from the synchronic survey in Part III (Synchrony), however, present evidence of several diachronic scenarios. Only a few of them will be sketched in the following sections.

9.4.1 Articles, definiteness and the evolution of adjective attribution marking in Indo-European

The rise of attributive articles and their (partial or complete) further development to definite markers in Baltic, Slavic and Germanic, as described above, took place on functionally and chronologically parallel paths in various other Indo-European languages of Europe. This has been observed by several scholars (cf. **brugmann-etal1916 gamillscheg1937 heinrichs1954** and, more recently, **nocentini1996 philippi1997 himmelmann1997**). It is not clear whether these parallel developments across western-Indo-European branches can be explained in

terms of areal typology, i.e., as the result of linguistic contacts, or whether they are inherited from a common ancestor language. Independent developments, though theoretically possible, seem rather unlikely given the close genealogical and areal connection between the languages in question.

In those western branches of the Indo-European family where definite markers have evolved, cognate formatives are also usually attested as adjective attribution markers. The attributive article in Rumanian, for instance (see §**??**), is also attested in Latin and other Romance languages, cf. Latin *Cato ille maior, Babylon illa magna*.[20] The suffixed definite marker in Rumanian evolved from this attributive article (**gamillscheg1937 nocentini1996**). Note also that the attributive article in Romance is polyfunctional and can mark adjectival, genitival and prepositional attributes as well as relative clauses.

In the two Albanian languages (see §**??**), the attributive article *i* NOM, *e/të* ACC and *të* OBL and the definite suffix *-i* NOM, *-in/-në* ACC and *it* OBL most likely have the same etymological source, i.e., Indo-European **-to* (cf. **himmelmann1997** with references), which is also the etymological source of the definite marker *to* and the homophonous attributive article in Ancient Greek (see §**??** for the corresponding constructions in Modern Greek).

Indo-European **-to* is the etymological source of secondary attributive articles in Slavic languages as well. The use of this marker in attributive apposition constructions is already well-attested in Old East Slavic documents.

(34) Attributive nominalization in Old E-Slavic (Indo-European)

a. *[…] sъ usmъ galiiei-sk-**ymъ***
with Jesus:COM Galilee-ADJZ-NMLZ:INSTR
'[…] with Jesus the Galilaen' (**mendoza2004**)

b. *vъ sareftǫ sidonъ-sk-**ǫjǫ***
to Sarepta:PREPOS Sidonia-ADJZ-NMLZ:ACC
'to Sarepta in Sidonia' (**mendoza2004**)

In Bulgarian, the former attributive nominalizer grammaticalized into a true definite marker. In an analogous manner (but much later in time), reflexes of the Proto-Baltic/Slavic pronoun **tъ* M developed into definite suffixes in northern Russian dialects (cf. **leinonen2006a**).[21]

[20] Cf. the secondary attributive articles in Germanic languages in similar constructions: English *Philip the Fair*, German *Friedrich der Große* which is also cognate (and homophonous) with the definite marker. The Germanic constructions have been dealt with in more detail in §**??**.

[21] Whereas Komi-Zyrian (Uralic) influence triggered the suffixation of these anaphoric markers in northern Russian dialects (**leinonen2006a**), a typologically similar grammaticalization

Dahl (**dahl2003** see also **dahl2015a**) shows that in some languages definite noun phrases with attributive adjectives (or other adnominal modifiers) show special behavior. He compares the "displaced"[22] definite marking with "long form" adjectives in the Baltic languages with, among others, the demonstrative *ille* linking postponed adjectives to proper nouns in Latin constructions like *Babylon illa magna* **dahl2003** But due to its function and syntactic behavior the attributive article in Romance can clearly be distinguished from definite markers (**gamillscheg1937**). As it was demonstrated for the Baltic languages (see §**??**), the so-called "long form" inflection (i.e., anti-construct state agreement inflection) of adjectives is not a true definiteness marker.

Dahl also gives examples of languages in which "displaced" definiteness markers (or "quasi-definiteness markers") evolved from other sources than local-deictic pronouns, as in Amharic where an attributive nominalizer grammaticalized from a (person-deictic) possessive marker in contrastive focus construction.

(35) Amharic (Afro-Asiatic; **hudson1997**)

a. Default construction

i. *təlləq bet*
large house
'(a) large house'

ii. *təlləq bet-**u***
large house-POSS:3SG
(1) 'his large house' (if the owner has only one house, which is large); (2) 'the large house'

iii. *təlləq bet-e*
large house-POSS:1SG
'my large house'

b. Contrastive focus construction

i. *təlləq-**u*** *bet*
large-?DEF house
'(a/the) LARGE house'

ii. *təlləq-**u*** *bet-**u***
large-?DEF house-POSS:3SG
'his LARGE house' (if the owner has more than one house but the

process due to TurkicTurkic languages influence is behind the chronologically much older suffixation of definite marking in Bulgarian (**kusmenko2008**).

[22] The term "displaced" is not used by Dahl but adopted from **melcuk2006**

expression is referring to the large one)

iii. *təlləq-**u*** *bet-e*
large-?DEF house-POSS:1SG
'my LARGE house'

The suffix *-u* [M] used for emphasizing the adjective in Amharic is homophonous with the definite noun marker and with the 3rd singular possessive marker. Note that the possessive and the definite suffixes of nouns (or noun phrases) are mutually exclusive (**hudson1997**). Hence, the examples in (**??**) are ambiguous; they could have a possessive or a definite reading. The "emphasizing" adjective suffix *-u* [M], however, does not co-occur with the definite suffix. Therefore, the reading of the examples in (**??**) is not ambiguous.

Consequently, the suffix *-u* [M] in Amharic should be analyzed as an adjective attribution marker rather than as a "detached" marker of definiteness.

(36) Amharic (Afro-Asiatic; **hudson1997**)

a. Attributive nominalization (contrastive focus)

i. *təlləq-**u*** *bet*
large-ATTR house(M)
'(a/the) LARGE house'

ii. *qonjo-**wa*** *dəmmät*
pretty-ATTR:F cat(F)
'(a/the) beautiful cat'

b. Attributive nominalization (headless noun phrase)

i. *təlləq-**u***
large-ATTR:M
'(a/the) big one'

ii. *qonjo-wa*
pretty-ATTR:F
'(a/the) pretty one'

Contrastive focus marking on adjectives in Amharic is thus very similar to the marking found in Udmurt. In both languages, attributive apposition is marked by means of attributive nominalization. The respective formatives in both languages originate from (person-deictic) possessor markers.

Consistently, data from northern Eurasian languages and Amharic do not provide evidence for the existence of "displaced" definiteness markers. From a diachronic perspective, however, there is much evidence for a functional overlap-

ping between attributive nominalization and definiteness marking. In all Indo-European languages dealt with so far, adjective attribution is the primary function. The former local-deictic marker in these languages always grammaticalizes into an attributive nominalizer first. The further development into true markers of definiteness comes only after this stage.

9.4.2 The emergence of head-marking attributive construct state in Iranian

As shown in §??, several Iranian languages of the northern Eurasian area exhibit a head-marking attributive construct state device as a licenser of adjective attribution. The Iranian construct state marker (aka *Ezafe*) originates from the Old Iranian relative particle *-hya*, which has undergone a process of grammaticalization, to end up as a part of nominal morphology in the modern Iranian languages (**haider-etal1984**; **samvelian2007b**). Since the Old Persian relative particle *-hya* itself originates from a demonstrative, the emergence of construct state marking in Iranian and anti-construct state marking in other Indo-European languages follow a similar path. Originally, *-hya* was a grammatical word marking the phrase or clause on its right as a syntactic modifier of the noun on its left (**haider-etal1984**). Syntactically, the marker was an attributive article hosted by the attribute. In BalticBaltic languages and Slavic,Slavic languages the article developed further into an anti-construct state agreement marker (see §??). In Iranian, however, the article attached phonologically to the head noun. According to **samvelian2007** this conflict between opposite directions of phonological and syntactic alignments was later resolved by the re-analysis of the article as a head-marking inflectional affix. As the result of this grammaticalization, syntactic and phonological attachments were alined to each other.

9.4.3 Innovation of juxtaposition

Two scenarios are attested where juxtaposition has been innovated: either by loss of agreement marking or by loss of anti-construct state marking.

9.4.3.1 Loss of agreement marking

Head-driven agreement (in number and case) of adjectival modifiers following the head noun can be reconstructed for Common Kartvelian. In Old Georgian, this pattern is more or less preserved. In modern Kartvelian languages, however, the unmarked constituent order of adjectival modifiers and head is noun-final,

although the opposite order is possible as well (**harris1991a**). As shown in §?? of Part III (Synchrony), the agreement features of Common Kartvelian are more or less preserved only in the marked (but inherited) head-initial noun phrase type. In the head-final noun phrase type, on the other hand, modern Kartvelian languages display a strong tendency to lose head-driven agreement. Preposed attributive adjectives in Mingrelian and Laz are juxtaposed to the head noun as a rule. In Modern Georgian and Svan, the agreement paradigm of preposed attributive adjectives shows a high degree of syncretism (cf. **harris1991a tuite1998**).

Two other non-related languages of the Southern Caucasus, Armenian and Ossetic have lost noun phrase internal agreement too (**2**)72–281]stolz2015a.[23] According to **johanson2002a** Turkic contact influence is the explanation for the loss of agreement in Armenian.

Interestingly, the loss of adjective agreement marking in Armenian and Kartvelian is connected to the shift of the default constituent order. Note, however, that juxtaposition can also be innovated without constituent order shift, as in English where the change is a result of the complete loss of the agreement inflection during the course of time from Middle to Modern English.

9.4.3.2 Loss of anti-construct state marking

Saamic languages present another evidence of a language change in which juxtaposition replaces an original morpho-syntactic device. The original anti-construct state marking, which is itself innovative in Proto-Saamic (see §??) is in dissolution in modern Saamic languages as the result of the merger of attributive and predicative adjective forms which were originally distinguished from one another.

9.5 Diachronic polyfunctionality

In §??, a few examples of polyfunctional adjective attribution marking devices were presented. It was shown, however, that the polyfunctionality parameter is less relevant to northern Eurasian languages because most languages of the area exhibit highly differentiated attribution marking devices. Polyfunctionality

[23] The innovation of juxtaposition in the Eastern Armenian standard language is not complete, though. There is a small class of adjectives which are marked by means of head-driven agreement, see§??.

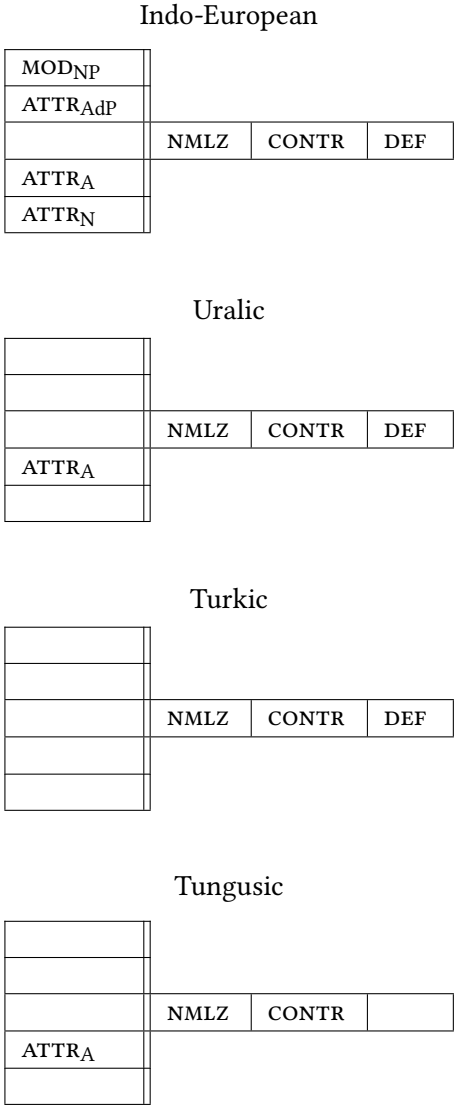


Figure 9.1: Functional map of markers cognate with the Old Iranian "relative particle" *-hya* (across Indo-European languages) and the possessive suffixes 3rd person singular (across Uralic, Turkic and Tungusic languages)

might, however, indicate a historical dimension if additional semantics of attribution marking devices is taken into consideration and if the languages of a whole genera are compared to each other. For instance, construct state marking of adjectives and other modifiers, as attested especially in Indo-European varieties (but also in Turkic and Uralic), seems to be inherently tied to the evolution of attributive nominalization, contrastive focus and even definiteness marking in several languages. Figure **??** shows functional maps similar to the one in Figures **??** and **??** in §**??** but with scope over cognate markers in whole language families.

