

Chapter 11

The diachrony of embedding

Carlotta Viti

Université de Lorraine

Embedded finite clauses are usually presented as the most typical form of subordination, since they are dominant in English and in most modern Indo-European (henceforth IE) languages of Europe. However, the situation is quite different if we consider not only other language families and areas, but also Indo-European itself in its early stages, as attested by Hittite, Vedic, Ancient Greek, Latin, etc. In these languages, embedded finite constructions compete with (and are often superseded by) participial and nominalizing strategies on the one hand, and non-embedded forms of clause linkage, such as parataxis and correlative structures on the other hand, so that an increase of embedded finite subordination is often observable in the history of the IE languages. I analyze this morphosyntactic tendency, as well as alternative patterns of change, with consideration of different relative, adverbial, and completive functions. I will show that factors of semantic complexity, as well as more or less formal sociolinguistic registers, also play a role in the selection of embedded or non-embedded constructions.

1 Introduction

Embedded clauses, that is, subordinate clauses (SCs) functioning as a constituent of the matrix clause, present a wide variety of structures, both across different languages and in the subordinating system of a language, which also reflect a variety of diachronic developments. In most cases, embedded clauses develop from non-embedded clauses (although the opposite change is also attested, as we will see). Their diachrony can therefore be better analysed in the framework of subordination and of clause linkage in general. In the following, I will discuss the diachronic change of embedded relative clauses (Section 2), adverbial clauses

(Section 3), and complement clauses (Section 4), before drawing some conclusions in the final paragraph (Section 5). Our materials are especially drawn from the IE languages, not only because these languages represent my main field of expertise, but also because they offer a longer attested diachronic record with respect to other language families.

2 The development of embedded relative clauses

2.1 Main strategies of relativization

An embedded relative clause (RC) modifies a noun phrase of the matrix clause (MC). It is often marked by a relativizer, as the relative pronoun *who* in example (1a). The RC and its nominal head noun form a relative construction (RCON), as indicated in (1b). The nominal antecedent can also be missing, as in (1c); in this case, RC and RCON coincide.

- (1) a. *The girl*_{RC}[*who is reading a book over there*] *is Peter's sister.*
- b. *RCON*[*The girl who is reading a book over there*] *is Peter's sister.*
- c. *RC/RCON*[*Who has written that letter*] *is certainly not a native English speaker.*

Embedded RCs can be distinguished according to whether the head noun (when present) is external or internal to the RC itself and, in the former case, whether they precede or follow their head noun. Moreover, embedded RCs also vary according to the type of relativizer (when present), which can be more or less explicit (cf. Lehmann 1984, Keenan 1985, Smits 1989, Andrews 2007, etc.; cf. also Meinunger & Alexiadou, this volume). Typological studies usually consider external postnominal RCs, as in (1a), to be the most common structure of RC (cf. Keenan 1985: 143ff). From a functional point of view, they also privilege restrictive RCs, which specify the extension of the RC domain, with respect to non-restrictive or appositive RCs, which depend instead on definite head nouns such as proper names or nouns modified by a possessor, e.g. *Peter's sister, who arrived yesterday, likes reading a lot.* In the case of non-restrictive RCs, we often have a pause in intonation between the head noun and the RC, which is graphically represented by a comma in many languages. Some languages also use different relative markers for restrictive and non-restrictive RCs, as in English, where the relative pronoun *that* tends to be used with a restrictive function, or in Italian, where the relative pronoun *il quale* can be only non-restrictive. This syntactic variety of RCs is, however, mitigated by some typological correlations. For example,

it is well-known that the RC is marked by more or less explicit strategies, such as full noun phrases, personal pronouns, relative pronouns or gapping, according to the position of the relativized syntactic function on the Accessibility Hierarchy (cf. Keenan & Comrie 1977). Higher positions, such as subject and direct object, are usually relativized by means of more implicit markers, while relativized functions of direct object, object of preposition/postposition, and possessor usually require heavier anaphoric resumption.

Based on the data of the ancient IE languages, I suggest that what is often considered to be the norm in typological studies of the RC represents a secondary development of alternative subordinating constructions, which were originally not embedded or not even structured as RCs. In particular, embedded RCs can replace, or be derived from, two kinds of relativization strategies, such as participles and correlative constructions, in the history of IE.

2.2 Relativizing participles

The participle is a non-finite verbal form so called by Graeco-Roman grammarians (*metokhé* in Ancient Greek, *participium* in Latin) because it “participates” in both verbal morphology and nominal morphology, in having tenses/aspects like verbs and grammatical cases like nouns. Its case inflection allowed to modify noun phrases with different syntactic functions in the MC. Its verbal inflection also permitted to express different relations of anteriority, posteriority or simultaneity with respect to the MC situation as well as to the speech situation. For Proto-Indo-European (PIE), we can reconstruct at least a present participle, an aorist participle, and a perfect participle, in addition to various verbal adjectives. As a strategy of RC, the participle is considered by Keenan (1985: 144) to be marginal, owing to the fact that it may relativize a limited number of syntactic functions in the languages of Keenan’s database. In German, for example, the present active participle can only paraphrase a RC that relativizes the function of the subject, e.g. *ich sehe den in seinem Büro arbeitenden Mann* = *ich sehe den Mann, der in seinem Büro arbeitet* ‘I see the man who works in his office’.¹ Moreover, this use of the participle is quite marked in German, mainly limited to the written language. When, however, a language has morphological devices available to convert non-agentive semantic roles into the subject function, e.g. a morphologically marked voice system, the productivity of participles may increase. In PIE, while the present participle in *-nt-* (later extended to the aorist) and the perfect participle in *-wos-* are active, the participle in *-meno-* is middle-passive and therefore

¹Instead, German perfect participles are passive and can therefore promote the original object to subject position, e.g. *ich sehe das von Klee gemalte Bild* ‘I see the picture painted by Klee’.

compatible with various non-agentive roles (e.g. Latin *alumnus* ‘nursling, pupil’ from **alo-menos*, the passive participle of *alō*, -ere ‘to nurture’). Moreover, some verbal adjectives, such as the formations in -no- and in -to-, which originally were not linked to any specific tense/aspect, later acquired functions similar to proper participles. The verbal adjective in -to-, in particular, was associated to the passive voice in various IE languages (e.g. Latin *laudātus* ‘praised’ vs. *laudāns* ‘who praises’). This also increased the possibilities of participial relativization.

These possibilities were not equally extended in all IE languages. Hittite, for example, only has a participle in -ant-, which is typically active with intransitive predicates and passive with transitive predicates (e.g. *pānt-* ‘gone’ vs. *kunant-* ‘killed’), and is, more generally, the ancient IE language where participles are minimally productive. This may be related to the fact that, beyond the participle, the whole verbal system is also simple in Hittite, where the finite verb has two tenses (present and preterit) and two moods (indicative and imperative). Ancient Greek, in contrast, has a very elaborate verbal inflection, and may be considered to be the IE language which is also richest in participles – already in the ancient grammatical tradition, Ancient Greek was celebrated for its *philo-metokhē* ‘love for participles’. Especially in Homeric Greek, we can observe how participles represent the favourite strategy to express a RC, in both restrictive and non-restrictive functions, although they may also have numerous other subordinating functions beyond relativization (they may also express adverbial relations of time, condition, cause, concession, purpose, and consecution, in addition to various complement relations, cf. Section 4). An example of this is illustrated in (2), where the participle *eóntas* ‘being’ is translated with a RC in modern European languages such as English. A proper RC, instead, introduced by a relative pronoun, is preferred in Homeric Greek with syntactically complex constructions, as in (3), where the head noun *ámpeloi* ‘vines’ is coordinated in a series of subjects in a right-dislocation.²

