

# Revising a syntactic isogloss: nominal modifiers marking in Indo-European languages

Artemij Keidan\*

## Abstract

The paper deals with a morpho-syntactic innovation shared by a group of IE languages: the grammaticalisation of the pronominal linkers in verbless relative NPs as markers of nominal dependency. Three goals are pursued: 1) defining a parametrised framework for a unified description of the highly variable outcomes that this pattern has produced in the historically attested languages (for the first time all collected together); 2) critically revising the existing partial assessments of this isogloss (starting from Benveniste 1966), in order to exclude some languages whose data are too scarce or inconsistent (e.g. Hittite, Latin), and to reassess some other data under a new viewpoint (e.g. the two competing verbless relatives in Greek); 3) suggesting some possible new members of the isogloss by applying the parametrised approach to previously unnoticed data (e.g. Khotanese, Middle Indo-Aryan, Carian). In conclusion, after a typological comparison with a parallel development in some Semitic languages, an attempt is made to describe and explain the grammatical motivations that could have initiated this phenomenon.

## 1 Defining the problem

### 1.1 *In medias res*: the data under consideration

In the present study, I am dealing with a grammaticalisation pattern shared by a compact group of IE languages. Though not unknown to the linguists, this innovation has never been considered as one single syntactic isogloss regarding the entire set of languages that I am going to present below. This pattern has to do with the marking of the constituency relationships within the NP, for instance with the marking of dependency between the head noun and its modifiers. One common feature of all these constructions is the grammaticalisation of a former pronominal linker as a “constituency marker”. I use this term to refer to any morphological means that mark the dependency relation between the head and the modifier, regardless of the locus of marking.

---

\*Sapienza University of Rome, email: [artemij.keidan@uniroma1.it](mailto:artemij.keidan@uniroma1.it), ORCID: 0000-0003-4944-0675. Published version: [doi.org/10.1163/9789004508873\\_006](https://doi.org/10.1163/9789004508873_006).



In order to stress the parallelism of these patterns within the languages under analysis, I wish to start by exposing the primary linguistic data. A more detailed discussion of each language will be presented in Section 2. For now, the data are presented in the most legible form, in simplified transcription and without being glossed over, just to show the common structure (the hyphen is added when needed for clarity; the pronominal linker is highlighted in boldface).

- (1) a. Sanskrit *áditir yá duhitá táva* ‘Adivi, the daughter of yours’  
 b. Avestan *azəm yō ahurō mazdā* ‘I, the Ahura Mazda’  
 c. Old Persian *Gaumāta haya maguš* ‘Gaumata the magian’  
 d. Parthian *mād čē dēwān* ‘the mother of demons’  
 e. Middle Persian *nōhzādag ī targumān* ‘Nōhzādag the interpreter’  
 f. Greek, I var. *Τεῦκρος δὲ ἄριστος Ἀχαιῶν* ‘Teucer, the best of the Achaeans’  
 g. Greek, II var. *Σόλων ὁ Ἀθηναῖος* ‘Solon the Athenian’  
 h. Armenian *hogi or i nma* ‘the spirit in him’  
 i. Latin *divi qui potes* ‘the mighty gods’  
 j. Hittite *memiyaš kuiš iyawaš* ‘the thing to be done’  
 k. Carian *en mwdonś ki* ‘the mother of Mwdon’  
 l. Lithuanian *aukšto-ji mokykla* ‘the high school’  
 m. Old Slavic *dobry-ję ženy* ‘the good wives’

The constructions in (1) present a high degree of similarity, notwithstanding some obvious differences. Three basic elements are to be observed: a nominal head, a linker of pronominal origin, plus a modifier, which, in its turn, can be an adjective, a nominal epithet, a modifying NP inflected in an oblique case, or a modifying PP. The most relevant element of variation among the languages presenting this pattern resides in the degree of grammaticalisation of the pronominal linker.

The structural parallelism among all these patterns is striking enough to call for an explanation, be it genetic or typological. The existing accounts of this apparent isogloss, among others Benveniste (1966) and Ivanov (1979), are variously unsatisfying, first of all, because they analyse only a subset of the languages listed in (1); more on this in §3.1. In the present paper, I wish to suggest a descriptive framework that enhances the similarities and accounts for the differences, see §??. I also argue that, with such a descriptive framework at hand, more languages could be added to those already considered in the classical studies, see §2.2. On the other hand, some of the languages that have been traditionally considered as belonging to this isogloss are expelled thereof. Some possible grammatical motivations behind this innovation are critically revised in the concluding sections of the paper.

## 1.2 Descriptive framework

The innovation described here can be thought of as a grammaticalisation process that leads from a free syntactic construction to a fixed syntactic pattern, or to a completely morphologised form. Considering that in almost all the languages under consideration the pronominal linker seems to go back to a relative pronominal stem (see §1.2.4 and §3.3 on this respect), we can consider this development a grammaticalisation of a verbless relative. It is schematically represented in (2); here, a NP with a CP as its modifier develops into a simple NP with a nominal head and a modifier linked by a dependency marker; note that the exact position, as well as the degree of syntactic autonomy of such marker, varies from language to language.

$$(2) \quad [_{NP} \text{ head } [_{CP} \text{ pronoun } [_{TP} [_{V} \emptyset [_{NP} \text{ modifier}]]]]] \rightarrow [_{NP} \text{ head marker modifier}]$$

Although this development is attested in several branches of the Indo-European family, it is only in the Iranian domain that the entire process is attested up to the complete grammaticalisation, in all the subsequent diachronic stages, from Avestan to the modern Iranian dialects (see the survey in Haider & Zwanziger 1984: §2). The other languages from our collection attest either the oldest or the latest stage of the grammaticalisation process. However, we can convincingly postulate the formula in (2) as a general rule and then analyse how the starting pattern was distorted in each involved language to the degree that we often cannot easily recognise it in the resulting constructions.

Observing the entire range of linguistic data at our disposal, we can distinguish at least the following stages of grammaticalisation, advancing from the relative proposition up to a fully morphological marker:

- a) Deletion of the relative predicate (copula) and the correlative demonstrative pronoun.
- b) Case agreement of the pronominal linker with the head noun, spreading over the entire case paradigm.
- c) Fixation of the mutual order of head, pronominal linker, and modifier.
- d) Morphologisation (total loss of syntactic autonomy).
- e) Partial demorphologisation (the newly created morpheme becomes a phrasal marker rather than a bound morpheme).

As we will see below, such a simple scale of degree of grammaticalisation is impractical in accounting for the observed variability. The grammaticalisation process involves a few independent factors, especially regarding the behaviour of the pronominal linker (its position and autonomy) and the resulting systemic changes. A multidimensional set of parameters seems to be more appropriate for describing such a manifold development.<sup>1</sup> What follows is a list of such descriptive parameters.

---

<sup>1</sup> Our approach can be considered a specific adaptation of the broader notion of “grammaticalisation cline” as defined in Hopper & Traugott (2003).

### 1.2.1 Agglutination status

First, we should consider the agglutination status of the pronominal linker with respect to its supporting word, be it the head or the modifier (see §1.2.2). Three degrees of agglutination are to be registered:

- a) free word;
- b) clitic;
- c) bound morpheme.

For the sake of simplicity, the agglutination status is assigned on a common-sense ground and is not checked against any general criteria of cliticisation since such criteria (e.g. Nevis & Joseph 1993: 94) are typically targeted on living languages while here we are mostly dealing with languages only attested in written sources. Sometimes there are some graphic means in the writing system that distinguish clitics from free morphemes or bound morphemes from clitics, but, more often, there is no such distinction at all, especially in the *scriptio continua*. Intuitively, however, few would deny that the adjectival endings in Slavic and Germanic are bound, the Avestan relative pronoun is clearly a free morpheme, while the Greek article and the Persian *ezāfe* are better represented as clitics.

Note that the degree of agglutination is not strictly correlated with the date of attestation. Thus, Old Slavic adjectival endings had already agglutinated soon after the beginning of the written tradition while Old Lithuanian adjectival endings never attained complete agglutination even up to the present day. Germanic strong adjectival endings enter the written tradition as already completely bound morphemes. Greek article showed an autonomous accent in some case forms but functioned as a clitic in others.

### 1.2.2 Locus of marking

Once the pronominal linker loses its syntactic autonomy, it can be evaluated in terms of locus of marking, while at stages when it is still free (as in the case of the oldest IE languages considered in Section 2), this value can only be established conventionally, depending on how we represent the internal structure of the NP. We have the following options:

- a) Dependent-marking, as in the case there is some specific morphology and/or syntactic behaviour (agreement) reserved to the modifiers of the head noun (e.g. the special adjectival endings in Slavic).
- b) Head-marking, when a special mark pre-signals that a given head has a modifier after it (e.g. the *ezāfe*-marked nominal heads in New Persian).
- c) Double marking, i.e. the redundant marking of both the head and the modifier word with some special morphological device to denote the fact that they both belong to the same constituent (one possible example is the redundant article in Greek).

- d) Alternant marking, when the locus of marking is not fixed and can vary depending on some contextual or structural conditions (a partial example is the Modern German NP marking, with the “strong” ending moving across the NP; possible new examples are discussed in §2.2).<sup>2</sup>

### 1.2.3 Position of the mark

Independently from the locus of the marking — which is an abstract notion — the linear position of the mark must also be taken into consideration. Two possibilities are observable:

- a) preposed mark (either a prefix or a proclitic);
- b) postposed mark (either a suffix or a “postclitic”).

The reference is, in both cases, to the lexeme which undergoes the marking, regardless whether it is the head or the modifier. Note that the linear fixation of the linker — in case the general word order is still relatively free — may lead to a major disruption in the initial linear order of the three elements represented on the diagram (2).

### 1.2.4 The source of the mark

One last parameter regards the etymology, i.e. the PIE source, of the pronominal linker undergoing the grammaticalisation. By etymology I mean the formal reconstruction of the antecedent, regardless its original relative or demonstrative semantics. Indeed, the distinction between relative and demonstrative pronouns is not always clear even in the attested languages, let alone the protolanguage. This will be debated in §3.3, and for now, a simple list of PIE source stems with an approximate semantic characterisation will suffice; for more etymological details see Dunkel (2014).

- a) PIE *\*iō-*, together with its athematic counterpart *i-*: these two closely related stems are attested as relative and/or demonstrative pronouns in many IE languages, which makes the original semantics uncertain, even if usually considered demonstrative (see Luján 2009: 223); the attributive pronominal linkers in Vedic, Avestan and Balto-Slavic (indirectly also in Old Persian and possibly in Armenian) originate here.
- b) PIE *\*so*, used as the pronominal linker in Greek, and indirectly, in Old Persian; original semantics commonly considered demonstrative (but see Hirt 1934: 163).
- c) PIE *\*k<sup>w</sup>i-/k<sup>w</sup>o-*, used as pronominal linker in Hittite, Carian, Parthian, and plausibly in Armenian; original semantics is debated between interrogative, relative and indefinite (Luján 2009: see) who concludes that the interrogative function was primary).

---

<sup>2</sup> The alternant marking type is not mentioned by Nichols (1986), yet seems a natural enlargement of her classification.