The polyfunctionality of the Persian Ezafe *-(y)e* was described in §**??**. This construct state marker licenses nominal (ATTR_{N}), adjectival (ATTR_{A}) and adpositional (ATTR_{AdP}) attributes as well as modification within an adposition phrase (MOD_{NP}). The cognate formative in the closely related Iranian language Northern Kurdish is even connected to definiteness marking (DEF) (**schroder2002** cf. also Table **??** on page **??**). In Old Iranian, Old Baltic and Old Slavic languages, a cognate marker was used as an attributive nominalizer (NMLZ, or as a "relative particle" marking non-verbal attributes; see §**??** and **??**). The marker's further grammaticalization into an anti-construct state agreement marker in Baltic and Slavic is connected to contrastive focus marking (CONTR).

The marker described in the functional map for Uralic is the possessive suffix 3rd person singular, which is used as a quasi-definite marker (DEF) in a variety of modern Uralic languages. In Udmurt the original possessive suffix is regularly used as a nominalizer (NMLZ) and has grammaticalized into a marker of contrastive focus of adjectives (CONTR) (see §**??**). In Saamic, finally, the cognate marker has grammaticalized into an anti-construct state marker (ATTR_{A}).

Turkic is similar to Uralic but without evidence for the grammaticalization of the possessive suffix 3rd person singular to a true adjective attribution marker. In Tungusic, finally, there is no evidence for definiteness marking but the possessive suffix 3rd person singular is used as dependent-driven agreement marker in (ATTR_{A}).

These diachronic functional maps demonstrate general synchronic paths of attribution marking devices and give the impression that nominalization and appositional attribution play an important role in the further development of the respective markers as attribution marking devices.

10 Areal typology in the Circum-Baltic area

The Circum-Baltic area can be defined geographically as the drainage area of the Baltic Sea. Autochthon languages belonging to this area are mostly from the Germanic, Baltic and Slavic branches of Indo-European as well as from the Finnic and Saamic branches of Uralic. Several authors have tried to establish a Circum-Baltic Linguistic Area (Sprachbund) based on shared linguistic features across member languages of this area (for instance **koptjevskaja-tamm2006**).

riesler2006a described areality in the morpho-syntax of noun phrase structure in the Circum-Baltic languages. It is conspicuous that both the languages of the two Uralic branches of the area and the languages of the three contacting Indo-European branches have innovated adjective attribution marking devices which deviate from the prototypes of their respective families.

Saamic innovated anti-construct state marking and Finnic innovated head-driven agreement. The prototype of adjective attribution marking in Uralic, however, is juxtaposition. Except in Saamic and Finnic, juxtaposition occurs in all Uralic languages as the default adjective attribution marking device (see §**??**) and is also reconstructed for Proto-Uralic (**decsy1990 janhunen1981**).

Head-driven agreement is the prototype of adjective attribution marking in Indo-European and is also the type reconstructed for Proto-Indo-European (**decsy1991**; **watkins1998**). In Germanic, Baltic and Slavic, however, a secondary type evolved from attributive nominalization. Consequently, several modern languages of these branches exhibit anti-construct state agreement marking as a default or secondary device.

All five Circum-Baltic branches (Germanic, Baltic, Slavic, Finnic, Saamic) of the "buffer zone" have thus undergone change and innovated adjective attribution marking devices which deviate from the prototypes of their respective families:

- **Finnic**:

  Juxtaposition ≠ head-driven agreement

- **Saamic**:

Juxtaposition ≠ Anti-construct state

- **Germanic, Baltic, Slavic**:

  Head-driven agreement ≠ Anti-construct state agreement

The developments in Saamic and in the three Indo-European branches can even be connected to each other in structural terms: the innovative anti-construct state (agreement) marking in these languages evolved from an attributive apposition construction marked by means of attributive nominalizers. The principal grammaticalization paths are thus similar:

(1) $[_{NP}\,[_{NP'}\,{}_{A}big\;{}_{HEAD}\text{Ø-}_{NMLZ}]\;{}_{N}house] \Rightarrow [_{NP}\,{}_{A}big\text{-}_{ATTR}\;{}_{N}house]$

Therefore, **riesler2006a** described the result of this areal innovation as a "grammaticalization area" (**heine-etal2005**), i.e., a linguistic area of geographically neighboring languages in which similar processes of grammatical changes took place as the result of language contact. According to **heine-etal2005** a *model language* must affect at least two different *replica languages* in a grammaticalization area. In the case described here, a pre-proto-stage of either Germanic or Baltic/Slavic could probably be the "model" since attributive nominalization by means of cognate markers evolved in several other branches of Indo-European. But even Uralic influence should be considered. Possible model and replica languages of the area are thus:

- Proto-Baltic/Slavic < **Pre-Proto-Germanic** > Proto-Saamic
- Proto-Germanic < **Pre-Proto-Baltic/Slavic** > Proto-Saamic
- Proto-Baltic/Slavic < **Pre-Proto-Saamic** > Proto-Germanic

Given the high age and the cognate constructions and formatives in other Indo-European branches (mostly Iranian) and considering other attested Baltic contact influence on Saamic,[1] it seems most plausible to locate the core of the grammaticalization area in the Baltic/Slavic groups of Indo-European. Saamic and Germanic have probably borrowed the model of attributive nominalization but realized the construction with their own inherited morpho-syntactic means.

Nonetheless the vast geographic spread of cognate constructions among several Indo-European, Uralic, Turkic and even Tungusic branches makes it also possible to assume a source outside both Indo-European and Uralic and a development preceding the proto-stages of these language families.

[1] See, for instance, **riesler2009** for lexical borrowings.

Part V

Conclusion

11 Results and conclusions

These concluding sections summarize the essence of this study and provide an overview of the main findings. In addition, they address a few questions relevant for future investigations towards a general description of noun phrase structures.

11.1 Aims and content

The aims of this study were: (1) a synchronic-typological description of adjective attribution marking devices in northern Eurasia, i.e., typologizing geographically relevant languages according to their syntactic and morpho-syntactic kinds of adjective attribution marking, (2) a synchronic survey of the geographic distribution of the attested kinds of adjective attribution marking devices across the northern Eurasian languages, and (3) a diachronic description and functional reconstruction of a hitherto underscribed pattern in the evolution of adjective attribution marking in the Indo-European and Uralic languages of the Circum-Baltic area of northern Europe.

(1) As the main result of the **synchronic-typological description**, an ontological classification of attested syntactic and morpho-syntactic types of adjective attribution marking devices was developed. For the purpose of comparison and achieving stringent classification standards, even interesting devices attested in languages outside the area were taken into consideration.

Central typological parameters for the morpho-syntactic description of noun phrase structure are *syntactic source* (i.e., the central syntactic operation which licenses attribution and belongs primarily either to agreement marking or to government), *syntactic pattern* (i.e., devices projecting embedded noun phrases, devices projecting simple adjective phrases, or incorporation) and *syntactic locus* of the respective formatives (on-head, on-dependent, floating).

The following overview lists all known devices (one single device, which is not attested in the northern Eurasian area, is given in parentheses).

- Juxtaposition (as in Komi-Zyrian)

- Incorporation (as in Chukchi)
- Construct state (as in Northern Kurdish)
- Anti-construct state (as in Skolt Saami)
- Attributive nominalization (as in Udmurt)
- Attributive article (as in Yiddish)
- Anti-construct state agreement (as in Russian)
- Head-driven agreement (as in Finnish)
- Appositional head-driven agreement (as in Georgian)
- Modifier-headed possessor agreement (as in Saliba)
- (Linker) (as in Tagalog [attested only outside northern Eurasia])

A more detailed overview of the attested types including definitions and an ontological cross-classification is presented in Table ?? and Table ?? on pages ??–?? and in Figure ?? on page ??.

(2) The **synchronic survey** showed that the most common types of adjective attribution marking devices are head-driven agreement (the Indo-European prototype which characterizes most parts of the European linguistic map) and juxtaposition (the prototype in Uralic, Turkic and Mongolic monotonously characterizing larger parts of North Asia). Modifier-headed possessor agreement is the least common type in northern Eurasia since it is known to occur only in Tungusic. The Mongolic and Turkic families of North Asia exhibit a very low degree of diversity in regard to their adjective attribution marking devices. A relatively high degree of diversity characterizes several branches of Indo-European (especially Germanic and Indo-Iranian) and Uralic (especially Saamic). Typological diversity is thus predominantly found in peripheral subareas of Northern Eurasia where different language families meet, for instance in the Circum-Baltic area in northernmost Europe and in Inner Asia (§??).

(3) The **diachronic description** revealed a re-occurring pattern of attributive nominalizers developing further into attributive state markers and various other types of attribution marking devices in different languages of the area and during different periods of time. These structurally similar diachronic paths, which

had not yet been systematically investigated from a cross-linguistic perspective, were reconstructed in detail for Baltic/Slavic (§??), Germanic (§??) and Saamic (§??). In the three Indo-European branches Baltic, Slavic and Germanic, anti-construct state agreement marking evolved from attributive nominalization. Anti-construct state marking arose in the Saamic branch of Uralic as the result of a structurally similar development from attributive nominalization. The developments in the geographically adjacent but genealogically unrelated languages present evidence for areality across Circum-Baltic languages (§??).

The thesis also provides an extensive appendix including a list of 234 languages sorted by their genealogical affiliation and coded for attested noun phrase types (table starting on page ??) as well as a collection of maps illustrating the spread of attested noun phrase types across a world sample of languages (Figure ?? and Figure ?? on pages ??–??), across all northern Eurasian genera (Figures ??–?? on pages ??–??) and across Europe languages (Figure ?? and Figure ?? on pages ??–??).

11.2 Innovative findings

The study presents the first systematic description and mapping of all attested adjective attribution marking devices in the languages of northern Eurasia. It also provides the first complete ontology of adjective attribution marking devices based on syntactic and morpho-syntactic noun phrase types found in northern Eurasian languages. The geographic spread of different adjective attribution marking devices across the main genera of all northern Eurasian language families is surveyed and mapped similar in a way similar to the surveys carried out by the EUROTYP program[1] but covering a larger area.

The present study has a strong diachronic component. Synchronic typological research certainly sheds light on the evolution of language; nevertheless, linguistic typology can scarcely be considered a historical discipline per se since the applied method is most often exclusively a synchronic comparison of linguistic data. The present investigation, however, achieved a historical reconstruction of adjective attribution marking in several languages by using the historical-comparative method in combination with synchronic typology. By applying this innovative methodological approach a new hypothesis about the origin of secondary adjective attribution marking devices in Germanic, Baltic, Slavic and Saamic can be

[1] Cf. http://www-uilots.let.uu.nl/ltrc/eurotyp/

proposed.