- (2) énth’ huías Priámoio dúō lábe Dardanídao ein heni
 then son(M):ACC.PL Priam:GEN two took Dardanus:GEN in one:DAT.M
díphrō eóntas...
 chariot(M).DAT.SG be:PART.PRS.ACC.M.PL
 ‘Then he took two sons of Priam, Dardanus’ son, who were in one
 chariot’ (lit. “being in one chariot”).” (Homer, *Iliad* 5.159–160)
- (3) allà tá g’ ásparta kai anérota
 but this:NOM.N.PL PTC unsown:NOM.N.PL and unploughed:NOM.N.PL

²In these and in the following examples, the / indicates the end of a verse in a poetic text.

pánta phúontai, puroì kai krithai
 all:NOM.N.PL grow:PRES.IND.3PL wheat(M):NOM.PL and barley(F):NOM.PL
 ēd' ámpeloi, hai te phérousin /
 PTC vine(F):NOM.PL which:NOM.F.PL PTC bring:PRES.IND.3PL
 oînon eristáphulon
 wine(M):ACC.SG made.of.fine.grapes:ACC.M.SG
 'But all these things grow there without sowing and ploughing – wheat,
 barley, and vines which bring a wine made of fine grapes.' (Homer,
Odyssey 9.1091-111)

As nominal morphology has decayed to a larger extent than verbal morphology in most IE languages, the adnominal possibilities of the participle have been reduced with time. It may be hypothesized that the change has started in relative constructions characterized by complex head nouns or by complex RCs. Eventually, even head nouns with syntactically short modifiers may have been rendered by relative constructions with a relative pronoun. This is an aspect of the more general change from synthesis to analysis which can be observed in the history of IE as well as of other language families, and which may be related to case loss or case reduction. Relativizing participles may resist longer in languages where participles represent a synchronically productive category and where nominal inflection is complex. The Baltic languages are an example of this. Moreover, contact factors may interact in the (non-)implementation of the syntactic replacement of participial RCs with embedded RCs marked by a relative pronoun. I consider the flexible use of the participle in the Baltic languages, in particular, to represent an old inheritance from PIE and, at the same time, to have been retained thanks to a Finno-Ugric substrate, as well as to a Finnish adstrate, along the lines of Heine & Kuteva (2003) "contact-induced grammaticalization". Finnish and Uralic in general have a very flexible use of the participle, which is often rendered by RCs in English (cf. Viitso 1998: 113).

2.3 Correlative relative clauses

Correlative constructions have been often considered to be a marginal relativization strategy in typological studies in that they do not function syntactically as a noun phrase and are not embedded (cf. Keenan 1985: 163–164, Dryer (2013), etc.) – again, embedding is more or less consciously presupposed as the standard in the literature, as this is the syntactic structure typically used in English. In a correlative construction, the RC is adjoined to the MC. It is adjoined at the periphery (usually the left periphery) of the MC, and it does not take determiners,

case markers, or adpositions, which are typical for a noun phrase. Although the presence of a relative marker, usually a relative pronoun, signals subordination, the adjoined RC of a correlative diptych is not morpho-syntactically deranked. Its verb has a finite form, with no concord in tense or mood with the verb of the MC, and all arguments present the same form as expected in a non-subordinate clause. The RC of a correlative construction often has an internal head. The syntactic linkage between RC and MC is loose, and the MC usually presents a correlative element, typically a demonstrative pronoun with anaphoric function, to resume the relativized element. The structure can be illustrated with “which story you have told me last time, this (story) is very interesting” (not acceptable in English), instead of the corresponding embedded construction “the story you have told me last time is very interesting”. De Vries (2002: 20) analyses the structure of correlative relative clauses as follows:

correlatives [S-matrix [RC (...) N ...] [S-matrix ... (Dem) ...]]

Correlative clauses are favoured in languages having an unmarked SOV major constituent order, especially loose verb-final languages, which allow right-dislocation of syntactically heavy constituents, as in Hindi. They are also common in languages with a syntactically free word order. Instead, rigid SVO or VSO languages do not have a productive use of correlative RCs. Keenan (1985: 165) mentions Sanskrit and Medieval Russian as examples of IE languages with correlative constructions. Lipták (2009a,b) presents further case studies from Dutch, Polish, Basque, Hungarian, Tibetan, etc. It must be remarked that although new publications such as de Vries (2002) and Lipták (2009a,b) have pointed out the importance of correlative RCs, typological studies continue to present them as a minor relativization strategy as compared to externally-headed embedded RCs (cf. Dryer 2013). This is probably due to the fact that typological studies especially rely on synchronic data from modern languages, which in IE, as in other language families, have often developed embedding from earlier patterns of correlation.

Unlike in most modern languages of Europe, correlative RCs represent the norm in early IE. We especially find them in Vedic, as in (4) and in Hittite (5), two of the earliest attested IE languages. It is not by chance that the very term “correlative diptych” (*diptyque correlative*) has been introduced by an Indologist, Armand Minard (1936), who identified various forms of it in the prose of the Śatapathabrāhmaṇa. One of his students, Jean Haudry (1973), later recognized the same construction in more or less residual expressions of other early IE languages.

- (4) *yám* *bhadréṇa* *śávasā* *codáyāsi*
 whom:ACC.M.SG auspicious:INS.N.SG force(N):INS.SG rouse:CAUS.SUBJ2SG
prajāvatā *rádhasā* *té* *syāma*
 rich.in.offspring:INS.N.SG generosity(N):INS.SG these:NOM.M.PL be:OPT.1PL
 ‘Whom you will rouse with auspicious force and with generosity rich in
 offspring – may we be these ones.’ (*Rig-Veda* 1,94,15cd)
- (5) *kuiš* *ammēl āppan* LUGAL-uš *kīšar[i]* *nu*
 who:NOM me:GEN after king:NOM.SG become:PRES.3SG CON
^{URU}Hattušan āppa ašāš[i] *n=an* *nepišaš*
 Hattusha:ACC again settle:PRES3SG CON-him:ACC heaven:GEN
^DIŠKUR-aš ḥazziē[(ttu)]
 stormgod:NOM strike:IMP.3SG
 ‘Who becomes king after me and again settles Hattusha, let the stormgod
 of heaven strike him!’ (CTH 1.A, edited by Neu 1974 in StBoT 18)³

The Vedic sentence in (4), drawn from a hymn to Agni, the god of fire, presents a binary structure of unequal extension. The first member, a headless RC introduced by the relative pronoun *yá-*, identifies a discourse topic (the one whom the god favours), while the subsequent MC presents the resumptive *te* ‘these ones’ and expresses a comment about it (i.e. the speaker wishes to be one of these favourites of the gods). The grammatical inconsistency here observable between the singular relative pronoun and the plural anaphoric pronoun – a construction *ad sensum*, as the RC has a generic referent – underlines the loose connection between RC and MC. The Hittite sentence in (5) is taken from a historical text, the famous celebration of king Anitta (18th or 17th century BC), who destroyed the city of Hattusha and laid a curse on it. The preposed and headless RC, marked by the relative pronoun *kuiš* ‘who’, is resumed by the demonstrative clitic pronoun *-an* in the subsequent MC. The latter is introduced by the typical clause connecting particle *nu* (etymologically a temporal adverb meaning ‘and then, and’, related to Ancient Greek *nûn* ‘now’ and **né(w)os* ‘new’), which marks the syntactic separation between the two clauses.