- d) PIE *\*e-/o-*, a difficult to detect pronominal stem with possibly a demonstrative meaning (see Dunkel 2014: 183; Hirt 1934: 162); in its inflected word-forms it can be easily confused with the pronominal endings as such; it can be hypothesised as a pronominal linker for Germanic and perhaps in Prakrit.

## 2 Data

In the following paragraphs, specific forms in each language sharing this innovation are surveyed. Due to the lack of space, I will not go into details of the peculiarities of each language, and will present just the examples that best fit our assumption. The glossing of the examples is limited to the relevant information, i.e. to the nominal inflectional categories, sometimes to the case only. For each language, an assessment of the parameters defined in §1.2 is provided.

### 2.1 Already known data

#### 2.1.1 Vedic

The pronominal stem *ya-* is attested in Vedic as a relative pronoun or a sort of “determiner”. Verbless relative modifiers of nominal heads governed by *ya* are attested in Late Vedic verses and prose. Some examples are shown in (3):

- (3) a. *áditir* [...] *yá* *duhitá* *táva*  
 Aditi-NOM [...] LNK-NOM daughter-NOM you-GEN  
 ‘Adivi, the daughter of yours’ (RV 10.72.5)
- b. *paúruṣeyaṃ vadhám yám*  
 human-ACC death-ACC LNK-ACC  
 ‘death caused by men’ (AV 19.20.1)
- c. *sá rátrī páritakmyā yá*  
 that-NOM night-NOM waning-NOM LNK-NOM  
 ‘that waning night’ (RV 5.30.14)

Although Gonda (1955: 22) considers the verbless use of *ya* “very common”, it is certainly quite far from being fully grammaticalised. It seems rather a fortuitous combination of some unnecessary and unrelated circumstances: the correlative pronoun may be lacking (or not, see the counterexample in 3c); the so-called *attractio relativi* (case agreement of the relative linker with the nominal head) may occur, not only in verbless but also in clausal relatives. Other conditions, of those listed in §1.2, are never met. Thus, the linker is not cliticised at all, as we see from the accent it bears, while the linear word order is still totally unconstrained, notwithstanding the vague and unsubstantiated claim by Benveniste (1966: 218) who affirms that the linker tends to be located after the modifier. There is no need to go deeper into this construction since it has been sufficiently analysed in the preceding literature (starting from Delbrück 1900: §139).

All considered, we could only include Old Indian data into our isogloss as an example of the initial stage of this grammaticalisation process, with virtually no continuation in the later stages of the language.

### 2.1.2 Avestan

The widely known Avestan “article-like relative” or “quasi-article” (see Caland 1891: §§24–51; Delbrück 1900: §139; West 2011: §§237–239) is a verbless relative headed by the linker *yā*, going back to PIE *\*iō-*, which also functioned as a plain relative elsewhere in the language. As for our parameters, some more signs of grammaticalisation are detectable here, in comparison to Old Indian. The word order is quite stable (see Seiler 1960: Ch. 4); the modifier phrase is practically never separated from the linker, which is pre-posed and agrees in case. A dependent-marking strategy can be inferred, even if the incomplete cliticisation prevents us from making a definitive conclusion on such respect. As often observed (cf. Haider & Zwanziger 1984: §2.3), the variety of case-forms in which the linker could stand becomes richer through the diachrony of Avestan: from almost exclusively nominative and accusative examples in Gathic (see 4a–b), to ablative and instrumental ones in Young Avestan (see 4c).

- (4) a. *azəm yō Ahurō Mazdā*  
 I-NOM LNK-NOM Lord-NOM Mazdā-NOM  
 ‘I, the Lord Mazdā’ (Y. 19.6)
- b. *vā [...] yāng daēuuōng*  
 you-ACC [...] LNK-ACC daēvas-ACC  
 ‘you, the Daēvas’ (Y. 32.5)
- c. *hača zəmaṭ yaṭ paθanaiiā*  
 from earth-ABL LNK-ABL wide-ABL  
 ‘from the wide earth’ (Yt. 17.19)
- d. *aom stārəm yim tištrīm*  
 this-ACC star-ACC LNK -ACC  
 ‘this star, the Tištrya’ (Yt. 8.50)

This construction is optional in Avestan, and identical phrases are attested both with and without the “quasi-article”. The correlative pronoun is not necessarily dropped (see 4d). Nevertheless, we can include Avestan in the isogloss since it clearly displays the start of the grammaticalisation process that is later witnessed in subsequent stages of Iranian.

### 2.1.3 Old to Middle Persian

The Old Persian construction is quite similar to the Avestan “quasi-article” but also presents some peculiarities. The linker, written  $\langle h^a y^a \rangle$  in Achaemenid spelling, is usually interpreted as *haya-* (oblique stem *taya-*), going back to the concatenation of PIE stems *\*so* and *\*iō-* (literally,

“that, which”), see Strunk (1967), Schmitt (2014: 193), and Dunkel (2014: 318, fn. 36); *contra* Meillet (1915: §331). This clearly distinguishes it from the parallel Avestan “article-like” *yā*. As a consequence, the two constructions must be considered parallel innovations, rather than reflexes of a common Proto-Iranian antecedent. Old Persian pronominal linkers are always preposed to the modifier (varying from a nominal epithet to a possessor in the genitive). This makes it a dependent-marking construction although there is no cliticisation as can be suggested from the word separation mark always occurring after the linker. The head nouns to which the linker agrees are usually in the nominative, accusative, or genitive (see Meillet 1915: §§380–384). A couple of examples are presented in (5).

- (5) a. *Gaumāta haya maguš*  
 Gaumata-NOM LNK-NOM magian-NOM  
 ‘Gaumāta the Magian’ (DB I.65–66)
- b. *kāra haya manā avam kāram tayam vahyazdātahyā*  
 army-NOM LNK-NOM I-GEN this-ACC army-ACC LNK-ACC Vahyazdāta-GEN  
*aja*  
 defeated  
 ‘My army defeated this army of Vahyazdāta’ (DB III.45)

The situation in Middle Iranian appears already considerably innovated. The Old Persian “quasi-article” construction is continued in the Middle Persian “quasi-*ezāfe*” construction even if a direct etymological link between the two is sometimes questioned. The pronominal linker is cliticised to the head NP. It is spelled <y> with the reading *ī*, or also <yg>, read *īg*. The latter form appears to have been reinforced with a suffixal *-g* from *-ka-* (see Maggi & Orsatti 2018: §2.10.2). There is no clear differentiation in the grammatical value between the two spellings. In the older texts, this “quasi-*ezāfe*” is still found as a plain relative pronoun while nominal modifiers, on the other hand, may still appear without the *ezāfe*-linker. However, in later texts, it appears almost as grammaticalised as the New Persian *ezāfe* (see Haider & Zwanziger 1984: §3; Durkin-Meisterernst 2014: §531). See the examples in (6). Note that Middle Persian “quasi-*ezāfe*” is a clitic attached to a phrase rather than to a word (see 6b) and allows recursive nesting of one modifier into another (see 6c).

- (6) a. *āštānag ī naxwistēn*  
 threshold LNK first  
 ‘the first threshold’ (Durkin-Meisterernst 2014: ex. 45b)
- b. *rah dō ī xwar ud mäh*  
 chariot two LNK sun and moon  
 ‘the two chariots, that is, the sun and the moon’ (Durkin-Meisterernst 2014: ex. 38)
- c. *iškōhān ī ba az dānišn ī yazdān*  
 poors LNK without knowledge LNK gods  
 ‘the poor without the knowledge of gods’ (Durkin-Meisterernst 2014: ex. 45)

A similar construction is also attested in Parthian, whose pronominal linker *čē* goes back to the PIE stem *\*k<sup>w</sup>i-*; see the phrase in (7). It is less frequent than in Middle Persian, however.

- (7) *srōd čē šādīft*  
 song LNK joy  
 ‘the song of joy’ (Durkin-Meisterernst 2014: ex. 24)

Summing up, on the one hand, no common antecedent of the linker can be reconstructed for all the Iranian forms under consideration. On the other hand, it is in the Iranian domain that the longest attestation of this construction is witnessed. The complete fixation of the *ezāfe* construction is reached only in Modern Persian.<sup>3</sup> Note that Bauer (2007: 109) denies a genetic link between the New Persian *ezāfe* and the Old Persian verbless relative construction on the grounds that the former, unlike the latter, does not mark the noun for definiteness; see the rejection of this claim in §3.2.

#### 2.1.4 Greek

The situation in Greek is complicated by the fact that two different pronominal stems were competing for the same linking function: the “demonstrative” *ὁ, ἡ, τό* (from PIE *\*so, \*seh<sub>2</sub>, \*tod*), and the “relative” *ὅς, ἣ, ὅ* (from PIE *\*ǵos, \*ǵeh<sub>2</sub>, \*ǵod*). Indeed, both sets of forms are known in both meanings, both in Homeric and in Classical Greek dialects, see Probert (2015: §6.1; Chap. II). In (8), for instance, two examples of *ὁ, ἡ, τό* used as relative are shown.<sup>4</sup>

- (8) a. *ἠπέλιησεν μῦθον ὃ δὴ τετελεσμένος ἐστί*  
 kept word-ACC LNK-NOM now accomplished-NOM is  
 ‘he uttered the word that now has come to pass’ (*Il.* 1.388)
- b. *λάβε τὸν Μανδάνη ἔτεκε παῖδα*  
 take LNK-ACC Mandane-NOM bore child-ACC  
 ‘take the child whom Mandane bore’ (*Hdt.* 1.108.4)

Phonologically, the two pronouns were indistinguishable in a subset of the paradigm, with only a later differentiation in spelling (cf. accentless *ῆ* vs. accented *ῆ̃*). Moreover, the diachrony of Greek shows a tendency to increase, rather than decrease, the phonological similarity as some distinctive allomorphs — such as Homeric *τοί* and *ταί*, plurals of *ὁ* and *ἡ*, but not of *ὅς* and *ἣ* — were gradually abandoned. Semantically, the relative use of *ὁ, ἡ, τό* is usually said to imply a non-restrictive interpretation of the relative, at least in Homeric Greek, see Probert (2015: 122). Still, the qualification of relative vs. demonstrative in the two stems is often interpretational (see Humbert 1960: §54) or reflecting a later scholarly standardisation.

<sup>3</sup> As one of the anonymous reviewers correctly observes, not all the modern Iranian varieties present the same degree of demorphologisation of this construction. Thus, the Kurdish *ezāfe* morpheme still presents the gender agreement with the head noun, see Rießler (2016: §4.5.1).

<sup>4</sup> Here and below the gloss LNK is used for both pronouns.

After a period of competing distribution, one of the two pronouns, i.e. the demonstrative  $\delta$ ,  $\eta$ ,  $\tau\acute{o}$ , underwent the complete grammaticalisation as a definite article in Classic Greek (and later), while  $\delta\varsigma$ ,  $\eta$ ,  $\delta$  functioned only as a relative linker (however, the latter was completely ousted by  $\delta$ ,  $\eta$ ,  $\tau\acute{o}$  used as relative in those dialects where the article never emerged). Even in the Classical period, the two pronouns were sometimes used interchangeably, e.g. in the oppositive constructions with  $\mu\acute{\epsilon}\nu$  and  $\delta\acute{\epsilon}$ .<sup>5</sup>

As for the verbless relative construction, both pronouns were used in Homeric Greek in the attributive postnominal position. The relative  $\delta\varsigma$ ,  $\eta$ ,  $\delta$  was the marginal choice in such a context, see the phrase in (9).