The three most important results of this study are (1) the discovery that STATE has to be included in the inventory of morpho-syntactic features, (2) the finding that adjectival modifiers can be phrasally embedded constituents, and (3) the diachronic attestation of contrastive focus constructions with phrasally embedded adjectival modifiers as a common source of innovative adjective attribution marking devices in the northern Eurasian languages.

11.2.1 The morpho-syntactic feature STATE

Morpho-syntax is commonly understood as phrase internal morphology, i.e., morphology assigned by syntax. The inventory of morpho-syntactic features thus excludes true morphological features which are assigned to phrasal constituents from (phrase internal) syntax. Prototypical examples of morphological features not assigned by noun phrase internal syntax are inflectional class of a noun (an inherent feature), definiteness marking of a noun (a feature assigned by semantics) or accusative marking of a noun phrase in object position inside a verb phrase (a morpho-syntactic feature assigned inside a verb phrase). The most typical morpho-syntactic features in noun phrase syntax are assigned by agreement triggered by one constituent, for instance adjective agreement in definiteness or in accusative case. If agreement of dependent constituents is triggered by a head noun the relevant feature has first to be assigned to the head from outside: either by semantics (e.g., definiteness) or by noun phrase external syntax (e.g., accusative).

Feature inventories (like the inventory presented by **kibort2010a**), however, do not yet include instances of morphological marking triggered not by constituents but by the syntactic structure as such. The present study provides an important contribution to the general typology of morpho-syntax by complementing the known inventory of morpho-syntactic features with truly morpho-syntactic devices, such as the well-known "construct state" in Persian. The state marker in Persian is not the result of either agreement or government but is assigned by syntax alone.

State markers (glossed in the following examples with the value MOD "modification") can occur with different loci, i.e., on-head (??), on-dependent (??) or floating (??).

(1) a. Head-marking STATE in Persian (Indo-European)
(cf. in more detail page ??)

xane-ye *bozorg*
house-MOD big
'a/the large house'

b. Dependent marking state in Kildin Saami (Uralic)
(cf. in more detail page **??**)
ēl'l'-es' *pērrht.*
high-MOD house
'a/the high house'

c. Floating state in Tagalog (Austronesian)
(cf. in more detail page **??**)
maganda-ng *bahay / bahay na* *maganda*
beautiful-MOD house house ATTR beautiful
'a/the beautiful house'

As a morpho-syntactic feature, however, STATE is not restricted to noun phrase structure. In the following example, a state marker (glossed as a "modification marker") licenses a noun phrase as the dependent constituent inside an adposition phrase.

(2) Dependent marking state in Kildin Saami (Uralic)
(cf. in more detail page **??**)
pērht *al'n*
house\MOD on
'on a/the house'

11.2.2 Embedded adjectival modifiers: synchrony

It is common knowledge that noun phrases can contain simple modifiers (like simple nouns: *stone house* or adjective phrases: *a big house*), embedded phrasal modifiers, i.e., modifiers which are projected as complex noun phrases themselves (like an adnominal possessor noun phrase: *John's sister's house*), or complex modifiers which are projected higher than noun phrases (like an adnominal adposition phrase: *a house in the village* or an adnominal relative clause: *a house which is huge*). It was demonstrated in the present analysis that even adjectival modifiers can constitute embedded noun phrases and occur in attributive apposition constructions, as in Udmurt:

(3) Embedded adjectival attribute in Udmurt (Uralic)
(cf. in more detail page **??**)

badǯym-ėz gurt
[NP [NP' Abig-NMLZ] Nhouse]
'a/the LARGE house'

Unexpected agreement features provided evidence for the embedded adjectival modifier in Udmurt (as well as in other languages). Such attributive apposition constructions are syntactically similar to the well-known nominalizations in Southeast Asian languages languages, **bickel1999** on the "Standard Sino-Tibetan nominalization pattern". In the northern Eurasian area such constructions with embedded modifiers are especially common in contrastive focus constructions and as the diachronic source of several other adjective attribution marking devices (see also §**??**).

Consequently, the syntactic ontology of adjective attribution marking presented in this study includes the phrasal projection of the attribution marking device as a central parameter with three values:

- embedded modifier
- simple modifier
- incorporated modifier

These parameters are applicable in a typology of general noun phrase syntax (including modifiers which are not adjectives and modifiers which are not simple constituents). Consider Table **??** (derived from Table **??** on page **??**) which includes a phrasally embedded attribute (like the juxtaposed relative clause in Minangkabau, example **??**), a simple attribute (like the juxtaposed adjective in Komi-Zyrian, example **??**) and an incorporated attribute (like the incorporated possessor in Västerbotten Swedish, example **??**).

(4) a. Juxtaposed relative clause in Minangkabau (Austronesian; **gil2005**)
batiak Kairil bali
papaya Kairil buy
'the papaya Kairil bought'

b. Juxtaposed adjective in Komi-Zyrian (Uralic; **lytkin1966a**)
bur mort-jas
good person-PL
'good people'

c. Possessor noun incorporation in Västerbotten Swedish (Indo-European; **gil2005**),
Pelle-äpple
Pelle-apple
'Pelle's apple'

Table 11.1: Ontology of general noun phrase structure (derived from Table **??** on page **??** and restricted to morphologically unmarked attribution marking devices, i.e., phrasally embedded, simple and incorporated attributes)

Phrasally embedded attribute	Simple attribute	Incorporated attribute
"juxtaposed Rel"	"juxtaposed A"	"incorporated Psr"
$[_{NP}[_{Rel}$ NP V] N]	$[_{NP}$ A N]	$[_{NP}$ N_{PSR} N_{PSD}]

11.2.3 Embedded adjectival modifiers: diachrony

Adjectival modifiers which are embedded as a noun phrase projection are common cross-linguistically in contrastive focus constructions (see also §**??**), as in Udmurt:

(5) Juxtaposed simple and embedded adjectival attribute in Udmurt (Uralic) (cf. more detailed page **??**)
a. Juxtaposition (default)
badźym gurt
$[_{NP}$ $_{A}$big $_{N}$house]
'a/the large house'
b. Attributive nominalization (contrastive focus)
badźym-ėz gurt
$[_{NP}$ $[_{NP'}$ $_{A}$big-nmlz] $_{N}$house]
'a/the large house'

In Udmurt, as in other languages where attributive nominalization is attested in constructions with adjectives in contrasted focus, focus always takes scope over a whole noun phrase (but not over an adjective phrase). This explains why

the adjective phrase has to be nominalized and occurs in an attributive appositional construction (i.e., embedded as noun phrase with an empty head), see Table **??**. This synchronic finding is directly connected to the diachronic ev-

Table 11.2: ??‼

Simple noun in contrastive focus		$[_{NP}$		$N]_{focus}$
Noun phrase with adjectival modifier in contrastive focus		$[_{NP}$	A	$N]_{focus}$
Embedded adjectival modifier in contrastive focus		$[_{NP}$ $[_{NP'}$	$A]_{focus}$	N]
Simple adjectival modifier in contrastive focus (impossible)	*	$[_{NP}$	A_{focus}	N]

idence for attributive apposition because attributive nominalization is a major (and chronologically re-occurring) diachronic source for the grammaticalization of new adjective attribution marking devices in different languages of the area.

The ultimate etymological source of attributive state marking formatives are prototypically local or person deictic markers (which also tend to be reanalyzed as markers of definiteness, cf. Figure **??** on page **??**). These markers are initially used as attributive nominalizers in contrastive focus constructions and later reanalyzed either as anti-construct state markers or anti-construct state agreement markers:

(6) a. $[_{NP}[_{NP'}$ A-NMLZ$]_{focus}$ N] ⇒ $[_{NP}$ A-ATTR N]
b. $[_{NP}[_{NP'}$ A-NMLZ:AGR$]_{focus}$ N] ⇒ $[_{NP}$ A-ATTR:AGR N]

11.3 Other findings

* Information structure and the evolution of attribution marking Cross-linguistic data show how relevant information structure is for the description of noun phrase syntax: secondary adjective attribution marking devices occur in contrastive focus constructions in Indo-European, Uralic, Turkic, Tungusic and Kartvelian,. Since contrastive focus has scope over a whole noun phrase (but not over an adjective phrase) in all attested cases, the adjective is used in an attributive appositional construction, i.e., in an embedded noun phrase.

Information structure is also relevant to diachronic noun phrase syntax because in several languages of northern Eurasia new primary devices were innovated from attributive appositional constructions. A typical grammaticalization path starts with attributive nominalization used as a secondary device in contrastive focus constructions. The original emphatic construction with a phrasally embedded adjective is later reanalyzed as a default attribution marking device (either as anti-construct state or as anti-construct state agreement).

Such a development started for instance in Proto-Baltic/Slavic and Proto-Germanic where attributive nominalization arose as a secondary adjective attribution marking device (alongside the original head-driven agreement device) in contrastive focus constructions and developed further into anti-construct state agreement:

(7) $[_{NP}[_{NP'}$ A-NMLZ:AGR] N] $\Rightarrow$ $[_{NP}$ A-ATTR:AGR N]

The etymological source of anti-construct state agreement markers in the Indo-European branches are local-deictic markers (demonstratives).

Similarly, in Proto-Saamic attributive nominalization arose as a secondary adjective attribution marking device (in addition to the original juxtaposition) in contrastive focus constructions and developed further into anti-construct state:

(8) $[_{NP}[_{NP}$ A-NMLZ] N] $\Rightarrow$ $[_{NP}$ A-ATTR N]

The etymological source of anti-construct state marking in Saamic is a person-deictic marker (possessive suffix).

Even Proto-Finnic head-driven agreement likely originated in a contrastive focus construction, specifically from appositional head-driven agreement which was reanalyzed as the default adjective attribution marking device under Indo-European influence:

(9) $[_{NP}$ $[_{NP'}$ A-AGR] N] $\Rightarrow$ $[_{NP}$ A-AGR N]

* Attributive nominalization and definiteness marking Data from Saamic and from other Uralic and Turkic languages in which attributive nominalizers originate from the possessive suffix 3rd person singular contradict Himmelmann's (**himmelmann1997**) assumption that a functional convergence between attributive nominalizers and definiteness markers with a person-deictic or a local-deictic etymological source is unlikely to occur.

The data is, however, in accordance with Himmelmann's (**himmelmann1997**) assumption about the functional extension of deictic elements to attributive and

definite markers if one acknowledges that definite markers with a local-deictic etymological source can evolve from attribution markers (but not vice versa), as in Indo-European:

- DEM ⇒ NMLZ (⇒ DEF)

By contrast, in the Uralic and Turkic languages, in which the etymological source is a person-deictic marker, attribution markers evolved from definite markers:

- POSS ⇒ DEF ⇒ NMLZ

This finding implies an implicational universal: *Possessive markers develop to attributive nominalizers only in languages in which similar possessive markers are already used as markers of (quasi-) definiteness* (cf. Universal **??** on page **??**).

* "Displaced" definiteness marking on adjectives Synchronic and diachronic data from the languages analyzed in the present study provide clear evidence against the existence of "displaced" definiteness marking on attributive adjectives (as proposed, for instance, for Baltic languages or for Amharic; cf. **dahl2015a**). The primary function of the respective markers is always the licensing of adjective attribution (by means of attributive nominalization in contrastive focus constructions). Even though there is a functional overlapping between attributive nominalization and definiteness marking from a diachronic perspective, the grammaticalization of definiteness marking is secondary in all attested cases.