In Indic, the correlative diptych has remained as the typical relativization strategy from the earliest texts of the Veda up to the modern era. In fact, many modern studies of correlative RCs are based on Hindi, where this construction is unmarked for relativization (cf. Srivastav 1991, Dayal 1996, Bhatt 2003 among

³In Hittite transcription, the sign [...] may indicate material lost in text break. Superscript writing indicates determinatives, that is, signs of the semantic category of the next noun. In (5), the superscript ^{URU} is the determiner of cities, and the superscript ^D is the determiner of gods.

others). That is, no change from adjoining to embedding is necessary. However, *if* a change occurs in the domain of relativization, then the adjoined RC of the correlative diptych turns out to change into an embedded RC much more typically than the other way round, on the basis of my data. That is, an anaphoric resumptive of the relative construction is more easily lost than added in the MC. Since anaphoric resumptives are especially common for RCs that rank low on the Keenan & Comrie (1977) Accessibility Hierarchy, it may be argued that the change from correlative to embedded RCs, if any, starts with RCs having a subject function, where anaphora can be more easily dispensed with.

Further restrictions may apply in embedded RCs, e.g. the RC tends to be adjacent to the head noun, usually postposed to it. The verb and the arguments of the RC, however, continue to be marked independently from the MC, as it would be expected in a non-subordinated clause. The reason is that on the one hand a RC presents an independent time reference with respect to the MC (e.g. *the cake that you ate / are eating / will eat is done with cricket flour*), that is, the tense/aspect of the MC does not automatically condition the tense/aspect of the RC, as it happens in some kinds of complement clauses instead (cf. Section 4). On the other, the RC only concerns a noun phrase of the MC and leaves other arguments or adjuncts unaffected. This change from correlative RC to embedded RC can be seen in the history of Latin, for example. The earliest documents of this language often offer examples of correlative diptychs, as in (6), besides examples of embedded RCs. With time, embedded RCs generalize (7), and correlative diptychs remain only in fixed expressions, e.g. in proverbs.

- (6) SEI. QVES ESENT. QVEI. SIBEI. DEICERENT.
 if any:NOM.M.PL be:SUBJ.IPF3PL RP:NOM.M.PL REFL say:SBJV.IPF.3PL
 NECESVS. ESE. BACANAL. HABERE. EEIS. VTEI. AD.
 necessary to.be Bacchanal:ACC to.have they:NOM.M.PL CONJ to
 PR. VRBANVM ROMAM. VENIRENT.
 praetor(M):ACC.SG urbanus:ACC.M.SG to.Rome come:SBJV.IPF.3PL
 DEQVE. EEIS. REBVS. VBEI. EORVM.
 about.and these:ABL.F.PL thing(F):ABL.PL where of.them
 VER[B]A. AVDITA. ESENT. VTEI. SENATVS
 words(N):NOM.PL heard:NOM.N.PL be:SUBJ.IPF.3PL CONJ senate(M):NOM
 NOSTER. DECERNERET
 our:NOM decide:SBJV.IPF.3SG

‘If there are any who say that it is necessary to have a bacchanal, they have to come to Rome to the praetor urbanus, and the senate has to decide on those things, when their words have been heard.’ (*Senatus Consultum de Bacchanalibus*, CIL I² 581, 3–6)

- (7) Gallia est omnis divisa in partes tres,
 Gaul(F):NOM is all:NOM.F.SG divided:NOM.F.SG in part(F):ACC.PL three
quarum unam incolunt Belgae,
 RP:GEN.F.PL one:ACC.F.SG inhabit:PRES.IND.3PL Belgae:NOM.PL
 aliam Aquitani, tertiam *qui*
 another:ACC.F.SG Aquitani:NOM.PL third:ACC.F.SG RP:NOM.M.PL
 ipsorum lingua Celtae, nostra Galli
 their:GEN.PL language(F):ABL.SG Celts:NOM.PL our:ABL.F.SG Gauls:NOM.PL
 appellantur
 call:PRES.IND.PASS.3PL
 ‘All Gaul is divided into three parts, one of which is inhabited by the
 Belgae, another by the Aquitani, the third by those who are called Celts
 in their own language and Gauls in our language.’ (Caesar, *De bello
 Gallico* 1,1)

The text in (6) is drawn from the *Senatus Consultum de Bacchanalibus*, an Old Latin senatorial decree, engraved on bronze and dating to 186 BC, which prohibited the popular festivals in honour of Bacchus. Here, the relative pronoun *quei* (corresponding to Classical Latin *quī* ‘who’) depends on a previous SC (the conditional ‘if there are any’) and is resumed in the subsequent MC by the demonstrative pronoun *eeis* (corresponding to Classical Latin *īī* ‘these, they’).⁴ The text in (7), instead, belongs to Classical Latin. It presents the beginning of Caesar’s commentary on the Gallic War. Here we have two RCs, both embedded. The former, introduced by the genitive relative pronoun *quārum* ‘of which’, is immediately postposed to its head noun *partēs trēs* ‘three parts’ (lit. “three parts, of which” etc.). The second RC, marked by the nominative relative pronoun *quī* ‘who’, is headless (‘who are called’ etc.). This does not imply that correlative RCs disappear in Caesar or in Classical Latin in general – they do not – or that embedded RCs are not found in Old Latin. As I said, both constructions coexist, partly overlapping and partly preferred in different functional domains. For example, correlative RCs seem to prevail in Latin in restrictive RCs with a generic interpretation, according to my data, while embedded RCs are rather favoured with

⁴The form *eeis*, from a demonstrative stem **eyo-*, represents the hyper-characterization of the regular NOM.M.PL marker *-ī* with the plural ending *-s*, common for many other nominal and pronominal stems, cf. the corresponding Oscan form *insc* “*īī*”. The section of the *Senatus Consultum de Bacchanalibus* in (6) would be rendered in Classical Latin as: *si qui essent, qui sibi dicerent necesse esse Bacchanal habere, īī ut ad pr(aetorem) urbanum Romam venirent, deque iis rebus, ubi eorum verba audita essent, ut senatus noster decerneret*. The correlative elements of the relative pronoun and of the demonstrative pronoun are here in bold.

restrictive RCs which identify a specific referent, as well as with non-restrictive RCs. I rather mean that correlative RCs are more common in Old Latin than in Classical Latin, whereas embedded RCs are more common in Classical Latin.

The correlative diptych decays in Romance. However, it may be interesting to notice that in spoken communication and informal registers, Romance languages such as Italian may still present examples of adjoined RCs. I have recorded the example in (8) in a natural conversation in Italian, where a mother addresses her son, who was used to taking a lot of food without eating it up. The Standard Italian correspondent construction would be embedded (*Devi mangiare la pasta che prendi*, with the RC after the MC). Corresponding constructions are also acceptable in the spoken form of other Romance languages. Spoken French, for example, also admits a sentence like (8) : *les pâtes que tu prends, tu dois les manger* (besides the Standard French construction *tu dois manger les pâtes que tu prends*). This suggests that the choice of an embedded or adjoined RC (the latter being characterized by an anaphoric resumptive) is subject to sociolinguistic variation, in addition to its diachronic and regional underlying factors. Synchronically, sentences like (8) are analysed as left dislocations of an NP with a regular RC plus resumption of the relative construction by a clitic (cf. Cinque 1990). It is important to notice, however, that they represent the same structure as the ancient correlative diptych.

- (8) La pasta che prendi poi la devi mangiare.
DEF.ART pasta:F.SG which you.take then it:F.SG you.must eat
'You must eat the pasta you take.' (lit. "The pasta that you take, then you must eat it.")