- (9) Τεϋκρός    θ    δς    ἄριστος    Ἀχαιῶν    τοξοσύνη  
 Teucer-NOM and LNK-NOM best-NOM Achaeans-GEN archery-DAT  
 ‘and Teucer, best of all the Achaeans in bowmanship’ (*Il.* 13.313–4)

The demonstrative  $\delta$ ,  $\eta$ ,  $\tau\acute{o}$  predominated in the verbless relative constructions in Homeric Greek and became the only possibility in Classical and Biblical Greek, see the phrases in (10).

- (10) a. ἄνακτες    οἱ    νέοι  
 lords-NOM LNK-NOM young-NOM  
 ‘the young masters’ (*Od.* 14.61)
- b. ἰχθύσι    τοῖς    ὀλίγοις  
 fishes-DAT LNK-DAT little-DAT  
 ‘to the little fishes’ (*Od.* 12.252)
- c. Σκύθαι    οἱ    ἄριστοί    τε καὶ    πλείστοι  
 Scythians-NOM LNK-NOM best-NOM and most\_numerous-NOM  
 ‘the best and most numerous of the Scythians’ (*Hdt.* 4.20.1)
- d. Ξενίας    ὁ    Ἀρκάς    [...] καὶ    Πασίων    ὁ  
 Xenias-NOM LNK-NOM Arcadian-NOM and Pasion-NOM LNK-NOM  
 Μεγαρεύς  
 Megarian-NOM  
 ‘Xenias the Arcadian [...] and Pasion the Megarian’ (*Xen. Anab.* 1.4.7)
- e. μετὰ Ἰησοῦ    τοῦ    Γαλιλαίου  
 with Jesus-GEN LNK-GEN Galilean-GEN  
 ‘with Jesus of Galilee’ (*Mt.* 26:69)

<sup>5</sup> It could be added that the term ἄρθρον ‘joint’ of the Ancient Greek grammatical tradition — upon which the Latin term *articulus* ‘article’ was later calqued — was initially used for both the relative and the demonstrative pronouns (see Schwyzler & Debrunner 1939–1950: II, 14; Delbrück 1900: 406), which likely reflects the *Sprachgefühl* of the Greek grammarians.

Traditionally, the pattern with  $\delta\varsigma$ ,  $\eta$ ,  $\delta$  (such as the one in 9) was the only one to be compared to the Indo-Iranian verbless relatives, thanks to a common etymology of the pronominal linkers (see Benveniste 1966 among others). However, in my opinion, the articulated NP in Greek represented a sort of “second take” of the same innovation, with just a different pronominal linker. The demonstrative  $\delta$ ,  $\eta$ ,  $\tau\acute{o}$  shows the most advanced signs of grammaticalisation, i.e. the case agreement with the head noun (see 10b, 10e), while  $\delta\varsigma$ ,  $\eta$ ,  $\delta$  appears only in the nominative.

The resemblance of the Greek constructions with a post-nominal article to the Old Persian “quasi-article” is impressive. Such Greek appellatives as  $\Sigma\acute{o}\lambda\omega\nu\ \delta\ \text{Ἀθηναῖος}$  ‘Solon the Athenian’ are a perfect match of the corresponding Old Persian appellatives, e.g. *Gaumāta haya maguš* ‘Gaumata the Magian’, especially with respect to the linear order “head – linker – epithet”. Speaking of which, already Aristotle distinguished three main constructions of the articulated noun:

- I. “article – modifier – head” (as in  $\delta\ \acute{\alpha}\gamma\alpha\theta\acute{o}\varsigma\ \acute{\alpha}\nu\eta\rho$ );
- II. “article – head – article – modifier” (as in  $\delta\ \acute{\alpha}\nu\eta\rho\ \delta\ \acute{\alpha}\gamma\alpha\theta\acute{o}\varsigma$ );
- III. “head – article – modifier” (as in  $\acute{\alpha}\nu\eta\rho\ \delta\ \acute{\alpha}\gamma\alpha\theta\acute{o}\varsigma$ ).

The first pattern is the default one, which prevails in the texts. The third construction is the most similar to the corresponding Old Persian construction but is diachronically declining. It is also the only possible construction with the relative linker  $\delta\varsigma$ ,  $\eta$ ,  $\delta$ . The second pattern, on the contrary, is a typical Greek innovation, with both the head and the modifier marked for dependency (the double-marking typology). It has been observed that, from Homer to Herodotus, Thucydides, Xenophon, Tragedy, Aristophanes and Attic Oratory, the third pattern presented a rapidly decreasing frequency of occurrence, while the double-marked pattern steadily increased in frequency (see the figures in Gildersleeve 1900–1911: §609). Later, in the Biblical Greek, the third pattern resurfaces again (see 8c), perhaps under the pressure of its Semitic substrate (cf. §4.2 for the Semitic NP construction). Interestingly, the 5<sup>th</sup> cent. Armenian translation of the Greek Bible renders the third pattern with its own verbless relative construction, see §2.1.5.

In summary, Homeric Greek attests two constructions competing for the same function, and using two very similar pronominal stems as linkers (note that Old Persian *haya-* is a concatenation of the same two stems). The syntactic similarity between the oldest Greek pattern and the Iranian construction cannot be denied. However, in my account, the oldest construction with  $\delta\varsigma$ ,  $\eta$ ,  $\delta$  has been superseded by the articulated construction of the type  $\acute{\alpha}\nu\eta\rho\ \delta\ \acute{\alpha}\gamma\alpha\theta\acute{o}\varsigma$ , which, in its turn, also gradually fell out of use. Instead, the double-marked articulated NP, unknown elsewhere, had gradually gained frequency as a typically Greek innovation. As for the agglutination status, the article has always remained a clitic up to the present days.

### 2.1.5 Armenian

Armenian data, presented by Meyer (2017: §5.1), have been usually neglected within the classical studies on the IE verbless relatives, with which they have much in common. The examples in (11) are from Meyer (2017), whose distinction between clausal verbless relatives and *ezāfe*-like verbless relatives is disregarded here.

- (11) a. *anun Astucoy or tearn n araracoc'*  
 name-NOM God-GEN LNK-NOM lord-GEN DET creatures-GEN  
 'the name of God, [which is that of] the lord of creation' (*Ezrik Kotbac'i* 358)
- b. *ordi [...] orum anun Ormzd*  
 son-NOM [...] LNK-DAT name-NOM Ormzd-NOM  
 'a son [...] whose name is Ormzd' (*Ezrik Kotbac'i* 145)
- c. *hogi mardoyn or i nma*  
 spirit-NOM man-GEN LNK-NOM in he-LOC  
 'the spirit of the man which [is] in him' (1 Cor. 2:11)

The etymology of the Armenian relative stem *or* is debated. Meillet (1906–1908) considers it a possible outcome of PIE *\*k<sup>w</sup>o-*. The loss of the word-initial labiovelar stop is difficult to explain, but it is paralleled by some other pronominal derivatives. Instead, Pisani (1950) was the first to propose PIE *\*io-* as its antecedent, based on other examples of word-initial deletion of PIE *\*j* in Armenian.

Armenian does not present a high degree of grammaticalisation of the verbless relative. Syntactically, it can stand far from the head noun. Neither does it show case agreement with the antecedent (see 11b). It is noteworthy, however, that *or* is used for marking precisely the constituency rather than the definiteness of the noun, which was marked separately by the determiner *-n* cliticised to the noun (as in 11a).

The patterns attested in Classical Armenian have been explained as a calque on the Middle Iranian *ezāfe*-like constructions, but, according to Meyer (2017: §5.1.5), the data at our disposal are insufficient to prove a contact-induced origin. On the other hand, it can also be a calque of the Greek article in attributive position, as in (11c), where *or i nma* 'that which [is] in him' translates Greek τὸ ἐν αὐτῷ. Given these uncertainties, as well as the lack of any continuation of the verbless relatives in Modern Armenian, we should regard it as belonging to the periphery of the isogloss under consideration.

### 2.1.6 Slavic

Slavic presents the pronominal stem *jǐ* (spelled *i*), *ja*, *je* used as a demonstrative and a relative pronoun (if enlarged with *-že*), see Večerka et al. (1994: 245). If agglutinated to nominal stems, it functions as a special adjectival marker. The strong phonological reduction at the word-end in Proto-Slavic prevents us from establishing the exact etymon of *jǐ*: it could be an outcome of either the thematic (*\*io-*), or the athematic (*\*i-*) PIE pronominal stem, or even a merger of the two (see Petit 2009: 350; Wissemann 1957: 62–63; Rosinas 1975). This relative linker is agglutinated directly to the dependent nouns functioning as epithets of nominal heads. The resulting forms are called “long” adjectives, as opposed to the simple or “short” adjectives, morphologically indistinguishable from nouns.

In the oldest sources, both the nominal stem and the pronoun were still fully inflected. As the process of agglutination advanced, the initial /j/ of the pronoun underwent phonological erosion, resulting in a hiatus (note that hiatuses are normally prohibited by the phonotactics

of Slavic). At the last stage, even the hiatus was deleted, which led to the complete fusion of the two endings into one (see on this Polivanova 2013: §394). All three stages are effectively attested in sources, sometimes even as spelling variants within one and the same manuscript; see (12).

- (12) a. *vide* *člověčsky* *dobryje* *ženy*  
 seeing human-ACC.PL good-ACC.PL#LNK-ACC.PL wife-ACC.PL  
 ‘seeing the good wives of men’ (*Codex Suprasliensis* 7.18–19)
- b. *dobraago* *sūkrovišta*  
 good-GEN.SG-LNK-GEN.SG treasure-GEN.SG  
 ‘of the good treasure’ (*Codex Assemanius*, Mt. 12:35)
- c. *dobrago* *sūkrovišta*  
 good-ADJ.GEN.SG treasure-GEN.SG  
 ‘of the good treasure’ (*Codex Marianus*, Mt. 12:35)

In (12a) the boundary between the nominal and the pronominal word-forms *dobry* and *je* is still well preserved (marked as # in the glosses); in (12b), following the deletion of /j/, nominal and pronominal word-forms are only “separated” by the hiatus; in (12c) the pronominal word-form is totally fused with the nominal ending, thus turning into a special adjectival ending (glossed as ADJ).

The grammaticalisation of the pronominal linker in Slavic must have been finalised soon after the beginning of the written tradition (end of the 1<sup>st</sup> millennium AD). Some signs of the preceding syntactic autonomy of the linker are observable in the oldest sources, see the data collected by Vaillant (1942). Thus, sometimes it functioned as a phrasal clitic, rather than as a word affix (as in 12a), or was affixed to bases other than plain nominals, such as: numerals, adverbs, frozen locatives with adverbial meaning, or even entire PPs.

### 2.1.7 Baltic

Baltic languages are strikingly similar to Slavic in having a special adjectival declension built up by agglutinating the pronominal element *-is*, *-ji*, almost identical to the anaphoric pronoun *jis*, *ji*, on the adjectival word-forms. As in Slavic, the source of the pronominal element can be either the thematic \**īo-*, the athematic \**ī-*, or a blend of the two. Such adjectives are usually called “definite”. In (13) two examples are provided, respectively from Modern and Old Lithuanian; the boundary between the adjectival form and the pronominal linker is glossed with #.