* The northern European "buffer zone" The Circum-Baltic branches Baltic, Germanic, Slavic (Indo-European), Saamic and possibly also Finnic (both Uralic) constitute a "buffer zone" (similar to Stilo's **stilo2005** notion of this term) between the Indo-European and Uralic prototypes of noun phrase structure.

The Circum-Baltic "buffer zone" is the result of areal grammaticalization processes (similar to the notion of "grammaticalization area" by **heine-etal2005**) in which new adjective attribution marking devices were grammaticalized from original attributive appositional constructions marking contrastive focus on the adjective. The developments are most likely the result of contact-induced changes and originate in Proto-Baltic/Slavic.

11.4 Prospects for future research

* General noun phrase structure The focus of the present study lies on noun phrases with adjectival modifiers, but taking a look at noun phrases with other

modifiers (using, for instance, the AUTOTYP database of **AUTOTYP-NP**) suggests that the central morpho-syntactic parameters for the typologization of adjective attribution marking (i.e., *source, pattern* and *locus*, see above) can be applied to a syntactic description of noun phrase structure in general. However, a systematic description of general noun phrase structure, including noun phrases with all possible kinds of adnominal modifiers (demonstratives, numerals, relative clauses, etc.) and performed on a world-wide sample of languages will most likely reveal several new noun phrase types and morpho-syntactic parameters. To illustrate this, one new parameter will be described below.

In AUTOTYP, several languages are coded in which the head-dependent relation in noun phrases has shifted in the sense that the semantic dependent shares at least some of the syntactic properties of the head. This resembles the type of modifier-headed possessor agreement found in Oroch or Saliba adjectives described in this study (cf. also **malchukov2000** for a typology of "dependency reversal in noun-attributive constructions" and **ross1998** who surveyed this type in Oceanic languages). Another prototypical example of such a modifier-headed noun phrase is found in Wari'.

(10) Wari' (Chapacura-Wanham; **everett-etal1997**)
mam mao 'in-on ca mixem pucun wom-u
with go:SG 1SG:REAL-3SG.M REAL black POSS:3SG.M cotton-POSS:1SG
'I went with my dirty clothes' (lit. 'with my cotton's blackness')

In the ontology presented in the present study, modifier-headed possessor agreement has been described as a device which is assigned by dependent-driven agreement (i.e., cross-referencing possessor agreement) and which is also phrasally embedded (because the attribute takes the slot of the possessed noun phrase). The shifted head-dependent relation, however, was not included as a parameter in the ontological cross-classification because modifier-headed possessor agreement was the only type of modifier-headed noun phrases relevant for adjective attribution marking.

The shifted head-dependent relation, however, can be relevant for the typologization of general noun phrase structure. In fact, several different types of modifier-headed noun phrases are attested with other kinds of modifiers, for instance in Russian and several other European languages in which numerals higher than one require special case marking on the head noun.

(11) Russian (personal knowledge)

a. *tri* *mal'čik-a*
three boy-GEN.SG.M
'three boys'

b. *kniga mal'čik-a*
book boy-GEN.SG.M
'the boy's book'

The noun 'boy' in the Russian construction with the numeral 'three' is marked with genitive case (**??**). Consequently, this construction is syntactically equivalent to the genitive marked possessive noun phrase (**??**). The use of the (dependent marking) possessor case in noun phrases with numeral modifiers suggests that the numeral is the syntactic head and the noun is the modifier. Since agreement is not involved in the assignment of the attribution marker, the type found in Russian is clearly distinguished from the above mentioned modifier-headed possessor agreement in Wari' and should therefore be labeled *modifier-headed case* (**AUTOTYP-NP**).

* Polyfunctionality In a typological survey of noun phrase structures, all types attested in a single language have to be coded if they are distinguished by a formal characteristic, such as a distinct marker, a distinct constituent order, a general marker with a distinct function, etc. Thus, this survey automatically accounts for the polyfunctionality of attribution marking if one and the same device is used with a similar function but for at least two different kinds of modifiers.

A survey of polyfunctional attribution markers in a world-wide sample of languages has already been presented by **gil2005** (see also §**??**). Gil's typology, however, is restricted to noun phrases with three different kinds of modifiers: possessor nouns, adjectives and relative clauses. A more thorough investigation of all kinds of multifunctional noun phrase markers in a restricted area (such as northern Europe) could trace the sub-areal distributions of various multifunctional types across certain genera. Together with a description of known evolutionary paths of attribution marking, such a survey would also help to develop a theory that accounts for polyfunctionality from both a diachronic and a synchronic perspective.

Language sample and maps

1 Genus abbreviations (families)

AB-AD=Abkhaz-Adyghe, AUA=Austroasiatic, AUN=Austronesian, C-SUD=Central Sudanic, CHAD=Chadic, CHAP=Chapacura-Wanham, CHU=Chukotkan, CUSH=Cushitic, DRAV=Dravidian, ESK-A=Eskimo-Aleut, GUNW=Gunwingguan, HM-MI=Hmong-Mien, IE=Indo-European, IROQ=Iroquoian, KAMCH=Kamchatkan, KARTV=Kartvelian, KOIS=Koisan, KOM=Kombio, K-KRO=Kadugli-Krongo, MONG=Mongolic, MUSK=Muskogean, NA-DA=Nakh-Daghestanian, NA-DE=Na-Dene, NIG-C=Niger-Congo, NIL=Nilotic, S-BOU=South Bougainville, SE-RA=Lower Sepic-Ramu, SEM=Semitic, SIN-T=Sino-Tibetan, SONG=Songhai, TAI-K=Thai-Kadai, TANG=Tangkic, TNG=Trans New Guinea, TUNG=Tungusic, TURK=Turkic, U-AZT=Uto-Aztecan, URAL=Uralic, YEN=Yeniseian, YUK=Yukaghir

2 Genus abbreviations (branches and subbranches)

5N=Five Nations, AAT=Avar-Andi-Tsezic, ABKH=Abkhaz, ADYG=Adyghe, ALBA=Albanian, ALT=Altay, ARAM=Aramaic, ARAP=Arapesh, ARME=Armenian, ATHA=Athabaskan, ATLA=Atlantic, BALT=Baltic, BANT=Bantoid, BE-CO=Benue-Congo, BRIT=Brittonic, BULG=Bulgar, BURM=Burmic, CELT=Celtic, CH-IN=Chechen-Ingush, CHIN=Chinese, CHUK=Chukchi, COM=Common, DARG=Dargwa, ENE=Enets, ENIN=Enindhilyagwa, ESKI=Eskimo, DAGH=Daghestanian, DAG=Dagur, FINN=Finnic, FOR=Formosan, GAE=Gaelic, GEOR=Georgian, GER=Germanic, GREE=Greek, HAUS=Hausa, HELL=Hellenic, HMON=Hmongic, HUNG=Hungarian, I-ARY=Indo-Aryan, I-IRA=Indo-Iranian, IRAN=Iranian, IT-W=Italo-Western, KARL=Karluk, KHAN=Khanty, KHOE=Khoekhoe, KIPCH=Kipchak, KIRA=Kiranti, KORAL=Koryak-Alutor, KRON=Krongo, L-BUR=Lolo-Burmese, L-SEP-Lower Sepik, LEZG=Lezgic, LEND=Lendu, M-KH=Mon-Khmer, MADA=Madang, MAL-P=Malayo-Polynesian, MANCH=Manchu, MAND=Mande, MANS=Mansi, MOGH=Moghol,

MONGO=Mongolian, MONGU=Monguor, MORD=Mordvin, NASI=Nasioi, NOU=Nanay-Orok-Ulcha, NENE=Nenets, NGAN=Nganasan, OCE=Oceanic, OR-UD=Oroch-Udege, OROM=Oromo, PERM=Permic, REMB=Rembargic, ROM=Romance, S-WEL=South Wellesley, SAAM=Saamic, SAMO=Samoyedic, SAY=Sayan, SELK=Selkup, SIN=Sinitic, SLAV=Slavic, SUND=Sundic, TSO=Tsouic, VIET=Vietic, W-MP=Western Malayo-Polynesian, YEN=Yenisey, YI-KA=Yimas-Karawari, YOR=Yoruboid, YUP=Yupik

3 Geographic (sample) abbreviations

EU=Europe, NA=North Asia, NE=North Eurasia, W=World

4 Type abbreviations

ACAgr=Anti-construct state agreement, AConstr=Anti-construct state, AHDAgr=Appositional head-driven agreement, Constr=Construct state, DConstr=Double-construct state, HDAgr=Head-driven agreement, Inc=Adjective incorporation, Juxt=Juxtaposition, Link=Linker, MHPAgr=Modifier-headed possessor agreement, Nmlz=Attributive nominalization

Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
(Pidgin and creoles)			Berbice Dutch Creole	–	–	–	X	Juxt	**kouwenberg1994**
AB-AD	ABKH		Abaza	X	–	–	–	HDAgr	**lomtatidze-etal1989**
AB-AD	ABKH		Abkhaz	X	–	X	X	HDAgr	**chirikba2003**
AB-AD	ADYG		Adyghe (Abzakh)	X	–	–	–	Inc	**paris1989**
AB-AD	ADYG		Karbardian	X	–	X	–	Inc	**colarusso2006**
AINU (isolate)			Ainu	–	X	X	X	Juxt	**refsing1986**
AUA	M-KH	VIET	Vietnamese	–	–	–	X	Juxt	**nguyen1987**
AUN	FOR	TSO	Tsou	–	–	–	X	AConstr	**szakos1994**
AUN	MAL-P	OCE	Saliba	–	–	–	X	MHPAgr	**mosel1994**
AUN	MAL-P	OCE	Takia	–	–	–	X	AConstr	**ross1998**
AUN	WEMP	M-PH	Tagalog	–	–	–	X	Linker	**schachter1987**
AUN	WEMP	SUND	Minangkabau	–	–	–	–	Juxt	**gil2005**
BASQUE (isolate)			Basque	X	–	X	X	Juxt	**hualde-etal2003**
C-SUD	E	LEND	Ngiti	–	–	–	X	HDAgr/Juxt	**kutsch-lojenga1994**
CARIBAN			Hixkaryana	–	–	–	X	n.a.	**derbyshire1979**
CHAD	W	HAUS	Hausa	–	–	–	X	ACAgr/HDAgr	**wolff1993**
CHAP			Wari'	–	–	–	X	MHPAgr	**everett-etal1997**
CHU	CHUK		Chukchi	–	X	X	X	Inc/HDAgr	**skorik1960**
CHU	KORAL		Alutor	–	X	–	–	Inc/HDAgr	**nagayama2003**
CHU	KORAL		Koryak	–	X	X	–	Inc	**zukova1997**
CUSH	E	OROM	Oromo (Boraana)	–	–	–	X	HDAgr	**stroomer1995**
DRAV	S		Tamil	–	–	–	X	Juxt	**asher1982**
ESK-A	ESKI	YUP	Yupik (Siberian)	–	X	X	–	Inc	**de-reuse1994**
ESK-A	ESKI		Western Greenlandic	–	–	–	X	HDAgr	**fortescue1984**
GUNW	REMB		Ngalakan	–	–	–	X	HDAgr/Juxt	**merlan1983**
HM-MI	HMON		Hmong Njua	–	–	–	X	Juxt	**harriehausen1990**
IE	ALBA		Albanian	X	–	X	X	HDAgr/Nmlz+HDAgr	**demiraj1998**
IE	ALBA		Arvanitika	X	–	–	–	HDAgr/Nmlz+HDAgr	**sasse1991**
IE	ARME		Armenian (Eastern)	X	–	X	–	HDAgr/Juxt	**ajello1998**
IE	BALT	E	Latvian	X	–	X	–	ACAgr/HDAgr	**nau1996**

Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
IE	BALT	E	Lithuanian	X	–	–	–	ACAgr/HDAgr	**press2005**
IE	CELT	BRIT	Breton	X	–	–	–	HDAgr	**ternes1992**
IE	CELT	BRIT	Cornish	X	–	–	–	HDAgr	**thomas1992b**
IE	CELT	BRIT	Welsh	X	–	X	–	HDAgr	**thomas1992a**
IE	CELT	GAE	Gaelic (Scots)	X	–	X	–	HDAgr	**macauley1992**
IE	CELT	GAE	Irish	X	–	–	–	HDAgr	**odochartaigh1992**
IE	CELT	GAE	Manx	X	–	X	–	HDAgr/Juxt	**phillips2004**
IE	GER	N	Danish	X	–	–	–	HDAgr[Nmlz]	own knowledge
IE	GER	N	Danish (W-Jutlandic)	–	–	–	–	HDAgr	**lund1932**
IE	GER	N	Faroese	X	–	–	–	ACAgr+HDAgr/HDAgr[Nmlz]	**lockwood1955**
IE	GER	N	Icelandic	X	–	X	–	HDAgr[Nmlz]	**kress1982**
IE	GER	N	Norwegian	X	–	–	–	ACAgr+HDAgr/HDAgr[Nmlz]	own knowledge
IE	GER	N	Swedish	X	–	X	–	ACAgr+HDAgr/HDAgr[Nmlz]	own knowledge
IE	GER	N	Swedish (Västerbotten)	X	–	X	–	Inc/HDAgr	**astrom1893**
IE	GER	W	Dutch	X	–	–	–	ACAgr[Nmlz]	**donaldson1997**
IE	GER	W	English	X	–	X	X	Inc[Nmlz]	own knowledge
IE	GER	W	Frisian (West)	X	–	–	–	ACAgr	**tiersma1985**
IE	GER	W	German	X	–	X	–	ACAgr[Nmlz]	own knowledge
IE	GER	W	German (Alemannic)	X	–	–	–	ACAgr	**reese2006**
IE	GER	W	German (Low)	X	–	–	–	ACAgr	**matras-etal2003**
IE	GER	W	Luxembourgeois	X	–	–	–	ACAgr	**schanen-etal2006**
IE	GER	W	Yiddish (East)	X	–	–	–	ACAgr(Nmlz+HDAgr)	**katz1987**
IE	HELL	GREE	Greek	X	–	X	X	HDAgr(Nmlz+HDAgr)	**ruge1986**
IE	I-IRA	I-ARY	Palula	–	X	X	–	HDAgr[Juxt]	**liljegren2016a**
IE	I-IRA	I-ARY	Parya	–	X	X	–	ACAgr/Juxt	**oranskaja2001**
IE	I-IRA	I-ARY	Romani (Burgenland)	X	–	X	–	HDAgr(Nmlz+HDAgr)	**halwachs-etal2002**
IE	I-IRA	I-ARY	Romani (Doplenjska)	X	–	–	–	HDAgr	**cech2006**
IE	I-IRA	I-ARY	Romani (Lithuanian)	X	–	–	–	HDAgr	**tenser2005**
IE	I-IRA	I-ARY	Romani (Sepecides)	X	–	–	–	HDAgr	**cech-etal2003**
IE	I-IRA	I-ARY	Romani (Sinte)	X	–	–	–	HDAgr	**holzinger1995**

Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
IE	I-IRA	IRAN	Kurdish (Northern)	X	–	–	–	Constr	**aygen2007**
IE	I-IRA	IRAN	Ossetic	X	–	X	–	Juxt(Constr)	**abaev1964**
IE	I-IRA	IRAN	Persian	–	–	–	X	Constr	**mahootian1997**
IE	I-IRA	IRAN	Rushani	–	X	–	–	AConstr(Constr)	**payne1989**
IE	I-IRA	IRAN	Shughni	–	X	–	–	HDAgr	**payne1989**
IE	I-IRA	IRAN	Tajik	–	X	X	–	Constr	**ido2005**
IE	I-IRA	IRAN	Talysh (Northern)	X	–	X	–	AConstr	**schulze2000**
IE	I-IRA	IRAN	Tati	X	–	–	–	Constr	**dzidalaev2000**
IE	I-IRA	IRAN	Yazghulami	–	X	X	–	HDAgr	**payne1989**
IE	ROM	E	Rumanian	X	–	X	–	HDAgr(Nmlz+HDAgr)	**beyer-etal1987**
IE	ROM	IT-W	French	X	–	–	–	HDAgr	**harris1997**
IE	ROM	IT-W	Galician	X	–	–	–	HDAgr	**perez-bouza1996**
IE	ROM	IT-W	Italian	X	–	X	X	HDAgr	**maiden-etal2000**
IE	ROM	IT-W	Portuguese	X	–	–	–	HDAgr	**gartner1998**
IE	ROM	IT-W	Romansch	X	–	–	–	HDAgr	**haiman1997**
IE	ROM	IT-W	Spanish (Castilian)	X	–	–	–	HDAgr	**torrego1998**
IE	ROM	IT-W	Spanish (Catalan)	X	–	–	–	HDAgr	**hualde1992**
IE	ROM	S	Corsican	X	–	–	–	HDAgr	**giacomo-marcellesi1997**
IE	ROM	S	Sardinian	X	–	X	–	HDAgr	**jones1997**
IE	SLAV	E	Belorussian	X	–	–	–	HDAgr	**mayo1993**
IE	SLAV	E	Russian	X	–	X	X	ACAgr	own knowledge
IE	SLAV	E	Ukrainian	X	–	X	–	HDAgr	**shevelov1993**
IE	SLAV	S	Bulgarian	X	–	X	–	HDAgr	own knowledge
IE	SLAV	S	Macedonian	X	–	–	–	HDAgr	**friedman2002**
IE	SLAV	S	Serbo-Croatian	X	–	–	–	HDAgr(ACAgr)	**kordic1997**
IE	SLAV	S	Slovenian	X	–	–	–	HDAgr(ACAgr)[Nmlz+HDAgr]	**priestly1993**
IE	SLAV	W	Czech	X	–	–	–	HDAgr	**janda-etal2000**
IE	SLAV	W	Kashubian	X	–	–	–	HDAgr	**stone1993b**
IE	SLAV	W	Polish	X	–	–	–	HDAgr	**feldstein-etal2002**
IE	SLAV	W	Slovak	X	–	–	–	HDAgr	**short1993b**

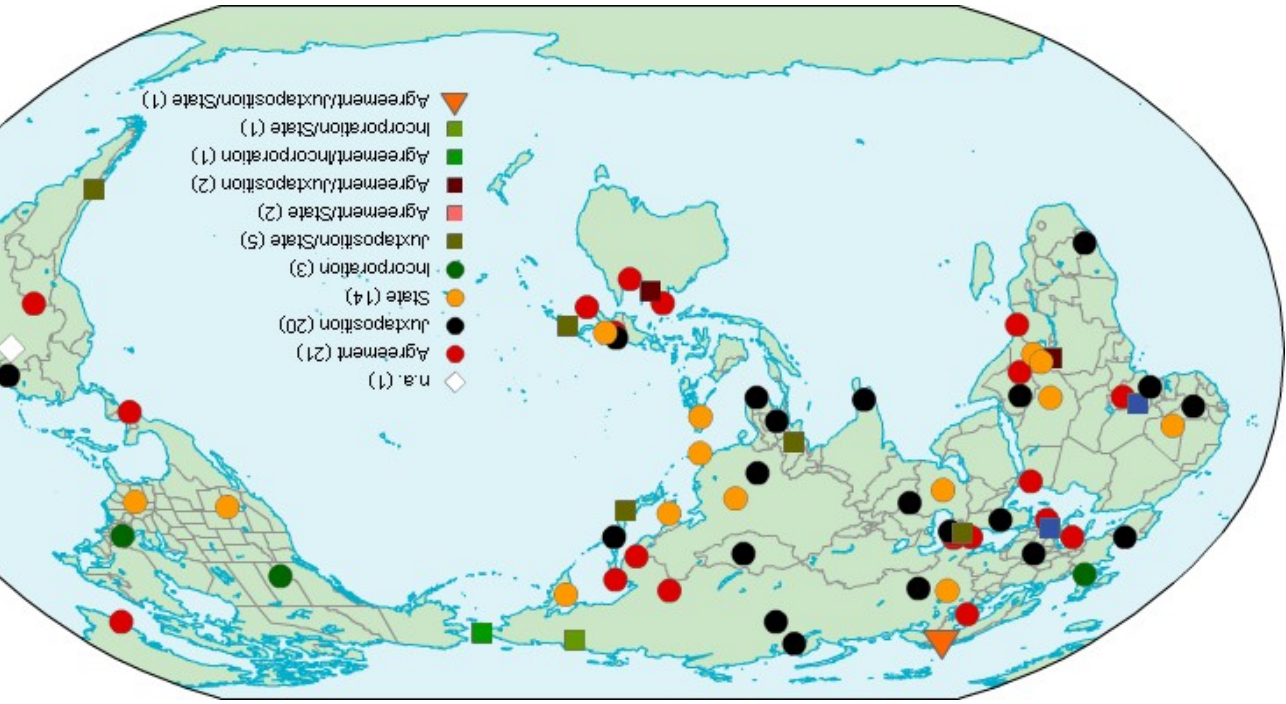
Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
IE	SLAV	W	Sorbian (Lower)	X	–	X	–	HDAgr	**stone1993a**
IE	SLAV	W	Sorbian (Upper)	X	–	–	–	HDAgr	**schaarschmidt2004**
IROQ	N	5N	Cayuga	–	–	–	X	Inc	**mithun-etal1982**
JAPANESE (isolate)			Japanese	–	X	X	X	AConstr/Juxt	**backhouse1984**
K-KRO	KRON		Krongo	–	–	–	X	ACAgr	**reh1985**
KAMCH	W		Itelmen	–	X	X	X	ACAgr	**georg-etal1999**
KARTV	GEOR		Georgian	X	–	X	X	HDAgr/Juxt(AHDAgr)	**cherchi1999**
KARTV	SVAN		Svan	X	–	X	–	HDAgr[Juxt]	**schmidt1991**
KARTV	ZAN		Laz	X	–	X	–	Juxt(HDAgr)	**holisky1991**
KARTV	ZAN		Mingrelian	X	–	–	–	Juxt(HDAgr)	**harris1991b**
KOIS	C	KHOE	Nama	–	–	–	X	Juxt	**hagman1977**
KOM	ARAP		Arapesh (Bukiyip)	–	–	–	X	Juxt	**conrad1991**
KOREAN (isolate)			Korean	–	X	X	X	AConstr	**martin-etal1969**
MAPUDUNGUN (isolate)			Mapudungun	–	–	–	X	ACAgr/Juxt	**zuniga2000**
MONG	DAG		Dagur	–	X	X	–	Juxt	**tsumagari2003**
MONG	MOGH		Moghol	–	X	–	X	Juxt	**weiers2003**
MONG	MONGO		Buryat	–	X	–	–	Juxt	**skribnik2003**
MONG	MONGO		Kalmyk	–	X	–	–	Juxt	**blasing2003**
MONG	MONGO		Khalkha	–	X	X	X	Juxt	**svantesson2003**
MONG	MONGO		Mongol (Khamnigan)	–	X	–	–	Juxt	**janhunen2005**
MONG	MONGO		Oyrat	–	–	–	–	Juxt	**birtalan2003**
MONG	MONGU		Mangghuer	–	–	–	–	Juxt	**slater2003**
MUSK	E		Koasati	–	–	–	X	Nmlz	**kimball1991**
NA-DA	DAGH	AAT	Akhvakh	X	–	–	–	HDAgr	**magomedbekova2000**
NA-DA	DAGH	AAT	Andi	X	–	–	–	HDAgr	**saidova2000**
NA-DA	DAGH	AAT	Avar	X	–	–	X	HDAgr(Juxt)	**alekseev-etal1997**
NA-DA	DAGH	AAT	Bagvalal	X	–	–	–	HDAgr	**magomedova2000a**
NA-DA	DAGH	AAT	Bezhta	X	–	–	–	HDAgr	**KibrikEtAl2004**
NA-DA	DAGH	AAT	Botlikh	–	–	–	–	HDAgr	**azaev2000**
NA-DA	DAGH	AAT	Chamalal	X	–	–	–	HDAgr	**magomedova2004**
NA-DA	DAGH	AAT	Godoberi	X	–	X	–	HDAgr	**saidova2004**