2.4 Paratactic constructions with demonstrative pronouns

Correlative RCs are not the only finite constructions that developed into embedded RCs. Alternatively, embedding may arise from the reanalysis of independent clauses, for example, when a demonstrative pronoun with anaphoric function is reinterpreted as a relative pronoun. Accordingly, a construction consisting of two asyndetically linked independent clauses, such as [*I see a man.*] [*That is sitting over there*], may be reanalysed as one complex sentence, formed by a MC and a postposed RC, such as [*I see a man* [_{RC} *that is sitting over there*]]. This is probably the origin of the RCs with the D-type of relative markers in Germanic, that is, of the RCs marked by a relativizer that is etymologically derived from a demonstrative pronoun, e.g. German *der, die, das*. D-relatives represent in fact the most ancient type of relative markers in Germanic – they are dominant not

only in the whole history of German, but also in Gothic and in Old English.⁵ A paratactic structure especially characterizes what Gärtner (2001) calls “integrated verb second” clauses, that is, a type of RC which presents the verb in the second position, like independent clauses and unlike regular RCs (the latter share the same verb-final word order as other subordinates). Gärtner shows how a regular RC (e.g. *das Blatt hat eine Seite, die ganz schwarz ist*) may be replaced by a verb-second integrated clause (e.g., *das Blatt hat eine Seite, die ist ganz schwarz*), and how further syntactic properties may nonetheless distinguish these clauses. Despite the presence of a relative pronoun, verb-second integrated clauses are therefore not really RCs in the syntax – according to Gärtner, they may be better analysed in terms of parataxis or as hybrid constructions between parataxis and hypotaxis. Cf. also Zifonun (2001).

Other IE branches also attest cases of the change from a sequence of independent clauses, with an anaphoric pronoun in the second clause, into a complex sentence having an exocentric and postnominal RC. Homeric Greek, for example, presents RCs marked by a demonstrative pronoun *ho, hē, tó* ‘this’ (M, F, N), the source of the definite article in Classical Greek, besides RCs introduced by the proper relative pronoun *hós, hē, hó* ‘who, which’ (M, F, N), cf. Schwyzler (1950: 19ff; 642). We probably deal with a polycentric syntactic change, independently occurring in different times and in different areas of IE (cf. Heine & Kuteva 2002: 113–115 and Schmidtke-Bode & Diessel, this volume). We therefore do not necessarily have to presuppose the reconstruction of RCs marked by a demonstrative relative marker for PIE.

3 The development of embedded adverbial clauses

3.1 Main strategies of adverbial clauses

An adverbial clause (AC) is a subordinate clause that modifies the MC like an adverb modifies the verb and that therefore has the syntactic function of a circumstantial element in the complex sentence. ACs are typically embedded, that is, they represent a constituent of their complex sentence, as in (9), where different ACs are written in brackets for different adverbial relations. Like adverbs,

⁵In Germanic, the ancient correlative construction has only left few traces, and other kinds of relative markers, such as W-relatives (derived from an interrogative-indefinite pronoun, e.g. English *who, which*) and 0-relatives or gapping, emerge later and are also synchronically more constrained (cf. Helgander 1971, Harbert 2007: 420ff. and more recently Bacskai-Atkari 2020, among others). For the development from correlative adjunction to embedding in Old English and Middle High German, cf. Bianchi (1999), who uses this diachronic scenario to justify Kayne’s antisymmetric view of syntax.

ACs are semantically heterogeneous and may code relations of time, as in (9a), condition (9b), cause (9c), concession (9d), purpose (9e), consecution (9f), etc.

- (9) a. Temporal AC: [*After arriving home*], *Peter takes a shower*.
- b. Conditional AC: [*If he arrives late*], *Peter skips his dinner*.
- c. Causal AC: [*Since it is raining*], *Peter stays at home*.
- d. Concessive AC: [*Although it is raining*], *Peter goes out*.
- e. Purposive AC: *Peter goes out* [*to buy some food*].
- f. Consecutive AC: *It rained a lot*, [*so that the street is completely wet*].

ACs are quite diverse from a formal point of view, as they may be marked by a particular subordinating conjunction, a form of the verb which is different from a verb in an independent sentence, or a special word order (cf. Thompson 1985). In (9), for example, we have both finite and non-finite subordinates, in the latter case marked with a participle, as in (9a) or with an infinitive, as in (9e). However, while adverbial relations of time, condition, cause, etc. are universal, ACs are not. In some languages, these relations may be left implicit and recovered from the context. In some others, adverbial relations may be explicitly expressed by non-subordinating strategies, such as coordination or clause-chaining. Embedded and non-embedded constructions may compete in the same language for the expression of an adverbial relation. The relation of cause, for example, can be rendered in English not only by embedded ACs marked by subordinators such as *since*, *because*, *as*, etc. but also by the conjunction *for*, which rather belongs to coordination. This can be clearly seen in German, where SCs and matrix/independent clauses are associated to different word orders, and where the particle *denn*, the functional equivalent of English *for*, selects the verb-second word order typical of independent clauses, unlike causal subordinators such as *weil* or *da*. Of course, subordinating and coordinating adverbial relations are not synonymous, and their usage is conditioned by various semantic and pragmatic features. For example, coordinating conjunctions of cause, such as English *for*, German *denn*, French *car*, etc., introduce focused information, while their subordinating counterparts commonly mark clauses representing a presupposed situation. That is, clauses introduced by subordinating connectives are usually backgrounded in their entirety, while clauses with non-subordinating connectives are not. Moreover, subordinating and non-subordinating strategies of adverbial relations may go back to different diachronic stages. As in the case of RCs, for ACs as well, we can observe that non-embedded constructions are generally used earlier than embedding, and are modified into (or replaced by) embedded structures only at

a later stage of a language. In my IE data, the earliest constructions used to express adverbial strategies are RCs (Section 3.2) and parataxis (Section 3.3), the latter presenting asyndeton or coordinating conjunctions.

3.2 Correlative adverbial clauses

The adjoined structure of an RC in the form of the correlative diptych, illustrated above in Section 2, can also be used to convey various adverbial relations. In this case, an adverbial subordinator of time, cause, purpose, etc. in the ACs is correlated with a semantically corresponding demonstrative adverb in the MC, e.g. “When he arrives home, then he watches the news” or “The sooner you’ll come, the quicker we will be finished”. This old correlative pattern is attested in most early IE languages. Axel-Tober (2023), for example, suggests that, in Middle High German, correlative *da*-clauses were reanalysed as embedded adjunct clauses because left-peripheral correlatives are surface-structurally identical to left-dislocated embedded clauses. Correlative adverbial clauses are still quite productive in Lithuanian, a language known for being conservative in several aspects of its grammar. Lithuanian presents correlations not only between *tàs* ‘that’ and *kàs* ‘who’ in the domain of relativization, but also between various adverbial elements such as *taĩp* ‘so’ and *kàd* ‘that’, *tíek* ‘so much’ and *kíek* ‘how much’, *tadà* / *tuomèt* ‘then’ and *kaĩ* / *kadà* ‘when’, *tòl* ‘until’ and *kòl* ‘until’, *teĩ* ‘there’ and *kuĩ* ‘where’, etc. According to Ambrazas (1997: 723), these correlations “serve to stress and reinforce the link between the clauses. In a way, they represent the subordinate clause in the structure of the main clause”. Among the many examples, he reports the following (1997: 738):

- (10) *kalbèsiu tiesą ir tada, kai pavojinga*
 speak:FUT.1SG truth:ACC even then when dangerous
 ‘I will speak the truth even when it is dangerous.’ (lit. “even then when”
 etc.)

My interpretation is different from that of Ambrazas (1997). I do not think that correlative structures have the function of reinforcing the linkage between SC and MC, otherwise we cannot explain the difference from other ACs favouring embedding, which represents a more integrated clause linkage with respect to correlation. I rather maintain that they serve to express separation and balance between the situations represented in the two clauses.⁶

⁶It may be remarked that the sentence in (10) would be much better with the main verb *sakysiu* (the FUT.1SG of *satyti* ‘say’, a cognate of English *say* and German *sagen*). The omission of the correlative element *tada* is possible but considered to be less natural than the correlative variant above (Daniel Petit, p.c.).