- (13) a. *aukštoji* *mokykla*  
 high-NOM#LNK-NOM school-NOM  
 ‘the university’ (lit. “the high school”)
- b. *awineta* *ne* *iokaltoia*  
 lamb-GEN NEG LNK-GEN#guilty-GEN#LNK-GEN  
 ‘of the innocent lamb’ (*Knyga Nobažnystės* 1.258.16)

The grammaticalisation of the pronominal linker in Baltic arose relatively late, postdating not only the supposed Balto-Slavic but even the Proto-Baltic stage. Today, the hypothesis of a genetic origin for this innovation from a common Balto-Slavic pattern, once assumed as self-evident, is considered untenable (Petit 2009; Hill 2014; Sommer 2018). Let us review the main differences between Slavic and Baltic constructions.

Chronologically, Baltic forms are a later acquisition than the Slavic ones. In the oldest Baltic sources, around mid-2<sup>nd</sup> millennium AD, the pronominalised forms are still rare and not fully grammaticalised, while the corresponding Slavic forms had undergone complete grammaticalisation already centuries prior. In Old Prussian, the oldest attested Baltic variety, the pronominal linker is still limited to a few NOM and ACC singular word-forms. Old Lithuanian adds further case-forms, but the linker still shows signs of a higher syntactic autonomy: it can be prefixed, rather than suffixed, to the word-form, and even reduplicated (cf. the two instances of the linker in 13b). It can be used as a phrasal clitic, rather than a word affix (see Petit 2009: 319). It post-dates the creation of the agglutinative locative markers in Old Lithuanian. Finally, the grammaticalisation was accomplished separately in Lithuanian and Latvian (more details in Zinkevičius 1958: 52–53).

In Lithuanian, the pronominal linker is added to adjectival forms that have specific adjectival endings (of pronominal origin), while Old Slavic simple (or “short”) adjectives were indistinguishable from nouns. Kuryłowicz (1975) considers this only fact as crucial for rejecting any Balto-Slavic explanation for this innovation. However, Kuryłowicz’s claim is less convincing if we consider that the creation of specific adjectival endings of pronominal origin represented, in its turn, a secondary and quite late innovation in Baltic. Only a few adjectives with pronominal endings are attested in Old Prussian (see Trautmann 1910: §159–161; Dini 2014: §6.3.2.5). There are only six such endings in Lithuanian, in the masculine declension only. In Modern Latvian, they are widespread in the adjectives and have even penetrated into the nominal paradigm. Effectively, it seems that Baltic presents two parallel innovations independent from each other: agglutination of the pronominal linker and transfer of pronominal endings to the simple adjectives. The first is, however, the older of the two.

Baltic and Slavic differ also in the distribution and the semantic value of their long adjectives. Baltic ones are rare; in today’s Lithuanian, they appear especially in frozen phrases and idioms (such as the one in 13a). The simple adjectives represent the default choice. In Slavic, on the contrary, the long adjectives have ousted the simple ones almost completely. Both Baltic and Slavic pronominalised adjectives are often described as marking the definiteness, but the mismatch in frequency makes such interpretation less convincing (see the discussion in §3.2).

### 2.1.8 Germanic

The so-called “strong” adjectives in Germanic appear at first similar to the pronominalised adjectives in Baltic and Slavic: a special adjectival declension is obtained affixing pronominal-style endings to adjectival bases. The question is whether this similarity is enough to include the Germanic branch in our isogloss. The data at our disposal are consistent with such a hypothesis.

Normally, a definite articulated NP implied a “weak” adjective while the strong form was used elsewhere. Some exceptions to this mechanism are also known. In (14) three examples

from Gothic are quoted; strong endings are glossed with STR followed by the case, while weak endings are only glossed for the case; note that the declension of articles and pronouns is automatically “strong”.

- (14) a. *hlaif unsarana þana sinteinan gif uns*  
 bread-ACC our-STR.ACC the-STR.ACC daily-ACC give we-DAT  
 ‘Give us our daily bread’ (*Codex Argenteus*, Mt. 6:11)
- b. *atgiban þana swaleikana unhulþin*  
 deliver the-STR.ACC such-STR.ACC devil-DAT  
 ‘to deliver such a man to the devil’ (*Codex Ambrosianus A*, 1 Cor. 5:53)
- c. *allata galaubeiþ, all-∅ weneiþ*  
 all-STR.ACC believes, all-STR.ACC hopes  
 ‘believes all, hopes for all’ (*Codex Ambrosianus A*, 1 Cor. 13:7)

The modifier of *hlaif* in (14a) is constituted by the article and a weak adjective (the standard pattern), plus a possessive pronoun with a strong ending. The phrase in (14b) presents the adjective *swaleiks* ‘such, similar’ which remains strong even when preceded by the article. The passage in (14c) presents a zero-marked strong adjective in the ACC.SG neuter form, which can be interpreted as a simple adjective (see below).

Strong endings are clearly of pronominal origin, with two possible explanations. According to the “agglutination hypothesis”, some very short pronominal forms, possibly from the PIE stem *\*e-/o-*, were added to the nominal stems. This idea was dismissed by Sievers (1876) and Leskien (1876: 138) but later proposed again, see the survey in Bammesberger (1990: §10.2.2.5).

According to the “transfer hypothesis”, pronominal endings were simply transferred to the nominal stems, possibly through lexical diffusion not yet accomplished at the time of the Gothic sources (see Ratkus 2015).

The agglutination hypothesis received some criticism, which — as I believe — can be convincingly addressed. First, it is claimed that only some of the strong endings, with even some language-specific variance, appear unambiguously pronominal (contrariwise, if agglutination had ever occurred, it should have involved all case forms). However, McFadden (2004) is successful in explaining the entire strong paradigm, not just some endings, as pronominal.

Second, it is sometimes claimed that Germanic does not attest any sort of bare adjectival stems, similar to the Russian short adjectives, which would have been natural to expect as a sub-product of the agglutination; yet, such forms are effectively attested in Gothic, even if limitedly to the zero-marked NOM/ACC.SG of the neuter (see 14c). After all, Russian short adjectives are likewise limited to the nominative form, only in the predicative position.

Third, if we assume that the agglutinated pronoun must go back to PIE *\*io-*, as Sievers (1876: 99) does, then the disappearance of the initial /j/ of such a pronoun is phonologically problem-

atic since the internal /j/ is normally retained in Germanic.<sup>6</sup> However, this problem will not arise if we consider other possible sources for the agglutinated pronoun.

Postulating *\*o-/e-* as the source stem for the strong endings, the two hypotheses are conciliated: the word-forms of this pronominal stem is practically identical to its own endings, which makes the difference between transfer and agglutination insubstantial, especially if we date it far back in time. The direct reflexes of PIE *\*o-/e-*, though unattested in Germanic, are known from other daughter languages, cf. Sanskrit oblique forms GEN *asyá*, DAT *asmaí*, LOC *asmín*, indistinguishable from their endings, or also Greek ἤ ‘than’ (Dunkel 2014: 183ff. Hirt 1934: 162). A similar, even if less explicit, conciliatory approach is that of Prokosch (1939: 261) who, in stressing the parallelism of the Germanic strong adjectives with the Balto-Slavic pronominalised adjectives, claims that “the strong declension adds pronominal elements”, but a few lines later, speaks about “transfer of pronominal endings”.

Scholars have drawn attention to the fact that the strong adjectives usually appear in non-definite NPs while it is the weak ones that express definiteness (see e.g., Hill 2014). If we consider this difference crucial, then Germanic strong adjectives are not comparable with the supposedly definite long adjectives in Balto-Slavic. Accordingly, some alternative comparisons have been formulated. Besides Kuryłowicz (1975)’s idea of paralleling strong adjectives to the bare adjectives provided with their special endings of pronominal origin in Baltic (see §2.1.7), it is the weak adjectives that have often been viewed as corresponding to the Balto-Slavic definite adjectives (see, e.g., Hill 2014). In this last view, the nasal suffix of the weak adjectives has been sometimes considered the outcome of the agglutination of an ancient pronominal stem to a nominal base. Such a reconstruction seems unconvincing: the same nasal affix is also known in other IE languages where it is in no way interpretable as the agglutination of a pronoun, being simply a derivative suffix.

This issue is easily solved if we give up to the definiteness marking as the primary function of these forms; see, among others, Rießler (2016: §9.1.2.6) and the discussion in §2.2.1 and §3.2 below.

### 2.1.9 Latin

Latin attests several verbless relative constructions with the relative linker *qui, quae, quod* (from the PIE stem *\*k<sup>w</sup>i-*); see the examples in (15).

- (15) a. *qui patres qui conscripti*  
 LNK-NOM fathers-NOM LNK-NOM registered-NOM  
 ‘the fathers, the enlisted ones’ (Festus, 338.25)

<sup>6</sup> Thus, to Sanskrit *madhyas* ‘middle’ corresponds Gothic *midjis*; cf. also the strengthened outcome between vowels: Sanskrit *dváyos* ‘of the two’ vs. Old Norse *tveggja*, Gothic *twaddje*; see Prokosch (1939: 91–93).

- b. *divi qui potes pro illo quod Samothraces*  
 gods-NOM LNK-NOM powerful-NOM for that-ABL which-ACC Samothracians-NOM  
*theoe dynathoe*  
 θεοὶ δυνατοί  
 ‘the powerful gods, for what Samothracians [call] θεοὶ δυνατοί’ (Varro, L.L., 5.58)
- c. *sed nihil quod crudele utile*  
 but nothing-NOM LNK-NOM cruel-NOM useful-NOM  
 ‘but nothing that is cruel is useful’ (Cicero, *Off.*, 3.46)

The scarce and unsystematic Latin data introduced into debate by Benveniste (1966: 220) and supported also by Ivanov (1979) should undergo a critical revision. The phrase in (15a) is a legal formula referring to the members of the Roman Senate. However, in many editions, the two occurrences of *qui* are interpreted as interrogative rather than relative pronouns (Benveniste is silent on this topic). The example in (15b) is a quote from *Augurum Libri*, an archaic divination text. Note, however, that the second clause also lacks the predicate (supplied in translation), which makes the first verbless clause less specific. In general, Latin is known for easily dropping the predicate for stylistic purposes (see Lavency 1998: 112–113). This tendency includes the copula of relative clauses, which produces a verbless relative NP. One can see such an example of the Classical Latin verb elision for stylistic reasons in (15c). No sign of grammaticalisation of *qui* is detectable, besides the lack of the verb. Nor is this construction significantly reinforced in later stages of Latin, or its Romance descendants. For all of these reasons, I suggest excluding Latin from the isogloss under discussion.

### 2.1.10 Anatolian

Hittite relative construction, based on the pronoun *kuiš* as the relative linker, is sometimes attested without the verb. Benveniste (1966: 218) includes such examples in the isogloss. So too does Ivanov (1979: §3.1.4). Three Hittite verbless relatives are presented in (16); note the article-like use of *kuiš* in (16b–c).