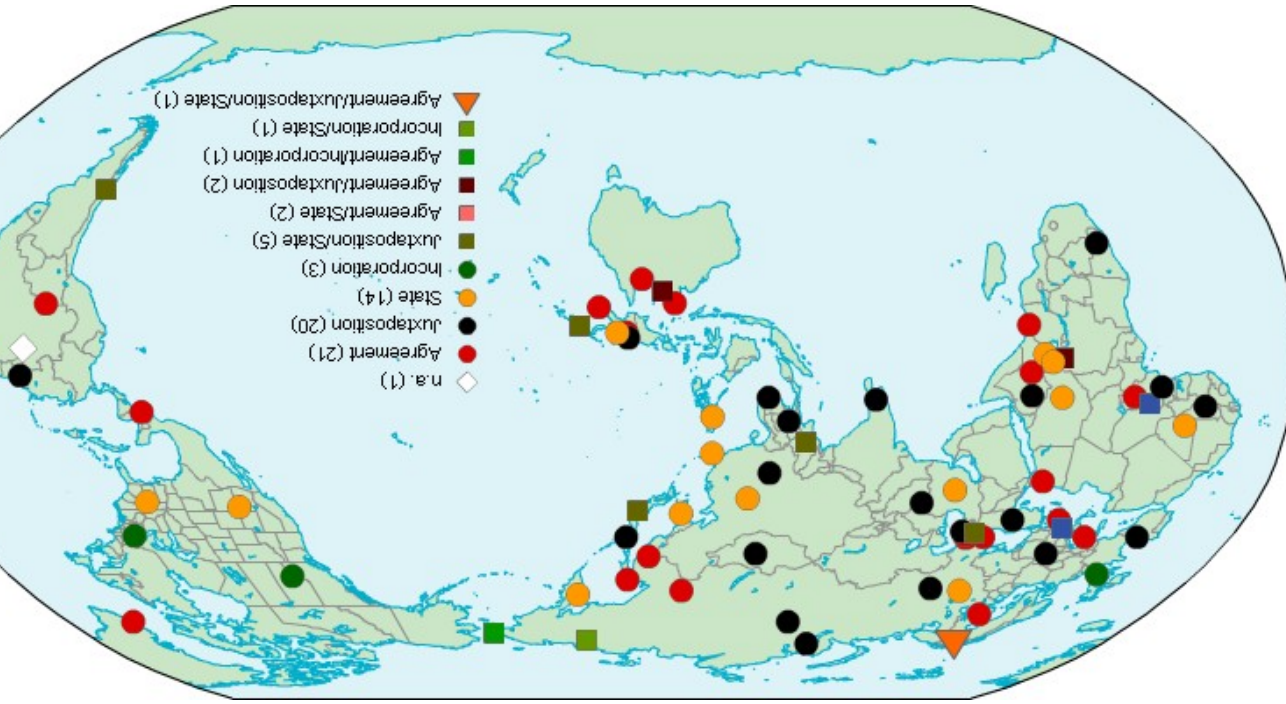
Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
NA-DA	DAGH	AAT	Hinukh	X	–	–	–	HDAgr	**isakov-etal2004**
NA-DA	DAGH	AAT	Hunzib	X	–	–	–	HDAgr	**van-den-berg1995**
NA-DA	DAGH	AAT	Karata	X	–	–	–	HDAgr	**magomedbekova1971**
NA-DA	DAGH	AAT	Tindi	X	–	–	–	HDAgr	**magomedova2000b**
NA-DA	DAGH	AAT	Tsez	X	–	X	–	ACAgr/Juxt(Nmlz)	**alekseev-etal2004**
NA-DA	DAGH	DARG	Dargwa	X	–	X	–	ACAgr(Juxt)	**isaev2004**
NA-DA	DAGH	LAK	Lak	X	–	X	X	HDAgr(ACAgr)	**abdullaev2000**
NA-DA	DAGH	LEZG	Agul	X	–	–	–	Juxt	**sauman1941**
NA-DA	DAGH	LEZG	Archi	X	–	X	–	HDAgr	**kibrik1994a**
NA-DA	DAGH	LEZG	Budukh	X	–	–	–	Juxt	**alekseev1994b**
NA-DA	DAGH	LEZG	Khinalug	X	–	–	–	Juxt	**deseriev1959**
NA-DA	DAGH	LEZG	Kryz	X	–	–	–	Juxt	**saadiev1994**
NA-DA	DAGH	LEZG	Lezgian	X	–	X	X	Juxt	**haspelmath1993**
NA-DA	DAGH	LEZG	Rutul	X	–	X	–	AConstr	**alekseev1994a**
NA-DA	DAGH	LEZG	Tabasaran	X	–	X	–	HDAgr/Juxt	**kurbanov1986**
NA-DA	DAGH	LEZG	Tsakhur	X	–	X	–	ACAgr/Juxt	**schulze1997**
NA-DA	DAGH	LEZG	Udi	X	–	–	–	Juxt	**schulze-furhoff1994**
NA-DA	NAGH	BATS	Bats	X	–	X	–	HDAgr	**holisky-etal1994**
NA-DA	NAGH	CH-IN	Chechen	X	–	X	X	HDAgr	**nichols1994a**
NA-DA	NAGH	CH-IN	Ingush	X	–	–	–	HDAgr	**nichols1994b**
NA-DE	ATHA		Sarcee	–	–	–	X	Inc	**cook1984**
NIG-C	ATLA	N	Fula (Gombe)	–	–	–	X	HDAgr	**arnott1970**
NIG-C	ATLA	S	Kisi	–	–	–	–	HDAgr	**tucker1995**
NIG-C	BE-CO	BANT	Babungo	–	–	–	–	HDAgr	**schaub1985**
NIG-C	BE-CO	BANT	Sesotho	–	–	–	–	HDAgr	**guma1971**
NIG-C	BE-CO	BANT	Swahili	–	–	–	X	HDAgr	**gromova-etal1995**
NIG-C	BE-CO	YOR	Yoruba	–	–	–	X	Juxt	**bamgbose1966**
NIG-C	MAND	W	Bambara	–	–	–	X	Juxt	**brauner1974**
NIL	S		Endo	–	–	–	X	ACAgr+Nmlz	**zwarts2003**
NIL	W	LWO	Lango	–	–	–	X	Nmlz	**noonan1992**

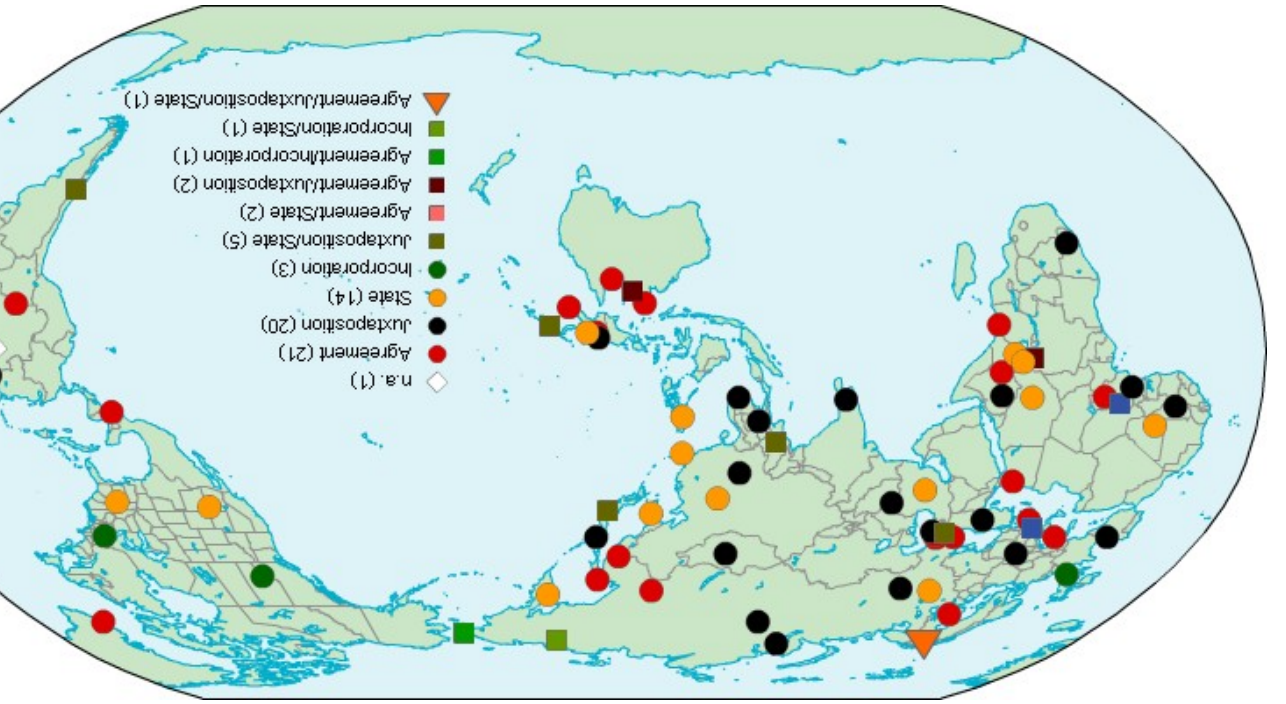
Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
NIVKH (isolate)			Nivkh	–	X	X	X	HDAgr	**gruzdeva1998**
S-BOU	NASI		Nasioi	–	–	–	X	ACAgr/Juxt	**rausch1912**
SE-RA	L-SEP	YI-KA	Yimas	–	–	–	X	HDAgr	**foley1991**
SEM	C	ARAM	Neo-Aramaic	X	–	X	–	HDAgr	**krotkoff1982**
SEM	W	C	Arabic (Cypriot)	X	–	–	–	HDAgr	**borg1985**
SEM	W	C	Arabic (Egyptian)	–	–	–	X	HDAgr	**gary-etal1982**
SEM	W	C	Maltese	X	–	X	–	HDAgr	**borg-etal1996**
SEM	W	S	Amharic	–	–	–	X	Juxt(ACAgr)	**leslau1995**
SIN-T	KIRA	E	Yakkha	–	–	–	–	Nmlz[Juxt]	**schackow2015a**
SIN-T	L-BUR	BURM	Burmese	–	–	–	X	Juxt/Nmlz	**wheatley1987**
SIN-T	L-BUR	LOLO	Lahu	–	–	–	–	Nmlz	**matisoff1973**
SIN-T	SIN	CHIN	Dungan	–	X	X	–	Juxt[Nmlz]	**kalimov1968**
SIN-T	SIN	CHIN	Mandarin	–	–	–	X	Nmlz(Juxt)	**li-etal1981**
SONG	S		Koyra Chiini	–	–	–	X	AConstr	**heath1998**
TAI-K	TAI		Nung	–	–	–	–	Juxt(Nmlz)	**saul-etal1980**
TAI-K	TAI		Thai	–	–	–	X	Juxt	**hudak1987**
TANG	S	S-WEL	Kayardild	–	–	–	X	HDAgr	**evans1995**
TIWI (isolate)			Tiwi	–	–	–	X	HDAgr	**osborne1974**
TNG	MADA	–	Mauwake	–	–	–	X	Juxt	**berghall2016a**
TUNG	AMUR	NOU	Nanay	–	X	–	–	Juxt	**avrorin1968**
TUNG	AMUR	NOU	Orok	–	X	X	–	HDAgr	**petrova1967**
TUNG	AMUR	NOU	Ulcha	–	–	X	–	Juxt[Nmlz]	**sunik1985**
TUNG	AMUR	OR-UD	Oroch	–	X	X	–	Juxt(MHPAgr)	**avrorin-etal1967**
TUNG	AMUR	OR-UD	Udege	–	X	X	X	HDAgr/MHPAgr	**nikolaeva-etal2001**
TUNG	MANCH		Manchu	–	X	X	–	Juxt	**avrorin2000**
TUNG	N		Even	–	X	X	X	HDAgr(MHPAgr)	**malchukov1995**
TUNG	N		Evenki	–	X	X	–	HDAgr/Juxt/Nmlz(MHPAgr)	**nedjalkov1997**
TUNG	N		Negidal	–	X	X	–	Juxt	**nedalkov2001**
TUNG	N		Solon	–	X	–	–	Juxt	**cincius1997**
TURK	BULG		Chuvash	X	–	X	X	Juxt(Nmlz)	**clark1998a**
TURK	COM	ALT	Altay (Southern)	–	X	X	–	Juxt	**baskakov1997b**