If we consider the functional load of correlative constructions, we can observe that they are not equally common for all adverbial relations. Diachronically, they turn out to be the structural source especially for relations of space (introduced in English by the subordinator *where*), time (cf. English subordinators such as *when*, *while*, *as*, *before*, *after*, *since*, *until*, *once*, *as soon as*, etc.), and manner (cf. English subordinators such as *like* and *as*), in both ancient and modern IE languages. In Ancient Greek, for example, a language where embedded ACs largely prevail over adjoined ACs in general, correlative structures are especially found for 1) expressions of location (*hoû / hó pou / hó thi* ‘where’ ... *én tha / enthá de / entaú tha / ekeî / autoû* ‘there’), direction (*hoî / hó poi* ‘where’ ... *ekeî se* ‘there’, etc.), origin (*hó then / hopó then* ‘from where’ ... *én then / enthé nde / enteú then / ekeî then* ‘from there’), perlocative movement (*hêi / hó pēi* ‘through where’ ... *têide / taútēi* ‘through there’) – all expressing space; 2) expressions of time (*hó te / hopó te* ‘when’ ... *tó te* ‘then’ or *tē ní ka / autí ka* ‘immediately at that point’) and age (*hē ní ka / hopē ní ka* ‘at which age’ ... *tē ní ka / tē ní ka de / tē ní ka ú ta* ‘at that age’); 3) expressions of manner (*hō s / hó pō s* ‘how’ ... *hō s / hó de / hó ú tō s* ‘so, like that’). Now, ACs of space, time, and manner are considered to be simple or primary in typological studies with respect to other ACs, since they can be paraphrased by an RC modifying a head noun of space, time, or manner (cf. Thompson 1985: 177ff). For example, the temporal AC *When he arrives home* can be reformulated as the relative construction *The time (at which) he arrives home*. Similarly, the space AC *Where you can buy croissants* can be reformulated as the relative construction *The place at which you can buy croissants*. The manner AC *How you do it* can be reformulated as the relative construction *The way (in which) you do it*, etc. As correlative structures represent a major source of embedded RCs, it is understandable that variants of the correlative diptych also serve to form those ACs which are semantically associated with RCs, as in the case of space, time, and manner. Alternatively, among other types of ACs, correlative constructions especially emerge for consecutive clauses, which express the consequence of the situation represented in the MC. The consecutive relation structured as an embedded AC in (9f) may have a correlative variant: *It rained so much that the street is completely wet*. The latter may be even more natural than the embedded construction in (9f).

I find it interesting that, although in most IE languages adverbial correlative structures have been superseded by embedded ACs introduced by an adverbial subordinating conjunction, they especially persist precisely in the two domains of space, time, and manner ACs on the one hand, and of consecution on the other. This can be seen in Italian, for example. In Italian, as in other Romance languages and in Latin already, embedded ACs are much more frequent than correlative ACs. However, correlative ACs are common in the specific domain

of consecutive clauses (*così ... che* ‘so ... that’, *tanto ... che* ‘so much ... that’) and of comparative clauses (*tanto ... quanto* ‘so much ... how much’, *tale ... quale* ‘as ... such’, *più ... che* ‘more ... than’, *meno ... che* ‘less ... than’, etc.), where the form *che* is to be analysed as an adverbial subordinator rather than as a complementizer. Comparative clauses are a kind of subordination of manner, whereby the situation of the MC is presented as equal, major or minor, better or worse than the situation of the subordinate with respect to a certain parameter, as in (11).

- (11) a. Il paesaggio era così bello come lo
 the landscape be:IPF.IND.3SG so beautiful how it
 ricordavo.
 remember:IPF.IND1SG
 ‘The landscape was as beautiful as I remembered.’
- b. Alla fine è andata meglio di quanto (non) mi
 in.the end be:PRES.IND3SG gone better than how.much NEG me
 aspettassi.
 expect:IPF.SBJV.1SG
 ‘In the end, it went better than I had expected.’

Comparative clauses intrinsically have a binary structure, which is properly underlined by the correlative pattern. The presence of a correlative element in the MC usually indicates that the sentence is not concluded and that another clause will follow (in (11b) as well, the MC *alla fine è andata meglio* implies the mention of a situation that has the function of the standard of comparison). The position of the two correlatives may vary – they can be adjacent, as in (11b) *meglio di quanto* ‘better than that’, or non-adjacent, as in (11a) *così ... come* ‘so ... how’ – but syntactically one belongs to the MC and one to the AC. Complex sentences with consecutive clauses also have a semantically balanced structure, describing in this case a relation of cause-effect, as in (12), again from Italian.

- (12) Ha passato così tante ore al computer che gli fanno
 has spent so many hours at.the computer that to.him make:PRES.IND
 male gli occhi.
 bad 3PL.the eyes
 ‘He has spent so many hours at the computer that his eyes are hurting.’

On the other hand, other ACs are less prone to be conceptualized in a binary, balanced structure. They usually express a situation that is presented against the

background of the MC situation. Purposive clauses, for example, are semantically similar to consecutive clauses, in that both refer to a situation that follows the MC situation. However, while consecutive clauses are factual, purposive clauses are not. They refer to an aim or a wish of the MC subject. The two clauses clearly are on different levels. This may be the reason why other ACs may lose the ancient correlative structure and are embedded earlier than consecutive clauses. According to my data, conditional clauses are the ACs that are least associated with correlative constructions in the ancient IE languages. That is, unlike many other ACs, clauses introduced by a conditional marker such as *if* usually do not trigger a resumptive such as *in that case*, *then*, or *so*, in the MC (while correlative elements such as ‘then’ are quite common in MCs connected to temporal subordinates, for example). We have seen in Section 2 that correlative constructions are especially common in Hittite relativization, and that the typical RC in Hittite is adjoined and not embedded. The same holds for most Hittite ACs – but not for conditional clauses. This can be seen in the laws, in which a conditional protasis and its following apodosis represent an offense and its punishment, respectively, as in “If anyone blinds a male slave in a quarrel, he shall pay 20 shekels of silver” (cf. KBo 6.4). In these cases, the conditional marker ‘if’ (*takku* in Old Hittite, *mān* especially in New Hittite) is usually not accompanied by a correlative element (and in Old Hittite neither by an introductory conjunctive particle) in the MC, cf. Friedrich (1959: 88); Hoffner (1997: 12).⁷ Among the modern IE languages, even in a language rich in correlative clauses such as Lithuanian, the conditional conjunction *jėi* / *jėigu* ‘if’ does not trigger a correlative. Condition is a logical operator (although the use of conditional clauses cannot be reduced to logical meanings). A conditional clause typically expresses an implication or a prediction that affects the realization of the MC situation. One situation depends on another. They are not represented on the same conceptual level. This may explain why conditional clauses are much more commonly embedded than structured in the adjoined structure of a correlative diptych.

3.3 Paratactic adverbial clauses

As anticipated above, in addition to correlation, ACs can also develop from originally paratactic constructions. The juxtaposition or coordination of two clauses referring to a series of chronologically ordered events can be easily reinterpreted as a temporal AC, in particular as a *when*-clause or as an *after*-clause (*Peter goes*

⁷In New Hittite, instead, the introductory conjunctive particle *nu* is used with conditional clauses as well, by analogy with the dominating correlative pattern used in subordination in this language.

to the restaurant. He orders lasagne > When/after Peter goes to the restaurant, he orders lasagne). When the situation of one of the two clauses is meant as depending on the occurrence of the other, we may also have the interpretation of a conditional clause (*If Peter goes to the restaurant*, etc.). The logical fallacy “post hoc ergo propter hoc” brings about the reinterpretation of a temporal clause into a causal clause (*Since/as Peter goes to the restaurant*, etc.). The etymology of adverbial subordinators may bear traces of these reinterpretations. In English, the subordinator *since* originally expressed a temporal relation (it derives from Old English *sipþan* ‘afterwards, since, ago’ < *sīþ þāem* ‘after that (time)’), and only later did it subsume a causal relation. In German, the subordinator *wenn* (cognate of German *wann* ‘when’) marks habitual temporal clauses and conditional clauses. An adverbial relation that implies another adverbial relation is semantically more complex. A causal relation is more complex than a temporal relation because a causal relation implies a temporal relation of anteriority, while a temporal relation does not necessarily entail a causal component. According to my data, the use of embedded rather than paratactic ACs is established according to a cognitive path of increasing semantic complexity. Embedding emerges earlier for temporal ACs (both *after*-clauses and semantically less specific *when*-clauses), for example, than for causal clauses. We have already seen that causal relations are still represented in most modern IE languages by coordinating conjunctions, such as English *for*, German *denn*, French *car*, Italian *infatti*, etc., besides the corresponding adverbial subordinators. Nowadays, coordinate causal clauses are rather marginal in comparison to subordinate causal clauses. In English, *for*-clauses are rarer than *as*-, *since*- or *because*-clauses. In the ancient languages, instead, we have the opposite situation. In Vedic texts, for example, causal clauses introduced by the particle *hi*, the functional correspondent of English *for* or French *car*, are much more common than causal clauses marked by the subordinator *yád* (cf. Viti 2007b: 134ff).