- (16) a. *šallayaš kan DINGIR<sup>MEŠ</sup>-aš kuiš šalliš*  
 great-DAT.PL PTC gods-DAT.PL LNK-NOM.SG great-NOM.SG  
 ‘the greatest of the great gods’ (KUB 31.141 3)
- b. *kuiēš daranteš kuiēš UL daranteš*  
 LNK-NOM.PL named-NOM.PL LNK-NOM.PL NEG named-NOM.PL  
 ‘those named and those unnamed’ (KUB 6.45 III 5–6)
- c. *kuit ḫandan apāt išša*  
 LNK-NOM.SG right-NOM.SG that-ACC.SG do  
 ‘do what is right’ (KUB 13.2 III 28)

The Hittite verbless *kuiš* presents almost no sign of grammaticalisation and seems to be the consequence of a general tendency to drop the verb (not unlike Latin, see §2.1.9). Ivanov

(1979) claims that the Hittite quality modifiers with a “postclitic” linker are exact parallels of the pronominalised adjectives in Balto-Slavic. However, this is not the case in every occurrence. It is also possible that *kuiš* was simply located in Wackernagel’s position, including the cases where it breaks the boundaries of the modifier phrase, as in (16a), where *kuiš* is inserted inside the PP. Consequently, Hittite data appear as unreliable and dismissible as Latin.

It has been observed that the bilingual Akkadian-Hittite “dictionaries” often translate Akkadian adjectives or participles with relative constructions in Hittite, whereby *kuiš* functions as a sort of article. Thus, Hittite *dammeškizzi kuiš* ‘who shows power’ translates Akkadian *ḫābilu* as ‘powerful’ (KBo I 42 II 31); see Ivanov (1979: 46–48). However, this is not enough to include Hittite in the isogloss since the pattern here is different (no nominal head, and furthermore, a finite verb instead of an adjective).

On the other hand, verbless relatives are sporadically attested in later Anatolian languages, including Luwian, and Lycian (see the examples in Hajnal 1997: 56), but it is the recently deciphered Carian that shows the richest evidence of a fully grammaticalised pronominal linker. The particle *ki* in Carian, descending from the same antecedent of Hittite *kuiš*, marked nominal modifiers, mostly patronymics and ethnic names. Two Carian funerary inscriptions from Memphis, around the 7<sup>th</sup> century BC, are displayed in (17); the original punctuation symbol “|” is preserved and glossed with a comma; some omitted words are supplied in the translation.

- (17) a. *iturowś* | *kbjomś* | *ki en* | *mwdonś ki*  
 Iturow-GEN , Kbjom-GEN , LNK mother-NOM , Mwdon-GEN LNK  
 ‘[Stele of] Iturow, [who is] the mother of Kbjom, of Mwdon [people]’ (E.Me 32)
- b. *šyins* | *upe* | *aries ki ted*  
 Šyin-GEN , stele-NOM , Arie-GEN LNK father-NOM  
 ‘Stele of Šyin, [who is] the father of Arie’ (E.Me 38)

According to a commonly accepted hypothesis (see Adiego 2007: 273; Hajnal 1997), the particle *ki* functioned as a relative pronoun if posed in proclitic position (cf. *ki en* ‘who [is] the mother’ and *ki ted* ‘who [is] the father’), but acted as a nominal determiner if in the “postclitic” position (cf. *mwdonś ki* ‘of Mwdon’); note that the position is inferred from the distribution of the vertical strokes, supposing that they separated prosodic units. In both cases, the relative phrase shows no verb. Therefore, *ki* should be considered another example of the relative linker grammaticalised as a modifier-marking clitic.

## 2.2 New suggested data

### 2.2.1 Modern German strong declension as a model

Let us return to the distribution of the strong/weak adjectives in Old Germanic, supposedly based on the notion of definiteness: weak adjectives appeared quite regularly with the definite article, while strong adjectives appeared elsewhere. Modern German preserves strong vs. weak endings, but the mechanism of their distribution became fully grammaticalised, with no more reference to definiteness. The Modern German NP is governed by the “once-per-phrase

principle” (cf. Evans 2019: §3.6): the “characteristic enlargement” — going back to the strong ending — appears only once per NP (or DP, if we prefer). Its presence is, in itself, a phrase-wide constituency marker, while the definiteness is fully expressed by the determiner. Observe the behaviour of the ending *-es/-s*, glossed as STR, in (18); the examples are from W. P. Lehmann (1957: 126) and Evans (2019: §3.9).

- (18) a. *das dunkle Bier*  
 the-NOM.STR dark-NOM beer-NOM  
 ‘a dark beer’
- b. *ein dunkles Bier*  
 a-NOM dark-NOM.STR Bier-NOM  
 ‘one dark Bier’
- c. *recht dunkles Bier*  
 quite dark-NOM.STR beer-NOM  
 ‘quite a dark beer’
- d. *Ihre Dunkles*  
 your-NOM dark-NOM.STR  
 ‘your dark one’ (i.e. “your dark beer”)
- e. *schlechten Bieres*  
 bad-GEN beer-GEN.STR  
 ‘of bad bear’

Where exactly the “enlargement” appears is governed by a subtle hierarchy among the word-forms of nouns, definite articles, adjectives, indefinite articles. Note that, syntactically, it goes on the dependent most of the times, however it can also mark the head, in case it happens to be a substantivated adjective (as in 18d), or a genitive nominal (as in 18e). This system represents, for us, a working prototype of the “alternant marking” strategy mentioned in §1.2.2. With this in mind, we are now able to analyse possible examples of the same strategy in two more languages.

### 2.2.2 Khotanese

In Khotanese, two nominal declension types are attested. One type is characterised by a set of endings that are phonologically longer than the corresponding endings in the other type. The longer endings are clearly preferred with adjectives and pronouns while quite rare, though not unattested, with nouns (see Sims-Williams 1990: 278; Emmerick 2009: 386). The enlargement consists in the insertion of the nasal element *-än-* before the ending proper, e.g.: ABL/INST.SG.F *-äñe* (vs. nominal *-<sup>i</sup>e*), LOC.SG *-äñā* (vs. nominal *-<sup>i</sup>a*). Another enlarged adjectival ending is GEN/DAT.SG *-ye* (vs. nominal M *-i* and F *-<sup>i</sup>e*).

Emmerick (1968: 257ff.) explains the nasal-enlarged endings as contaminated by the Old Iranian nominal declension of the nasal stems. But a more fascinating reconstruction is suggested by Sims-Williams (1990: 276ff.) who sees, in the nasal enlargement, the outcome of an

agglutinated pronominal stem *\*ana-* (correlated to Slavic *onŭ* ‘he’, Lithuanian *anàs* ‘that’, Vedic *anā* ‘therefore’, Hittite *anniš* ‘that’). The GEN/DAT.SG adjectival ending *-ye* is likewise explained by Sims-Williams (1990: 279) as a transfer from the pronominal paradigm, while Konow (1949: 41) considers it the outcome of *\*yahya*, which looks quite similar to the genitive word-form of the Old Iranian relative pronoun.

Moreover, there also exist “nasal” endings that do not correspond to any short counterparts: ABL/INST.SG.M *-āna* and GEN/DAT.PL *-ānu*. However, in a group of coreferential ABL/INST.SG.M nominals, one word-form may lose the nasal element from its case ending so that the otherwise unattested “shortened” forms come out, such as ABL/INST.SG.M *-ā*. Emmerick (1968: 258) explains this “shortening” as analogical with the aforementioned “enlargement”. The whole mechanism is referred to as “group inflection” by Emmerick (2009: 399); however, it may also be described as a tendency to allow only one occurrence of the nasal element per NP, not unlike the once-per-phrase principle in Modern German (though applied less regularly). Sometimes the dropping of the nasal element is explained *metri causa*. However, a functional burden can effectively be attributed to this phenomenon: the alternation of short and enlarged endings can encode the relation between a nominal head and its qualifiers.

Some examples of Khotanese large vs. short endings are given in (19). The large endings (either enlarged, or “natural”) are glossed LRG, while the deleted nasal element of the “shortened” endings is transcribed as  $\emptyset$ .

- (19) a. *natāña*            *balysāña*            *rahāśśa*  
 deep-LRG.LOC Buddha’s-LOC secret-LOC  
 ‘in deep Buddha’s secret’ (Suv. 4r.2)
- b. *hauva*            *prribhāvana*  
 strength-INST power-LRG.INST  
 ‘with powerful strength’ (Bcd 51v.1)
- c. *dārye*            *hārā*  
 bad-LRG.GEN thing-GEN  
 ‘of the bad thing’ (Z. 24.433)
- d. *purokā-∅* *Ysarkulna*  
 son-INST Zarkula-LRG.INST  
 ‘with the son [called] Zarkula’ (manuscript E, add. 12).

The form *balysāña* in (19a) is a short (i.e. nominal) locative of the possessive derivative *balysāna-* from *balysa-* ‘Buddha’. In (19b), which comes from Late Khotanese, the two word-forms are both nouns, but from the fact that a “group inflection” is attested, we can conclude that one is marked as the head noun, while the other is intended as its nominal epithet. The same goes for (19d), which contains a common noun and a coreferential proper name.

To sum up, if the suggested syntactic interpretation is correct, the Khotanese pattern belongs to the same type of “strong” adjectival declension as the one attested in Modern German. The origin of the enlarged endings remains uncertain, but if we accept the pronominal agglutination hypothesis, then it appears to be another case of grammaticalisation of a pronominal

linker. Note that the locus of marking seems alternant, unless this instability is simply explainable with the difficulty of individuating the head within a NP on a purely semantic ground.

### 2.2.3 Middle Indian

Another example of the alternant strategy in assigning dependency marker through longer and shorter case endings is attested in some Prakrit varieties, including Pāli. The alternation regards only a few grammatical cases, namely locative in Prakrit and locative plus ablative in Pāli. The old LOC.SG ending *-i/-e* alternates with the pronominal ending variously spelled as *-ammi*, *-ahmi*, *-amhi*, and *-aṃsi* (from Sanskrit *-asmin*). In Pāli, there is also the ABL.SG termination *-ā* alternating with *-asmā* and later *-amhā* (from Sanskrit *-asmāt*). So far, no convincing explanation for the distribution of the two variants has been proposed, if we except the not very convincing *metri causa* motivations claimed by Pischel (1900: §366a).<sup>7</sup> Moreover, it has been observed that the two variants are frequently used alongside each other on coreferential nominals (cf. Woolner 1928: 36), not only in verses but also in prose. It cannot be excluded that this alternating marking encoded the dependency relation within the NP, not unlike the adjectival vs. nominal endings in Khotanese, or the strong vs. weak adjectival declension in Germanic.

For example, in Hala's *Sattasāi* almost every long locative ending agrees with a short one in the same NP, see (20) for some examples (long endings are glossed LRG).

- (20) a. *diṭṭhe sarisammi guṇe*  
 viewing similar-LRG.LOC quality-LOC  
 'viewing similar quality' (*Sattasāi* 44cd)
- b. *putte samāruhattammi*  
 son-LOC climb.up-PART-LRG.LOC  
 'when the son has climbed up' (*Sattasāi* 11ab)
- c. *vāsuikaṃkaṇammi osārie*  
 snake.bracelet-LRG.LOC remove-PART-LOC  
 'being the snake-bracelet removed' (*Sattasāi* 69cd)

Here, as in Khotanese, the locus of marking is not fixed, at least at first sight: the long ending is on the modifier in (20a–b) but on the head in (20c). We should not be distracted by our translational equivalents of the Prakrit phrases. The headedness of a NP is always questionable. What is translated as 'a bracelet that is removed' could be also understood as 'a removed thing that is a bracelet'. Note, however, that the "once-per-phrase principle" is normally respected here too.