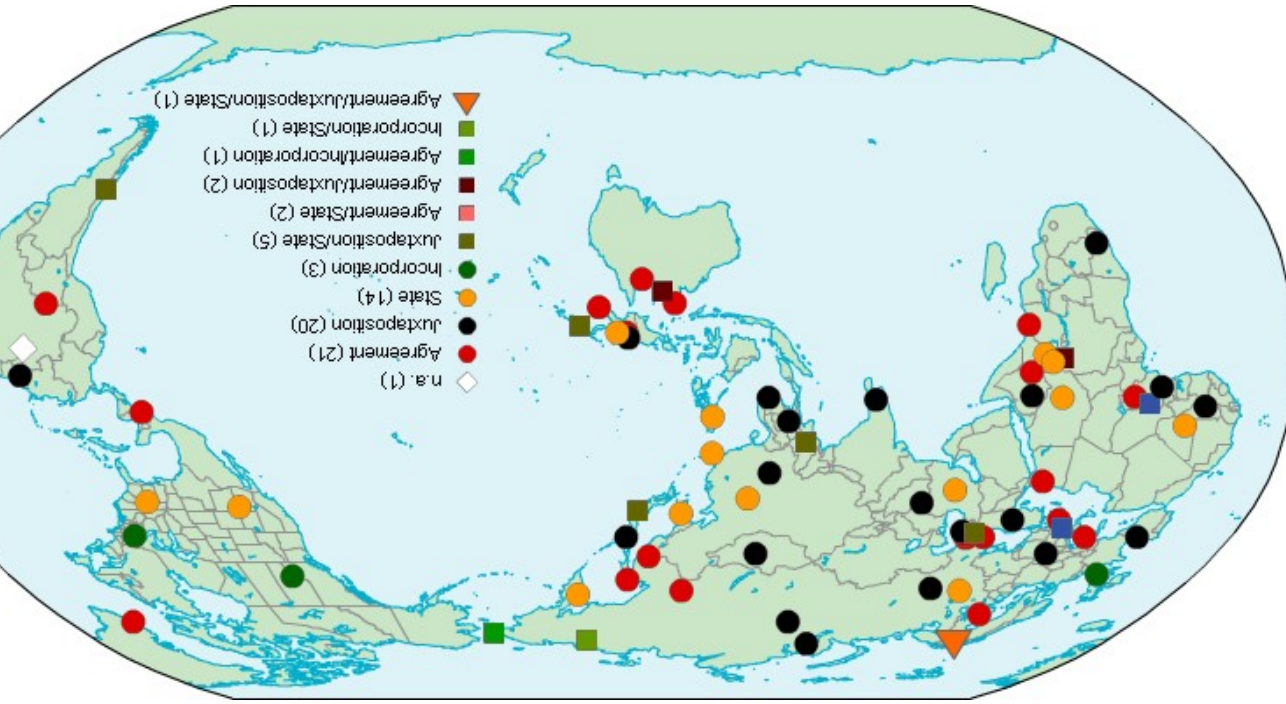
Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
TURK	COM	ALT	Kirghiz	–	X	–	–	Juxt	**kara2003**
TURK	COM	KARL	Uygur	–	X	–	–	Juxt[Nmlz]	**nadzip1971**
TURK	COM	KARL	Uzbek	–	X	X	–	Juxt[Nmlz]	**boeschoten1998**
TURK	COM	KIPCH	Bashkir	–	X	–	–	Juxt	**poppe1964**
TURK	COM	KIPCH	Karachay-Balkar	X	–	–	–	Juxt	**seegmiller1996**
TURK	COM	KIPCH	Karaim	X	–	–	–	Juxt	**kocaoglu2006**
TURK	COM	KIPCH	Karakalpak	–	X	–	–	Juxt	**baskakov2001**
TURK	COM	KIPCH	Kazakh	–	X	X	–	Juxt	**kara2002**
TURK	COM	KIPCH	Kumyk	X	–	–	–	Juxt	**kadyradziev2000**
TURK	COM	KIPCH	Nogay	X	–	–	–	Juxt	**baskakov1940**
TURK	COM	KIPCH	Tatar	X	–	–	–	Juxt	**poppe1963**
TURK	COM	LENA	Dolgan	–	X	–	–	Juxt	**ubratova1985**
TURK	COM	LENA	Sakha	–	X	X	–	Juxt	**krueger1962**
TURK	COM	OGUZ	Azerbaijani	–	X	–	–	Juxt	**budagova1982**
TURK	COM	OGUZ	Gagauz	X	–	–	–	Juxt	**pokrovskaja1997**
TURK	COM	OGUZ	Turkish	X	–	X	X	Juxt[Nmlz]	**kornfilt1997**
TURK	COM	OGUZ	Turkmen	–	X	–	–	Juxt	**clark1998b**
TURK	COM	SAY	Tuvan	–	X	X	–	Juxt	**anderson-etal1999**
TURK	COM	YEN	Khakaz	–	X	X	–	Juxt	**anderson1998**
TURK	COM	YEN	Shor	–	X	–	–	Juxt	**donidze1997**
U-AZT	N		Hopi (Toreva)	–	–	–	X	DConstr	**whorf1946**
U-AZT	S		Pipil	–	–	–	X	HDAgr	**campbell-l1985**
URAL	FINN		Estonian	X	–	–	–	HDAgr	**viitso1998**
URAL	FINN		Finnish	X	–	X	X	HDAgr	**buchholz2004**
URAL	FINN		Ingrian	X	–	–	–	HDAgr	**laanest1997**
URAL	FINN		Karelian	X	–	–	–	HDAgr	**zajkov1999**
URAL	FINN		Livonian	X	–	–	–	HDAgr	**moseley2002**
URAL	FINN		Vepsian	X	–	–	–	HDAgr	**sajceva1981**
URAL	FINN		Votian	X	–	–	–	HDAgr	**nikolaeva1997**
URAL	HUNG		Hungarian	X	–	X	X	Juxt	**kenesei-etal1998**

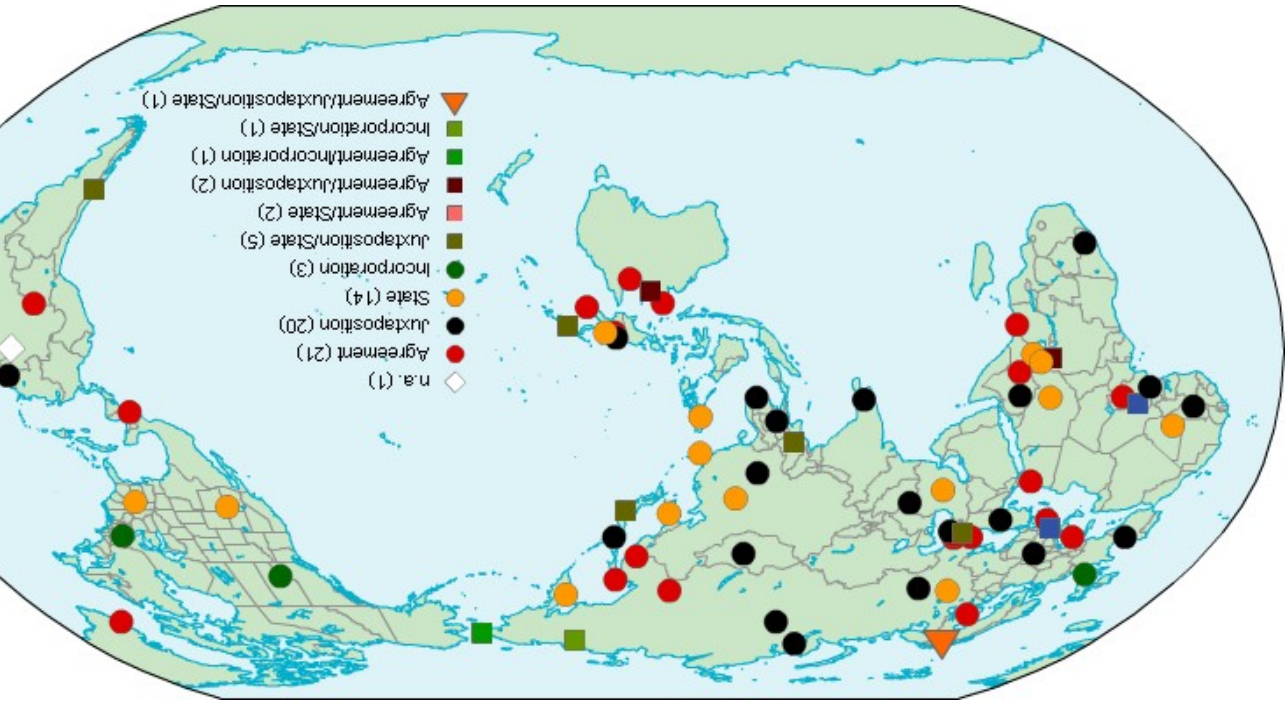
Genealogical affiliation				Geographic sampling				Noun phrase type(s)	
Family	(Sub-)Branch		Language	EU	NA	NE	W	#1/#2(#3)[#4]	Reference
URAL	KHANT		Khanty (Eastern)	–	X	X	–	Juxt	**nikolaeva1999**
URAL	KHANT		Khanty (Northern)	–	X	X	–	Juxt	**nikolaeva1999**
URAL	MANS		Mansi (Eastern)	–	X	X	–	Juxt	**riese2001**
URAL	MANS		Mansi (Northern)	–	X	X	–	Juxt	**riese2001**
URAL	MARI		Mari (Eastern)	X	–	X	–	Juxt[Nmlz+AHDAgr]	**alhoniemi1993**
URAL	MARI		Mari (Western)	X	–	–	–	Juxt[Nmlz+AHDAgr]	**alhoniemi1993**
URAL	MORD		Mordvin (Erzya)	X	–	X	–	Juxt	**zavodova1964**
URAL	MORD		Mordvin (Moksha)	X	–	–	–	Juxt	**zavodova1964**
URAL	PERM		Komi-Permyak	X	–	–	–	Juxt(Nmlz+AHDAgr)	**lytkin1966b**
URAL	PERM		Komi-Zyrian	X	–	–	–	Juxt(Nmlz+AHDAgr)	**kokkonen1984**
URAL	PERM		Udmurt	X	–	X	–	Juxt(Nmlz+AHDAgr)[AHDAgr]	**winkler2001**
URAL	SAAM	E	Saami (Inari)	–	–	–	–	AConstr/Juxt	**olthuis2000**
URAL	SAAM	E	Saami (Kildin)	X	–	–	–	AConstr/Juxt[HDrAgr]	own knowledge
URAL	SAAM	E	Saami (Skolt)	–	–	–	–	AConstr/Juxt	**feist2015a**
URAL	SAAM	E	Saami (Ter)	–	–	–	–	AConstr/Juxt	own knowledge
URAL	SAAM	W	Saami (Lule)	–	–	–	–	AConstr/Juxt	**spiik1989**
URAL	SAAM	W	Saami (Northern)	X	–	X	X	AConstr/DConstr/HDAgr/Juxt	own knowledge
URAL	SAAM	W	Saami (Pite)	–	–	–	–	AConstr/Juxt	**wilbur2014a**
URAL	SAAM	W	Saami (Southern)	X	–	–	–	AConstr/Juxt	**bergsland1994**
URAL	SAAM	W	Saami (Ume)	–	–	–	–	AConstr/Juxt	own knowledge
URAL	SAMO	ENE	Enets (Forest)	–	X	X	–	Juxt	**siegl2013a**
URAL	SAMO	ENE	Enets (Tundra)	–	–	–	–	Juxt	**sorokina2010a**
URAL	SAMO	NENE	Nenets (Tundra)	–	X	X	X	Juxt	**salminen1998a**
URAL	SAMO	NGAN	Nganasan	–	X	X	–	Juxt	**wagner-nagy2002a**
URAL	SAMO	SELK	Selkup (Central)	–	X	X	–	Juxt	**helimski1998b**
URAL	SAMO	SELK	Selkup (Northern)	–	X	X	–	Juxt	**helimski1998b**
URAL	SAMO	SELK	Selkup (Southern)	–	X	X	–	Juxt	**helimski1998b**
YEN	N		Ket	–	X	X	X	Juxt(HDAgr)[Nmlz]	**vajda2004**
YUK			Yukaghir (Kolyma)	–	X	–	–	AConstr/Inc	**maslova2003b**
YUK			Yukaghir (Tundra)	–	X	X	X	AConstr/Inc	**maslova2003a**