3.4 Factors of semantic complexity

Different degrees of semantic complexity can also be identified within the domain of temporal relations. *When*-clauses and *after*-clauses refer to states of affairs that are anterior to, or simultaneous with, the MC situation. They are therefore semantically less complex than *before*-clauses, which locate the MC situation with respect to an event that has *not* yet happened. The sentence *Before I go to work I take a rich breakfast* implies that, when I take my breakfast, I have not yet gone to work. Occurred events are much more natural as temporal frames than non-occurred events. This may explain why the explicit coding of *before*-clauses

by embedded strategies gets established relatively late. Vedic, for example, possesses several subordinating structures for *when*- and *after*-clauses, introduced by (non-synonymous) subordinating conjunctions such as *yád, yadā́, yádi, yád*, etc., which are also used for relations other than temporal. Despite their polysemy, however, Vedic temporal subordinators do not express *before*-clauses. Although we must be careful in evaluating the absence of a structure from a dead language – it may occur that the structure is not attested by chance – we have to consider that Vedic is attested by a huge corpus. When we observe, therefore, that *when*-subordinate clauses (more commonly meant as *after*-clauses, in the sense of temporal anteriority, than in the sense of temporal simultaneity in Vedic) are largely attested while *before*-subordinate clauses are not attested at least as a productive strategy, we can well infer a different weight of these subordinate clauses in this language. This does not mean that relations of temporal posteriority lack expression in Vedic. They were rather expressed by non-subordinating strategies. An embedded sentence such as *The sacrificer eats before other men have eaten* would be rendered in this language by paratactic constructions such as *The sacrificer eats; other men have not eaten*, or by a single clause where this temporal relation is coded on an adjective or adverb meaning ‘first, anterior, before’. Similarly, Śatapathabrāhmaṇa 1,1,1,8 defines as unbecoming the situation of somebody ‘who takes food before men (i.e. his guests) have eaten’: *yó manuṣyēṣv ānaśnatsu pūrvo ’śnīyād*, lit. “who takes food earlier than men that do not eat”.⁸ The compositional and often transparent form of subordinators of *before*-clauses confirm their relatively late emergence. In Classical Sanskrit, for example, *before*-clauses receive an explicit marking employing the polyrhematic conjunction *yāvan na*, where the negation *na* precisely expresses that the denoted situation has not yet happened. This complex subordinator usually has a correlative element *tāvat* in the matrix clause. Since the correlation *yāvat ... tāvat* typically means “as long as ... so long”, in Vedic already, the Classical Sanskrit complex marker of *before*-clauses literally means ‘as long as not ... so long’. That is, even at this relatively late stage, *before*-clauses are not embedded. They are adjoined in the form of the correlative diptych, as in (13).

⁸Here the adjective *pūrva*- ‘in front, previous to’, inflected in the nominative in agreement with the clause subject, governs the locative complement *manuṣyēṣv ānaśnatsu* ‘men that do not eat’. Alternatively, if the latter phrase is interpreted as a locative absolute, the sentence above may be literally translated as “who eats first, while men do not eat”. Absolute constructions are usually motivated by clausal syntax in Vedic, in this case by the locative meaning implied by *pūrva*- (with a metaphor from space to time). In any case, the relation of temporal posteriority is expressed by an adjective of time rather than by subordination in this passage.

- (13) tad yāvan na kaścīd vetti tāvac chīghraṃ
 then as.long.as NEG anybody:NOM know:PRES.IND3SG so.long quickly
 gamyātam iti
 go:IMP.PASS.3SG QUOT
 ‘Then before anybody knows (of it), go away quickly!’ (*Pañcatantra* 1.276,
 edited by Ramchandra Jha, 6th edition, Varanasi, Vidyabhavan (1991))

Similarly, Hittite possesses many more resources for *when*-clauses (introduced by the subordinators *mān* in Old Hittite, and *mahḥan* and *kuwapi* in New Hittite), as well as for *while*- and *until*-clauses (marked by the subordinator *kuitman* in both stages of Hittite) than for *before*-clauses. The latter may be expressed by some uses of the relatively recent composite subordinator *kuitman ... nāwi*, lit. ‘while ... not yet’, also in the opposite order *nāwi ... kuitman*. E.g. *kuitman ... nāwi eṣḥat* ‘before I sat’ (KBo III 4 I 3f), which some scholars still prefer to render ‘as long as I had not yet sat’ (cf. Friedrich 1960: 138). The use of expletive negation with *before*-clauses can find several linguistic parallels. Krifka (2010) analyses this phenomenon in German sentences such as *Bevor wir nicht gegessen haben, gehen wir nicht nach Hause* ‘Before we have eaten, we do not go home’. He relates this construction to the specific type of negation: expletive negation is possible with a negation working as an adverbial modifier, as German *nicht*, but is not possible in English, where negation is expressed by a separate head such as the *do*-support. The latter, Krifka observes, is rather tied to assertion.

A semantic component of negation is also implied in concessive clauses, which express a contrast with respect to a certain presupposed situation. The concessive sentence *Although it is raining, Peter goes out*, reported in (9d), presupposes that under normal conditions, people do *not* go out when it rains. In the domain of coordination, a similar relation is conveyed by the adversative conjunction *but*. It is therefore understandable that concessive relations, as other semantically complex adverbial relations, are coded by embedded SCs at a relatively late stage of a language. This can also be seen in the transparent form of their subordinators, which are often polymorphemic, or even polyrhematic, containing expressions of negation, contrast, or focus. In the modern languages of Europe, consider concessive subordinators such as English *although*, *even though*, German *obwohl*, *wenngleich*, French *bien que*, *malgré que*, Italian *nonostante*, Spanish *aunque*, etc. (cf. König 2006). A similar situation can be observed in the ancient IE languages, where concessive subordinators are also synchronically transparent. Moreover, in the ancient languages, embedded SCs explicitly marked by a concessive subordinator are much rarer than implicit subordination, especially

represented by the participle; they are also rarer than coordination or juxtaposition, optionally reinforced by discourse particles, for the concessive function. Originally, a complex sentence such as *Although they were few, they won the war* would have been better reformulated into a participial structure such as *Being few, they won the war* or into paratactic constructions, such as *They won the war. And yet they were few*, or *They were few, but they won the war*. Correlation was also possible in the form *Although they were few, nevertheless they won the war* – recall that correlative structures are not embedded.