The etymology of Prakrit long case-endings, unrelated to the Vedic "quasi-article" discussed in §2.1.1, can be individuated in the agglutination of the same short, "ending-like", demonstra-

<sup>7</sup> Note that Pischel's use of the *metri causa* explanation differs from what is usually done in Classical philology: there, metre is used to *confirm* some apparently aberrant forms as genuine; here, instead, metre is invoked when one needs to *dismiss* some aberrant forms as poetic artifacts.

tive pronouns going back to PIE *\*e-/o-*, already mentioned with reference to Germanic. Such pronouns are even attested as such in MIA, cf. Pāli inflection of *ayaṃ* ‘it’: ABL *asmā*, GEN *assa*, and LOC *asmīṇ*.

Again, as in Khotanese, the enlarged endings are attested only in a small part of the nominal paradigm. Contrary to Khotanese (and Germanic), this system was not diachronically reinforced and disappeared completely after the phonological simplification of the word ending occurred in Late MIA.

### 2.3 Parametrical table

Table 1 summarises the data surveyed in the preceding sections. The parameters defined in §1.2 are given in the columns. The first column represents the overall stage of grammaticalisation; the next column shows the agglutination status, with three values: “free”, “clitic”, and “bound”; the next one shows the locus of marking; then, the linear position (as “pre” or “post”); the last column shows the source of the mark.

	GRAMMATICALISATION STAGE	AGGLUTINATION STATUS	LOCUS OF MARKING	POSITION OF THE MARK	SOURCE OF THE MARK
Persian	(c) to (e)	clitic	head	post	<i>*so+io-</i>
Greek	(b) to (c)		double		<i>*so</i> (also <i>*io-</i> )
Avestan	(b)	free	dependent	pre	<i>*io-</i>
Armenian	(a)				<i>*k<sup>w</sup>o-</i> or <i>*io-</i>
Slavic	(d)	bound	dependent	post	<i>*io-</i>
Baltic	(c) to (d)				
Germanic	(d) to (e)		alternant		<i>*o-/e-</i>
Khotanese	(d) to (e)				
Prakrits	(d) to (e)				
Hittite	(a)	free	dependent	<i>*k<sup>w</sup>l-</i>	
Carian	(c)	clitic			

Table 1: Summary parametrical table

Each row summarises the most specific values for each language. Thus, for Greek, the double-marking typology is mentioned (corresponding to the II pattern of the articulated NP), even if it also attests the dependent marking (in the III pattern). “Persian” subsumes both Old and Middle Persian features. Vedic and Latin are not included since their verbless relative NPs, for different reasons, cannot be considered part of the isogloss. The degree of syntactic autonomy can be shown as a range in case the language under consideration presents significant diachronic change.

### 3 Discussion

#### 3.1 Previous studies

The isogloss under consideration is not a new discovery. However, preceding studies have almost exclusively analysed subsets of the entire list of languages considered in the present paper. Interestingly, no preceding author has ever made the conclusive step unifying all such partial results into one major isogloss.

Let us review the most convincing attempts of connecting single languages or groups of languages made so far. To start with, a strong defence of the Irano-Slavic convergent grammaticalisation of the relative pronoun as an adjectival marker is presented by Meillet (1934); cf. the criticism in Parenti (1996: 36) and the approach of Petit (2009: 355), who accepts the Balto-Slavic common development, but rejects the connection to other IE branches. On the other hand, the parallelism of the Slavic long adjective with the strong declension in Germanic has been suggested already by Leskien (1876) and is supported by some modern scholars, e.g. Dyen (1990); see also Bauer (2007), who puts together Balto-Slavic and Germanic but excludes Old Persian.

Benveniste (1966) has considered a larger set of verbless relative constructions, including Vedic, Avestan, Balto-Slavic, Greek (limitedly to the relative linker  $\delta\varsigma$ ,  $\eta$ ,  $\delta$ ), Hittite, and Latin. The same data are discussed in Ivanov (1979), with the explicit exclusion of the Germanic strong adjectives. Hajnal (1997) adds more Anatolian data, including Carian. Finally, Meyer (2017) adds Armenian verbless relatives to the discussion.

On the other hand, the Greek article  $\delta$ ,  $\eta$ ,  $\tau\acute{o}$  in postnominal position has been compared with the Iranian “quasi-article”, e.g. by Schwyzer & Debrunner (1939–1950: II, 26, fn. 1) and Seiler (1960). Probert (2015), while focusing on Greek relatives, including verbless ones with both linkers, gives reference to the parallel patterns in the other IE branches. Interestingly, the parallelism of the (New) Persian *ezāfe* with the (Modern) Greek article is treated as obvious in many synchronic studies dealing with the linking function, such as C. Lehmann (2018) and Franco et al. (2015).

Table 2 sums up these partial results; “+” means that such language is mentioned in connection to the discussed isogloss while “–” means that such a connection is explicitly criticised or denied; a blank cell means that this language is not mentioned at all.

A set-theoretical unification of all the partial pairings proposed so far makes, therefore, the next logical step towards the whole picture. For example, if we accept Benveniste’s unification of the Avestan, Old Persian, Hittite, Greek, and Latin verbless relatives with Balto-Slavic long adjectives, while, at the same time, supporting the idea of the equivalency between Balto-Slavic and Germanic pronominalised adjectives, then the Germanic data are automatically included in the isogloss. The same is true for the Armenian and Carian data, unknown or neglected by most of the preceding scholars. Furthermore, if my presentation of the Middle Indian and Khotanese data as a parallel to the Germanic NP construction is accepted, then these two languages must be added to the isogloss as well.

Such a unification has not been done basically for the following two reasons: 1) not all the languages that I include in the isogloss assign the same value of definiteness to the pronomi-

	Ved.	Av.	Pers.	Gr.	Slav.	Balt.	Germ.	Lat.	Hit.	Car.	Arm.
Leskien (1876)					+	+	+				
Dyen (1990)					+	+	+				
Bauer (2007)			–		+	+	+				
Meillet (1934)		+			+						
Vaillant (1942)	+	+	+	+	+	+	+				
Parenti (1996)					–	–					
Petit (2009)	–	–	–	–	+	+					
Benveniste (1966)	+	+	+	+	+	+		+	+		
Ivanov (1979)	+			+	+	+	–	+	+		
Seiler (1960)	+	+	+	+	+	+					
Hajnal (1997)	+	+	+	+	+	+	+	+	+	+	
Meyer (2017)	+	+	+	+				+	+		+

Table 2: Previous studies

nalised pattern; 2) not all the patterns use the same pronominal stem as the source of grammaticalisation. I address these apparent difficulties, respectively, in §3.2 and §3.3.

### 3.2 Definiteness or constituency?

Most of the existing studies of the pronominalised attributive constructions focus on the idea that the grammaticalisation of the pronominal linker primarily served the needs of marking definiteness. Some scholars, e.g. Benveniste (1966), consider the relative pronouns to have been a standard marker of definiteness already at the PIE stage. Others, e.g. Bauer (2007), suggest that there must have been an innovative typological drift towards definiteness marking shared by a series of attested IE languages. Consequently, all those languages that do not use the pronominalised pattern for expressing definiteness were automatically excluded from the isogloss. This would be the case of the strong adjectives in Germanic, which usually appear in non-definite NPs, whence the rejection of Germanic by Ivanov (1979: 62) and others. Likewise, the Armenian verbless relative construction could be expelled from the isogloss since this language uses a specific marker of definiteness, not coinciding with the relative linker. The same situation is attested also in Old Prussian where long adjectives were used alongside the definite article (see SOMMER 2018). Similarly, Bauer (2007: 109) fails to recognise the New Persian *ezāfe* as the natural development of the Old Persian quasi-article only because the latter cannot be explained in the same terms of definiteness marking as the former.

However, using definiteness as a criterion in diachronic analysis could be problematic not only in this specific case but also in general. Thus, we postulate gender as a grammatical category for a language in case the nouns with feminine referents are systematically different in morphological respect to the nouns with masculine reference, but we cannot use definiteness as a base for postulating a grammatical category because definiteness is not referential. We infer the definiteness of a noun from its grammatical marking, not the other way around. Thus, in reading a text in, say, Classical Greek, we hardly know in advance which nouns to consider

definite and which indefinite. The only certain data at our disposal is whether a noun is provided with a definite article, or not. Greek being a well-studied language, moreover, with an indigenous tradition of grammatical analysis, we have a good classification of the article's typical usages, yet no simple formal criterion for distinguishing them is available. When approaching a less well-known language, e.g. Old Lithuanian, predicting definiteness becomes virtually impossible.

Native speakers are not always helpful either: there are often inconsistencies in the distribution of the definiteness markers within one language and across parallel texts in different languages. Let us survey a collection of facts supporting this claim.

- Parenti (1996: 29) shows three different Modern Lithuanian renderings of a verse from the New Testament, of such kind that no pair of them shares the same distribution of definite vs. simple adjectives. This makes the definiteness appear to be a stylistic feature rather than a grammatical value.
- Similarly, Evans (2019: §5.2.1) provides examples of alternating definiteness status of adjectives in parallel Bible verses across Germanic languages.
- In the oldest Slavic written texts, there are instances of definite and indefinite adjectives used in conjunction with each other and referring to one and the same noun, which appears highly inconsistent with the logic of definiteness.
- The presumably definite adjectives of Old Lithuanian and Old Slavic evolved today to two exactly opposite systems: pronominalised adjectives have prevailed in Slavic (esp. Russian) whereas the simple adjectives have done so in Modern Lithuanian. If the long adjectives did effectively mark definiteness in both languages, then we would have to explain why most adjectives are definite in Russian, but indefinite in Lithuanian.
- There exist defective “definite only” as well as “indefinite only” adjectives in both Germanic and Slavic. However, as Orr (1983: 114) shows, there is no agreement in such defectiveness between translational equivalents in the two languages, e.g. Gothic *sums* ‘some’ (always strong, i.e. indefinite) vs. Slavic *někotoryi* ‘some, certain’ (always pronominalised, i.e. definite). This is difficult to explain from the logical point of view.

Consequently, I suggest considering definiteness as only a secondary additional value of the pronominal linkers, the primary function of which should be the marking of dependency or constituency. From a general point of view, constituency seems to me more pivotal as a value with respect to definiteness. Constituency is related to the referential value of a message while definiteness pertains only to the so-called “information packaging” within the message and is mainly stylistic rather than semantic.

If we consider constituency pivotal, then the main goal of the pronominal element becomes that of linking the head and the dependent words together, rather than marking definiteness. This approach is not unprecedented: the linking function of the pronoun in the verbless relatives has been already highlighted by such authors as Wissemann (1957), Seiler (1960), Kuryłowicz (1975), Franco et al. (2015), and Rießler (2016). Indeed, by ignoring definiteness we avoid

some false problems (definiteness mismatch across languages) and allow for the inclusion of more languages in our isogloss.