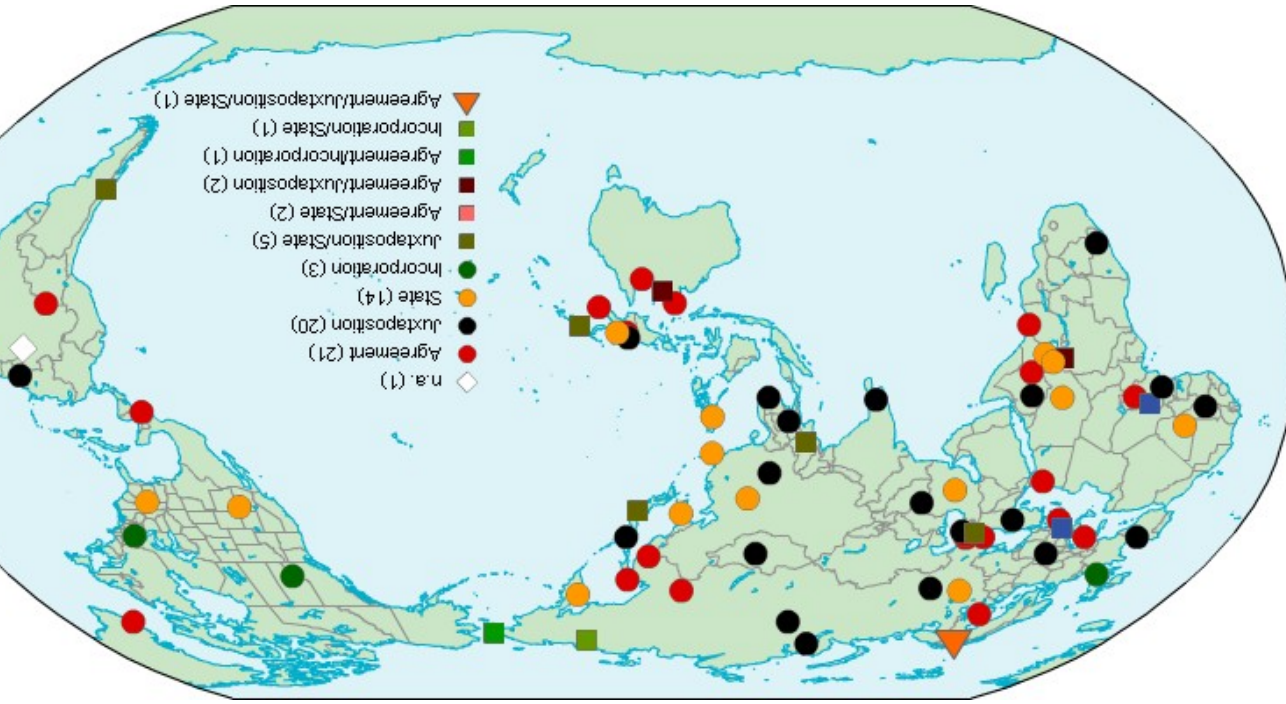

n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)

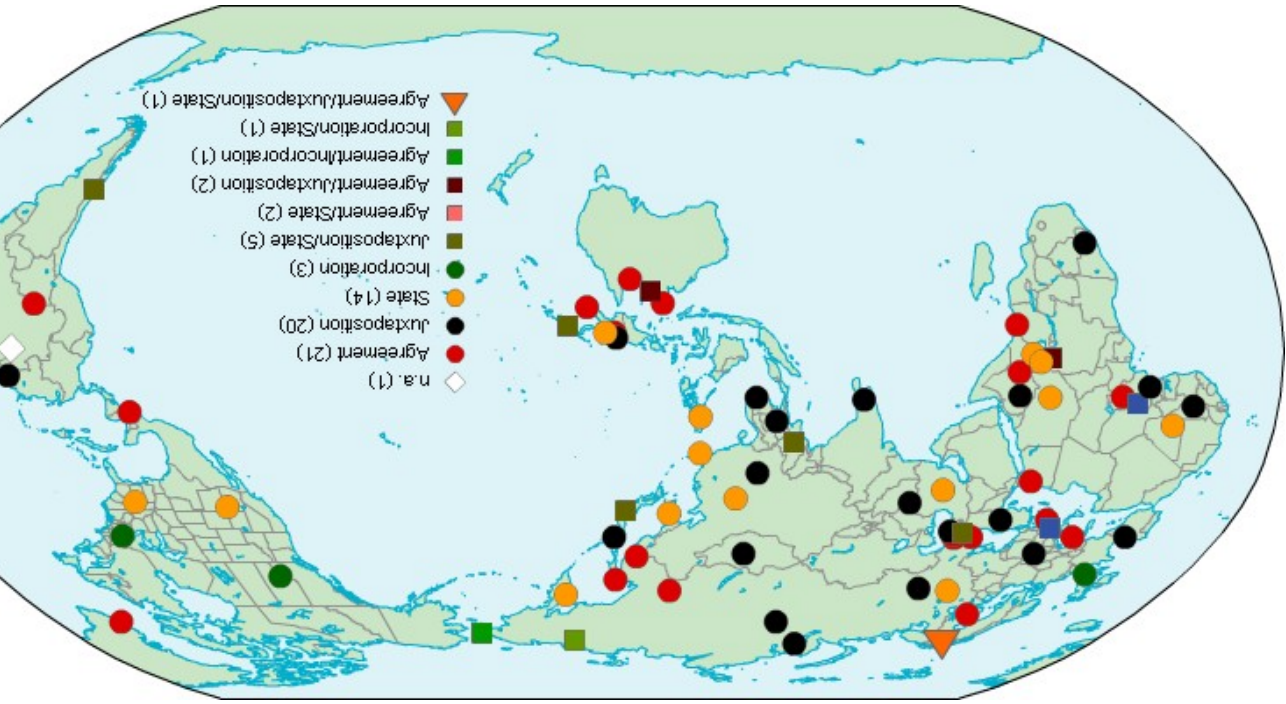

n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)

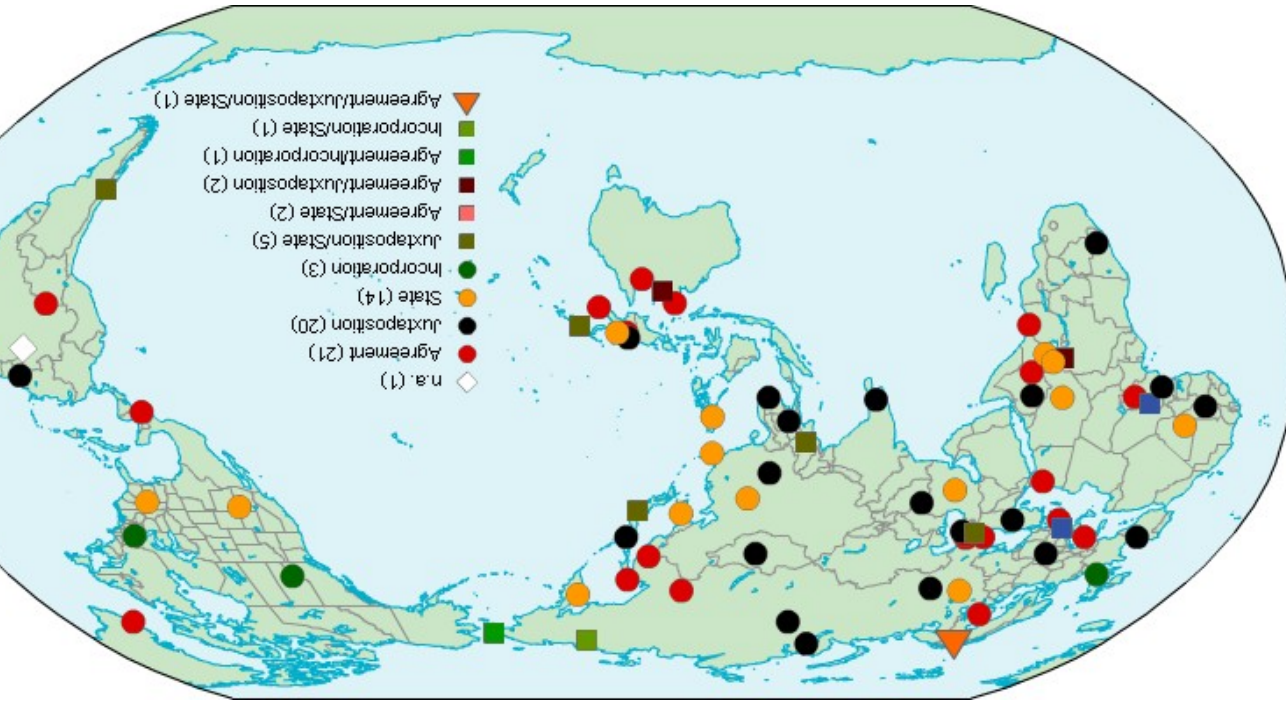

n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)


n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)


n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)


n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)


n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)


n.a. (1)
Agreement (21)
Juxtaposition (20)
State (14)
Incorporation (3)
Juxtaposition/State (5)
Agreement/State (2)
Agreement/Juxtaposition (2)
Agreement/Incorporation (1)
Incorporation/State (1)
Agreement/Juxtaposition/State (1)