Even Ancient Greek, which we have seen to privilege embedding over correlation in general, attests concessive subordinators formed based on the coordinator *kaí* ‘and’, cf. *kai ei* and *kai eán*, meaning “although, even though”, as well as participial constructions marked by *kaíper*, *kai mála*, *kai taúta*, etc. In these cases, Ancient Greek *kaí* can also have the function of an additive focus adverb, in addition to its basic coordinating function. In Hittite, the polysemic subordinator *mān*, meaning ‘when, if, although’ typically presents a correlative structure in concessive clauses, such as *mān ... (y)a* ‘although ... nevertheless’. Similarly, in Vedic, concessive subordinates introduced by the composite adverbial conjunction *yác cid dhí* ‘although’ (formed by the relativizer *yád*, the indefinite particle *cid*, and the causal coordinator *hí*) occur more rarely with concessive participles, optionally marked by particles. In Classical Sanskrit, the concessive marker *yady api* ‘even if, although’ (based on the subordinator *yadi* ‘when, if’ with the focus particle *api* ‘and, also, even, very’) is usually accompanied in the MC by the correlative *tathāpi* ‘nevertheless’ (formed correspondently by affixing the particle *api* on the demonstrative adverb of manner *tathā* ‘in that manner, so, thus’). Again, Old Indic prefers the adjoined structure of the correlative diptych to embedding to code an adverbial relation, as can be seen in (14), illustrating this time a concessive conditional. All this indicates a syntactic change from parataxis or correlation to embedding for concessive relations.

- (14) **yady api** tvadiya-vacanam na karoti **tathāpi**
 although your-word(N):ACC.SG NEG make:PRES.IND3SG nevertheless
 svāmī sva-doṣa-nāśāya vācyaḥ
 master(M):NOM.SG own-fault-destruction:DAT to.be.blamed:GV.NOM.M.SG
 ‘Even if he does not listen to your words, you must blame your master, in
 order that he may amend his faults.’ (*Pañcatantra* 1.171)

3.5 Adverbial infinitives

This example can also illustrate the early non-subordinating coding of another type of adverbial relation, that is, purpose. In (14), the purposive sentence of the

English translation *in order that he may amend his faults* is expressed in Classical Sanskrit by the dative of an action noun (in this case, a compound literally meaning ‘for the destruction of his own fault’). The verbal noun or infinitive was the primary expression of purposive relations in Old Indic, as well as in other early IE languages, and in PIE. Embedded subordinates with a purposive function are also attested, but are secondary to the infinitive in IE. In Vedic, embedded purposive subordinates marked by conjunctions such as *yád* or *yáthā* are much rarer than purposive infinitives. In IE in general, purposive SCs are formed with language-specific strategies, which cannot be reconstructed for PIE.

It must be stressed that the infinitive was originally more a noun than a verb in IE, according to a scalar view of parts of speech, as it was marked for case, like nouns, but not for tense, mood or voice, like verbs. In Old Indic, the infinitive usually had the case ending of the dative, which is called *sampradāna*- ‘bestowal’ in the Sanskrit grammatical tradition (cf. Pāṇini I.4.32), and which expresses the recipient, the beneficiary, and the aim of an action. The subject of the purposive infinitive is usually the same as the subject of the MC situation, and is therefore left implicit.⁹ For the direct object of the infinitive, if any, we have more variation. In (14), the direct object is coded as the first member of a compound. In Vedic, where compounding is much more restricted than in Classical Sanskrit, we would rather have an infinitive with a genitive object or an accusative object (*sva-doṣasya nāśāya* or *sva-doṣam nāśāya*). The genitive is the form expected in a possessive structure, and represents the originally nominal syntax of the verbal noun, while the accusative indicates the acquisition of a verbal syntax for this category. Irish, where verbal nouns are extremely productive from the earliest texts up to now, only maintains the objective genitive in this use.

Later, on the one hand, the IE infinitive attained an increasing number of verbal features in syntax and morphology, e.g. tense and voice marking, as in Greek and Italic. Once integrated in the verbal system, the infinitive could express many more subordinating relations (especially in the domain of complementation, as we will see below) which were originally impossible for this category. This may be because a verb is more typically used to represent events than a noun. It is

⁹In (14), the different subjects of the MC (“you”) and of the purposive clause (“in order that he” etc.) are only shown in the translation, as the passive construction would be unnatural in English in this case. In the Sanskrit original, however, the name of the master (*svāmī*) is the subject of a nominal passive clause, expressed by the gerundive *vācyaḥ* ‘that is to be blamed’ (‘the master is to be blamed, in order that he may destroy’ etc.). That is, the MC and the action noun share the same subject. In case they differ, we can have a double dative (a sentence such as *The Hotṛ sings a poem in order that Indra may drink the soma* would be rather formulated as *The Hotṛ sings a poem for Indra, for drinking the soma* (or for drinking of the soma with a genitive complement).

therefore understandable that verb-like infinitives may be more prone to express clausal meanings. On the other hand, purposive embedded clauses introduced by subordinating conjunctions also become more common with time, especially in the case of different subjects between MC and SC. We can therefore observe the development from a circumstantial element or a clause adjunct expressing aim, syntactically included into the MC in origin, to an embedded subordinate. It may be argued that this relatively late use of embedding for purposive relations can also be explained by their semantic complexity, as for other ACs analysed above. Similarly to *before*-clauses, purposive clauses refer to a situation that has not yet occurred at the time the MC situation takes place. The sentence *Peter goes out to buy some food*, illustrated in (9e), implies that the action of buying some food has not yet happened when Peter goes out. Also in this case, it is more natural to frame a situation with respect to an already occurred event than with respect to an event which has not yet happened, and which we do not know whether it will happen, as purposive clauses are not factual. Also in this case, as for *before*-clauses, it is understandable that embedded adverbial clauses are neither the earliest nor the most common strategy attested for purposive clauses.

4 The development of embedded complement clauses

4.1 Main strategies of complementation

Similarly to ACs, complement clauses (CCs) are also extremely diverse in meaning and, consequently, in structure – not only across different languages, but also within the same language. They may depend on different complement-taking predicates, such as predicates of utterance (cf. English *say, tell, ask* (in order to know), *report*, etc.), propositional attitude (*think, believe, suppose, assume, doubt, deny*, etc.), knowledge (*know, discover, realize, find out, forget*, etc.), immediate perception (*see, hear, feel*, etc.), pretence (*pretend, imagine, trick (into thinking)*, etc.), fearing (*fear, be afraid, worry, be anxious*, etc.), achievement (*manage, chance, dare, happen to, try, fail, avoid*, etc.), as well as on commentative predicates (*regret, be odd, be important*, etc.), desiderative predicates (*want, desire, wish, hope*, etc.), manipulative predicates (*make* someone do something, *let* something occur, *force, persuade, ask* (to do), etc.), modal predicates (*can, be able, may, must, should*, etc.), phasal predicates (*start, begin, continue, stop, finish*, etc.) (cf. Noonan (1985), Cristofaro (2005: 95ff), Dixon (2006)). Different complement-taking predicates may select different types of CCs, represented by embedded subordinate clauses, paratactic clauses, nominalizations, participles, as well as by morphological strategies. Sometimes, even the same complement-taking predicate may

appear in more than one construction, as in the following examples, where the utterance predicate *say* is associated with a paratactic structure (15a), with an embedded subordinate clause characterized by a complementizer and a finite verb (15b), and with an embedded infinitive (15c), the latter with a passive verb in the MC. Competing structures are not synonymous and may imply formal adjustments in the expression of the subordinate situation. The passive construction in (15c), for example, is used when one wants to demote the agent of the complement-taking predicate, as is typically the case for this grammatical voice. The two variants of parataxis (15a) and of embedding with a finite subordinate (15b) differ in the absence or presence of anaphoric shifters, which are typical of the indirect speech. Moreover, direct speech tends to be more common in the spoken or informal language, while indirect speech and subordination in general have a tighter connection with the writing and with the formal register (cf. Givón 1979: 207ff, Coulmas 1986, Chafe 1994: 215ff, etc.).