### 3.3 Relative or demonstrative?

Another misconception is, in my opinion, the idea that the pronominal linker must have necessarily been a relative pronoun in its prehistory. This is true for many of the languages under consideration, but apparently not true for others. Thus, the Greek construction with the postnominal article used as a linker has not been included in his analysis by Benveniste (1966) for the simple reason that Greek  $\delta, \eta, \tau\acute{o}$  is not traditionally considered a relative pronoun. However, the relative/demonstrative distinction seems to be very weak and interpretable, if not non-existent, in the oldest sources, especially in the ancient oral poetry with its typical “adding” syntax. Consider the incipit of the *Iliad* in (21):

- (21) μῆνιν ἄειδε θεὰ Πηληϊάδεω Ἀχιλῆος οὐλομένην, ἣ μυρὶ Ἀχαιοῖς ἄλγε’ ἔθηκε  
 ‘Sing, goddess, of the baneful wrath of Achilles son of Peleus, *which* gave countless griefs to the Greeks’

What is usually translated as ‘which’ is a Greek pronoun originally spelled  $\eta$ , that corresponds, in later orthography, to  $\text{ἣ}$  ‘which’, but also to  $\text{ἣ}$  ‘that’. Therefore, it may be translated as either a relative linker governing a subordinate relative clause, or a pronominal subject of an independent clause. Importantly, both versions would be contextually acceptable. The choice is mainly stylistic, rather than grammatical or semantic.

The interchangeability of relative and demonstrative pronouns has already been illustrated for Greek (see §2.1.4), but is also widespread elsewhere in the IE languages, starting from the English pronoun *that*, which can fulfil both functions. This circumstance makes the “true” PIE relative pronoun nearly impossible to reconstruct. Indeed, several contrasting proposals have been made but none have gained the general consensus, cf. the survey in Probert (2015: 23). Therefore, the relative semantics of the original pronominal linker should not be considered as pivotal as some scholars might have believed. The linking function can be successfully fulfilled regardless of the specific semantics of the pronoun.

## 4 Conclusion

### 4.1 The new shape of the isogloss

Whatever the origin of the isogloss might be, the first result of the present paper should be a reshaping of the isogloss’ boundaries, based on the above-mentioned theoretical assumptions and the revision of the linguistic data. Concerning the classical studies, the following additions and deletions are to be made.

Three languages should be excluded from the isogloss, namely Vedic, Hittite, and Latin. In all three, the degree of grammaticalisation of the pronominal linker is minimal. The presumed

verbless relatives are indistinguishable from the occasional omitting of the copula, which happens also in other constructions. The absence of a diachronic increase in grammaticalisation suggests that there was no grammaticalisation at all. In the case of Latin, the scarcity of occurrences is a major problem. In the case of Vedic and Hittite, the extreme antiquity of the sources speaks against the inclusion of these languages into the isogloss.

Certain languages must be added to the isogloss, namely Armenian and Carian. Germanic strong adjectives can be included in the isogloss if one ignores, as we do, the definiteness issue. Consequently, Middle Indian and Khotanese can also be included since they show enough analogies with Germanic. The fact that this pattern appears limited to just a few case-values is also not a major issue. We observe the same limitedness in other languages of the isogloss as well.

Finally, the status of Greek must be redefined. The idea suggested here is that the grammaticalisation of the pronominal linker was “passed” from one pronominal stem to another and that the articulated NP must be considered the ultimate outcome of this development.

In general, the grammaticalisation of the pronominal linker as a constituency marker can be considered one of the distinctive features of what can be defined as “second generation IE languages”, see Keidan (2013).

## 4.2 A typological parallel: possessive NP in Semitic

Before suggesting some tentative explanations for this isogloss, I wish to mention another language, Syriac, which can be considered a “minor scale” typological parallel to some odd features that we have met in analysing our IE data. In (22) three different variants of the possessive construction in Syriac are displayed (the examples are borrowed from Ciancaglini & Alfieri 2014: 118).

- (22) a. *bar nāšā*  
 son-STC man-STE  
 ‘son of man’ (i.e. “person”)
- b. *brā d-ʾalāhā*  
 son-STE LNK-God-STE  
 ‘the son of God’
- c. *br-eh d-ʾalāhā*  
 son-3SG.M LNK-God-STE  
 ‘the son of God’

The construction in (22a) is the most archaic (in fact, used as an idiom), and is a typical Semitic construction: the possessee is head-marked by the construct state (glossed STC), while the possessor is in the unmarked emphatic state (glossed STE). The construction in (22b) puts both nouns in the emphatic state and marks the dependent (possessor) with a genitival mark, etymologically going back to a third person pronoun. In (22c) the same construction is pre-

marked with another 3<sup>rd</sup> person marker put on the head, which anticipates the possessor. The situation in Syriac illustrates the following points:

- Different constructions can coexist and compete for the encoding of the same grammatical function. One can observe the same situation in Greek (with the two competing pronominal linkers), in Baltic (with two almost parallel innovations: pronominalisation of the adjectival endings and agglutination of pronominal linkers), and in Germanic (strong adjectival declension but also a new definite article).
- Pronominal linkers can be easily transformed in dependency markers. Indeed, the genitive marker of the second construction is etymologically a plain relative pronoun (see HUEHNERGARD 2018), still used as an uninflected relative linker in Syriac. Moreover, the pre-marking 3SG.M morpheme in the third example is also of pronominal origin.
- During the process of grammaticalisation, the same initial states can bring about different outcomes. Indeed, in North-Eastern Neo-Aramaic, a close relative of Syriac, the same pronominal marker has been cliticised to the head noun rather than the modifier.
- The latter case also illustrates the fact that NP encoding can be influenced, if not induced, by neighbouring languages: the North-Eastern Neo-Aramaic development seems to be correlate with the New Persian and Kurdish *ezāfe* constructions.

### 4.3 Possible explanations

In concluding this survey, I wish to discuss possible explanations of the isogloss under consideration. Three explanations are possible, in such respect: 1) genetically shared innovation, 2) contact-induced innovation, or 3) a naturally coincidental innovation (a Sapirian drift).

The genetic explanation of the isogloss, in the terms defended by Benveniste and Ivanov, is not tenable for at least two reasons (both equally disregarded by the two scholars). First, there is no common source of the relative pronominal linker shared by all the daughter languages belonging to the isogloss (cf. Probert 2015: 52), so that no common pattern can be reconstructed for PIE. Second, the chronology is also problematic: it is highly unlikely that a pattern of the protolanguage could have been “activated” in different languages separately in different moments ranging over more than two millennia.

Now, the contact borrowing hypothesis is tempting. In this view, the innovation must have started in the Iranian domain and then spread around to the neighbouring languages (and from these, to still others), due to the cultural and political prestige of Iran. However, this is only plausible for some of the connections. Hypothetically, we can imagine a situation of contact pressure from Persian to Armenian and from Persian to Carian. On the other side, a contact connection can be easily imagined between Slavic, Baltic, and Germanic groups. However, linguistic data per se are insufficient to prove a contact explanation: we need some historical evidence attesting the bilingualism. In the present case, such information is simply not available. For example, although contacts between Old Iranian tribes with Slavic populations seem possible, due to a series of lexical borrowings from Iranian to Slavic and even some shared syntactic

patterns (on which see Patri 2001), such contacts must have occurred in the prehistory, so we cannot be certain about them. Another problematic, yet not impossible, connection is the one between Persian and Greek. The two civilisations have been in conflict for centuries, which means the Greeks would not perceive Persian as a prestige language. Nevertheless, Hellenic and Iranian languages are known to share also other specific linguistic features (weakening or loss of /s/, to name one).

Typologically, the contact explanation seems plausible: it has been observed that the coding of the nominal modifiers (especially, possessors), are at the top of the scale of borrowability within nominal morphology (Matras 2007: 44), and that the grammaticalisation in general is often related to language contact (Hill 2014: 185).

Importantly, neighbouring languages would most likely not have developed these strikingly similar patterns by mere coincidence. If the contact borrowing of the pattern cannot be proven and the described isogloss is only the result of a drift, we need to find some structural explanation in the protolanguage that would explain such parallel developments in the daughter languages.

One such explanation can be individuated in the lack of grammaticalised adjectives in PIE. To put it more accurately: no standard encoding system for nominal modifiers can be reconstructed for PIE based on the data of the oldest attested daughter languages, all of which show no morphological distinction between adjectives and nouns. The latter can serve as nominal epithets, often allowing even the gender agreement with the head noun. The lack of noun/adjective distinction is still well preserved in some IE languages of later attestation. Such Slavic primary nominal stems as *drugŭ* could be interpreted as either a noun, ‘friend’, or an adjective, ‘other’, according to the context of the occurrence (Polivanova 2013: §268). The same situation is attested in Latin, see the NP in (23a), allowing two readings depending on which of the two nominals is the head and which is the modifier. The typically adjectival category of gradation could be applied to stems usually considered nominal, in Latin; see *homines magis asinos* ‘men more [similar to] asses’ in (23b), quoted from C. Lehmann (2018: 56).

- (23) a. *inimici*                      *Germani*  
 enemy-M.NOM.PL    German-M.NOM.PL  
 ‘German enemies’/‘hostile Germans’
- b. *homines*            *magis asinos*            *numquam vidi*  
 man-M.ACC.PL    more    ass-M.ACC.PL    never            saw  
 ‘I have never seen human beings who were more similar to asses’ (Pl. *Pseud.* 136).

The distinction between nouns and adjectives can reside at the level of lexicon rather than in morphology: some lexemes are consistently heads, while others are consistently modifiers, but in the oldest IE languages only a few such primary adjectives are attested, e.g. around a dozen in Vedic (see Alfieri 2016).

Unsurprisingly, all historically attested IE languages have tried to address this “breach in the system”. In the outcome of this process, we see the birth of a special adjectival declension in some languages (Germanic, Baltic, Slavic, perhaps Khotanese), and the creation of a special

linking morpheme in others (Persian *ezāfe*, Greek article). Note that some of these languages could have developed more than one means for the same goal. Thus, Greek, Slavic, Sanskrit, Latin, etc have also developed a class of “secondary” adjectives derived by suffixation, apophony, or accentual change from nominal stems.

The question, however, remains open. To make a comparison, proving the shared nature of an innovation can be virtually impossible even for closely related languages, see the study by Morpurgo Davies (2012) on the loss of /s/ in Greek dialects. What is there to say then about some not so closely related languages?

## 5 List of abbreviations

ABL	ablative	M	masculine
ABL/INST	ablative-instrumental (in Khotanese)	NEG	negation
ACC	accusative	NOM	nominative
ADJ	adjectival declension (in Slavic)	NP	noun phrase
CP	complementiser phrase	PART	participle
DAT	dative	PL	plural
DP	determiner phrase	PP	prepositional phrase
F	feminine	PTC	particle
GEN	genitive	SG	singular
GEN/DAT	genitive-dative (in Khotanese)	STC	<i>status constructus</i> (in Syriac)
INST	instrumental	STE	<i>status emphaticus</i> (in Syriac)
LNK	pronominal linker and its outcomes	STR	strong ending (in Germanic)
LOC	locative	TP	tense phrase
LRG	enlarged endings (in Khotanese and Middle Indian)		

## References

- Adiego, I. J. (2007). *The Carian Language*. Leiden & Boston: Brill.
- Alferi, L. (2016). “The definition of the root between history and typology.” *Archivio Glottologico Italiano* 101.2: 129–169.
- Bammesberger, A. (1990). *Die Morphologie des urgermanischen Nomens*. Heidelberg: Winter.
- Bauer, B. L. M. (2007). “The definite article in Indo-European: Emergence of a new grammatical category?” In: E. Stark et al., eds. *Nominal Determination. Typology, Context Constraints, and Historical Emergence*. Amsterdam & Philadelphia: Benjamins.
- Benveniste, É. (1966). “La phrase relative, problème de syntaxe générale.” In: *Problèmes de linguistique générale*. Paris: Gallimard: 208–222.

- Caland, W. (1891). *Zur Syntax der Pronomina im Avesta*. Amsterdam: Müller.
- Ciancaglini, C. A. & L. Alfieri (2014). "Iranian and Greek Influence on the Syriac Lexicon: the Emergence of Compound Words." *Incontri Linguistici* 36: 109–140.
- Delbrück, B. (1900). *Vergleichende Syntax der Indogermanischen Sprachen*. 3. Theil. Strassburg: Trübner.
- Dini, P. U. (2014). *Foundations of Baltic Languages*. Vilnius: Vilnius University.
- Dunkel, G. E. (2014). *Lexikon der indogermanischen Partikeln und Pronominalstämme*. Vol. 2. Heidelberg: Winter.
- Durkin-Meisterernst, D. (2014). *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Dyen, I. (1990). "The homeric argument for a Slavo-Germanic subgroup of Indo-European." In: P. Baldi, ed. *Linguistic Change and Reconstruction Methodology*. De Gruyter Mouton: 385–390.
- Emmerick, R. E. (1968). *Saka Grammatical Studies*. London: Oxford University Press.
- (2009). "Khotanese and Tumshuqese." In: G. Windfuhr, ed. *The Iranian Languages*. London & New York: Routledge. Chap. 7: 377–415.
- Evans, E. (2019). "The Origin, Functions, and Histories of Germanic Adjective Endings." PhD thesis. Indiana University. Department of Germanic Studies.
- Franco, L. et al. (2015). "Linkers and agreement." *The Linguistic Review* 32.2: 277–332.
- Gildersleeve, B. L. (1900–1911). *Syntax of Classical Greek. From Homer to Demosthenes*. New York, Cincinnati & Chicago: American Book Company.
- Gonda, J. (1955). "The original character of the Indo-European relative pronoun *jo-*." *Lingua* 4: 1–41.
- Haider, H. & R. Zwanziger (1984). "Relatively attributive: the 'ezāfe'-construction from Old Iranian to Modern Persian." In: J. Fisiak, ed. *Historical syntax*. Berlin & New York: Mouton de Gruyter: 137–172.
- Hajnal, I. (1997). "Definite nominale Determination im Indogermanischen." *Indogermanische Forschungen* 102: 38–73.
- Hill, E. (2014). "Inheritance and secondary similarities in the inflectional morphology of Baltic and Slavic." In: T. Civjan et al., eds. *Baltai ir slavai: dvasinių kultūrų sankirtos*. Vilnius: Vesmė: 177–194.
- Hirt, H. (1934). *Indogermanische Grammatik. Teil VI: Syntax I*. Heidelberg: Winter.
- Hopper, P. J. & E. C. Traugott (2003). *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.
- Humbert, J. (1960). *Syntaxe grecque*. 3rd ed. Paris: Klincksieck.
- Ivanov, V. V. (1979). "Sravnitel'no-istoričeskij analiz kategorii opredelennosti – neopredelennosti v slavjanskix, baltijskix i drevnebalkanskix jazykax v svete indoevropėistiki i nostratiki." In: T. M. Nikolaeva, ed. *Kategorija opredelennosti – neopredelennosti v slavjanskix i balkanskix jazykax*. Moskva: Nauka: 11–63.
- Keidan, A. (2013). "Branch-crossing Indo-European isoglosses: a call for interest." *Indoevropskoe ŗzykoznanie i klassičeskaŗ filologiŗ* 17: 406–417.

- Konow, S. (1949). *Primer of Khotanese Saka. Grammatical sketch, chrestomathy, vocabulary, bibliography*. Oslo: H. Aschehoug.
- Kuryłowicz, J. (1975). „Die Flexionen des Adjektivs im Baltischen und im Slavischen“. In: *Esquisses Linguistiques II*. Hrsg. von E. Coseriu. München: Fink: 426–438.
- Lavency, M. (1998). *Grammaire fondamentale du latin. V.2. La proposition relative*. Louvain & Paris: Peeters.
- Lehmann, C. (2018). „Adjective and attribution. Category and operation.“ In: C. Baumann et al., eds. *Adjektive. Grammatik, Pragmatik, Erwerb*. Berlin: De Gruyter: 13–76.
- Lehmann, W. P. (1957). „Structure of Noun Phrases in German.“ In: L. Dostert, ed. *Eighth Annual Round Table Meeting on Linguistics and Language Studies. Research on Machine Translation*. Washington, D.C.: Georgetown University Press: 125–134.
- Leskien, A. (1876). *Die Declination im Slavisch-Litauischen und Germanischen*. Leipzig: Hirzel.
- Luján, E. R. (2009). „On the grammaticalization of \**k<sup>w</sup>i-|k<sup>w</sup>o-* relative clauses in Proto-Indo-European.“ In: V. Bubeník et al., eds. *Grammatical change in Indo-European languages*. Amsterdam & Philadelphia: Benjamins.
- Maggi, M. & P. Orsatti (2018). „From Old to New Persian.“ In: A. Sedighi & P. Shabani-Jadidi, eds. *The Oxford Handbook of Persian Linguistics*. Oxford: Oxford University Press: 7–51.
- Matras, Y. (2007). „The borrowability of structural categories.“ In: Y. Matras & J. Sakel, eds. *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter: 31–73.
- McFadden, T. (2004). „On the pronominal inflection of the Germanic strong adjective.“ *University of Pennsylvania Working Papers in Linguistics* 10.1: 123–136.
- Meillet, A. (1906-1908). « La phrase nominale en indo-européen ». *Mémoires de la Société de Linguistique de Paris* 14 : 1-26.
- Meillet, A. (1915). *Grammaire du vieux Perse*. Paris : Guilmoto.
- (1934). *Le slave commun*. Sous la dir. d'A. Vaillant. 2<sup>e</sup> éd. Paris : Champion.
- Meyer, R. (2017). „Iranian-Armenian language contact in and before the 5<sup>th</sup> century CE. An investigation into pattern replication and societal multilingualism.“ PhD thesis. Oxford: University of Oxford.
- Morpurgo Davies, A. (2012). „Phonetic laws, language diffusion, and drift: the loss of sibilants in the Greek dialects of the first millennium BC.“ In: P. Probert & A. Willi, eds. *Laws and Rules in Indo-European*. Oxford: Oxford University Press: 103–126.
- Nevis, J. A. & B. D. Joseph (1993). „Wackernagel affixes: evidence from Balto-Slavic.“ In: C. Booij & J. von Marie, eds. *Yearbook of Morphology 1992*. Dordrecht: Kluwer Academic Publishers: 93–111.
- Nichols, J. (1986). „Head-marking and dependent-marking grammar.“ *Language* 62.1: 56–119.
- Orr, R. (1983). „The Twofold Adjective Declension in Germanic and Slavic (with some Reference to Baltic): A contrastive/comparative Analysis.“ *Zeitschrift für vergleichende Sprachforschung* 96.1: 104–118.
- Parenti, A. (1996). „L'aggettivo determinato lituano e l'evoluzione delle marche di definitezza.“ In: R. Benacchio et al., eds. *Determinatezza e indeterminatezza nelle lingue slave*. Padova: Unipress: 27–42.

- Patri, S. (2001). "Une correspondance irano-slave en novgorodien." *Historische Sprachforschung* 114.2: 284–297.
- Petit, D. (2009). "La préhistoire des adjectifs déterminés du baltique et du slave." *Bulletin de la Société de Linguistique de Paris* 104: 311–360.
- Pisani, V. (1950). "Studi sulla fonetica dell'armeno. I. Palatalizzazioni ed esiti di s + gutturale; esiti delle semivocali." *Ricerche Linguistiche* 1: 165–193.
- Pischel, R. (1900). *Grammatik der Prakrit-Sprachen*. Strassburg: Trübner.
- Polivanova, A. K. (2013). *Staroslavjanskij jazyk. Grammatika. Slovare*. Moskva: Institut Slavjanovedenija.
- Probert, P. (2015). *Early Greek Relative Clauses*. Oxford: Oxford University Press.
- Prokosch, E. (1939). *A comparative Germanic grammar*. Philadelphia: Linguistic society of America.
- Ratkus, A. (2015). "Gothic possessives, adjectives, and other modifiers in *-ata*." *Journal of Germanic Linguistics* 27.3: 238–307.
- Rießler, M. (2016). *Adjective attribution*. Berlin: Language Science Press.
- Rosinas, A. (1975). "Ar baltų \*i-, \*io- resp. \*ī, \*iā- kamienai buvo reliatyviniai?" *Baltistica* 11.2: 165–170.
- Schmitt, R. (2014). *Wörterbuch der altpersischen Königsinschriften*. Wiesbaden: Reichert.
- Schwyzler, E. & A. Debrunner (1939–1950). *Griechische Grammatik*. München: C. H. Beck.
- Seiler, H. (1960). *Relativsatz, Attribut und Apposition*. Wiesbaden: Harrassowitz.
- Sievers, E. (1876). „Kleine Beiträge zur deutschen Grammatik III. Die starke Adjectivdeclination.“ *Beitrage zur Geschichte der deutschen Sprache und Literatur* 2: 98–124.
- Sims-Williams, N. (1990). "Chotano-Sogdica II. Aspects of the development of nominal morphology in Khotanese and Sogdian." In: G. Gnoli & A. Panaino, eds. *Proceedings of the first European conference of Iranian studies, held in Turin, September 7th–11th, 1987*. Vol. 1. Rome: IsMEO: 276–296.
- Sommer, F. (2018). "The historical morphology of definiteness in Baltic." *Indo-European Linguistics* 6: 152–200.
- Strunk, K. (1967). "Wortstruktur und Pronomen im Altpersischen." *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 81.3–4: 265–275.
- Trautmann, R. (1910). *Die altpreussischen Sprachdenkmäler*. Göttingen: Vandenhoeck und Ruprecht.
- Vaillant, A. (1942). «L'article en vieux slave». *Revue des Études Slaves* 20 : 5-12.
- Večerka, R. et al. (1994). *Staroslavjanskij slovar' po rukopisjam X–XI vekov*. Moskva: Russkij jazyk.
- West, M. L. (2011). *Old Avestan Syntax and Stylistics*. Berlin: De Gruyter.
- Wissemann, H. (1957). „Zur nominalen Determination.“ *Indogermanische Forschungen* 63: 61–78.
- Woolner, A. C. (1928). *Introduction to Prakrit*. 2nd ed. Lahore: Punjab University.
- Zinkevičius, Z. (1958). "Nekotorye voprosy obrazovanija mestoimennyx prilagatel'nyx v litovskom jazyke." *Voprosy slavjanskogo jazykoznanija* 3: 50–100.