- (15) a. *Peter said to Sandra: ‘You are a good friend.’*
 b. *Peter said to Sandra that she was a good friend.*
 c. *Sandra is said to be a good friend.*

Parataxis, finite subordination, nominalizations, and participles are all attested for complement relations since the earliest stages of IE (cf. Viti 2015: 315ff). As we have seen for RCs and ACs, CCs as well often attest a development towards embedding in the history of a language. This development, however, proceeds at a different pace for different complement-taking predicates. According to my IE data, the earliest embedded constructions attested for complementation are implicit CCs represented by participles depending on predicates of immediate perception (Section 4.2).

4.2 Non-finite, participial complement clauses

Despite their amazing variety in the expression of complement relations, most early IE languages agree in their constructions of predicates of immediate perception such as ‘see’ or ‘hear’, as in the following examples from Vedic (16a), Homeric Greek (16b), and Old Latin (16c), where the participial clause is a syntactic constituent embedded in the MC.

- (16) a. yátredānīm pásyasi jātavedas
 where.now see:PRES.IND.2SG Jātavedas:VOC
 tíṣṭhantam agna utá vā
 standing:PART.PRES.ACC.M.SG Agni:VOC and or

cárantam / yád vāntárikṣe
 moving:PART.PRES.ACC.M.SG when or.atmosphere(N):LOC.SG
pathíbhiḥ **pátantam** tám
 path(M):INSTR.PL flying:PART.PRES.ACC.M.SG him:ACC.M.SG
 ástā vidhya śárvā
 archer(M):NOM.SG pierce:IMP.2SG dart:INS.SG
 śísānaḥ
 sharpen:PART.PRES.NOM.M.SG

‘Where you see him now standing or moving, o Agni Jātavedas, or flying along the paths in the atmosphere, pierce him, as archer, with your dart, sharpening (it).’ (Rig-Veda 10,87,6)

- b. ò pópoi ê phílon ándra
 alas say:IPF.3SG dear:ACC.M.SG man(M):ACC.SG
diōkómenon perí teíkhos / ophthalmōisin
 pursued:PART.PRES.MED.ACC.SG around wall(N):ACC.SG eye(M):DAT.PL
 horōmai
 see:PRES.IND.MED.1SG

‘Alas, he said, I see a man who is dear to me being pursued around the walls.’ (Homer, *Iliad* 22.168–169)

- c. Sed quis hic est,
 but who(INT):NOM.M.SG this:NOM.M.SG be:PRES.IND.3SG
 quem **astantem** video
 who(RP):ACC.M.SG standing:PART.PRES.ACC.M.SG see:PRES.IND.1SG
 ante ostium?
 before door(N):ACC.SG

‘But who is it this one that I see standing before the door?’ (Plautus, *Bacchides* 451)

Complement-taking predicates of immediate perception typically share an argument with their subordinate clause, as the MC object is the subordinate subject (e.g., *I see you running* implies that I see you and that you are running; this is also clear when the object is a nominalization, as in *I see your running*). Moreover, the subordinate situation has a determinate time reference that is simultaneous with the situation represented in the MC. As is typically the case in complementation, sharing of information determines syntactic deranking of the CC, that is, the arguments or the verb of the CC are formally different from their corresponding categories in an independent clause. In this case, the verb is implicit, or non-finite, lacking information for person, and the subject of the CC is in the oblique case

– in the ancient IE languages, typically in the accusative, as the predicate ‘see’ is regularly transitive.¹⁰

Formally, the participial clauses in (16) cannot be analysed as relative constructions, which in IE studies are limited to subordinate clauses marked by a relative pronoun. From a functional point of view, however, these clauses are similar. In fact, I consider the use of the participle depending on immediate perception predicates in CCs to be derived from the participle used as a strategy of relativization. We have seen in Section 2 that participles, similarly to adjectives, originally agree in gender, number, and case with their head noun, and may be functionally equivalent to RCs in ancient IE. This construction is maintained when the head noun is the object of a complement-taking predicate. In this case, the modifying participle and its head noun form a noun phrase, which is the direct object of an immediate perception predicate. Everything remains in the domain of the simple sentence. A sentence such as *I see a man running* in (17a) had originally the same constituent structure as a sentence such as *I see a blond man* in (17b). It may be hypothesized that the change occurs when the verbal features of the participle, such as tense/aspect, mood, or voice, prevail over its nominal features of gender, number, and case. As a verbal constituent, the participle is now to be connected with a new subject, that is, with the object of the immediate perception predicate, which is its syntactically closest noun phrase. Together, this noun and the participle are reinterpreted as a subordinate clause, depending on the immediate perception predicate, as in (17c). At this point, the sentence does not mean that I see a man who is running or a running man (English may use a different word order to express the attributive or predicative function of the participle). It means that I see the event of a man who is running. An original clause constituent, consisting of the participle and its head noun, is externalized and acquires clause status.

- (17) a. [S [NP I] [VP [V see] [NP a man running]]]
 b. [S [NP I] [VP [V see] [NP a blond man]]]
 c. [S1_MC [NP I] [V see]] [S2_SC [NP a man] [V running]]]

Later on, participles extend from their original relativizing domain to other complement relations, as well as to other subordinating relations outside com-

¹⁰The predicate ‘hear’, instead, displays more variation. It can govern an accusative but also another oblique case in IE. In Ancient Greek, for example, the verb *akoúō* ‘I hear’ properly selects the accusative of the entity heard and the genitive of the person, but one also finds the genitive of the entity, as well as a prepositional phrase for the person. The lower transitivity of ‘hear’ compared to ‘see’ may be explained by the fact that hearing implies a less direct source of information.

plementation. Their old use with complement-taking predicates of immediate perception is maintained in several IE languages, such as English, Lithuanian, and Russian, up to the present. In Germanic languages such as German and Icelandic, instead, we have the infinitive in this function. Cf. German *ich sehe einen Mann rennen*, Icelandic *ég sé mann hlaupa* ('I see a man run:1F'), with an infinitival complement clause, meaning 'I see a man running', vs. German *ich sehe einen rennenden Mann*, Icelandic *ég sé hlaupandi mann* (I see running man), with an attributive, relativizing participle, meaning 'I see a man who runs'. Some Romance languages also use the infinitive with complements of immediate perception predicates, cf. Italian *vedo un uomo correre*, French *je vois un homme courir*, as Latin used both the infinitive and the participle with these predicates (*video hominem currentem*, *video hominem currere*). The construction of the infinitive, however, represents an innovation with immediate perception predicates – the original use of the infinitive, as we have seen in Section 3, was circumstantial, to mark purposive relations, in IE.

4.3 Paratactic and correlative complement clauses

Complement-taking predicates of immediate perception imply an intermediate level of syntactic integration with respect to predicates of utterance, propositional attitude, and knowledge, on the one hand, and to modal, phasal and desiderative predicates, on the other (cf. Lohninger & Wurmbrand, this volume). The former group of predicates is originally characterized by loose types of clause linkage, such as parataxis or correlation. This is because they express situations that are scarcely conditioned by the situation represented in the MC. The two clauses do not necessarily share arguments (*I say that I / you / he go(es) to the city*). Moreover, the time reference of the subordinate is also not conditioned by the time reference of the MC (*I say that I go / went / will go to the city*). Utterance predicates selecting paratactic structures can be observed virtually everywhere in IE, but in some languages, as in ancient Indo-Iranian and in Hittite, they represent the most common strategy for reporting someone's words. That is, these languages favour direct speech over indirect speech. In Hittite, in particular, direct speech is marked by the quotative particle *-wa(r)-*, which commonly occupies the second position (after the conjunctions *-a / -ia* 'and' and *-ma-* 'but+', and before the enclitic pronouns, if any) in the clause-initial chain of clitics that is so typical of the Anatolian languages. The following text is drawn from the annals of the king Muršiliš II (14th-13th century BC), who informs a local chieftain about his intention to conduct a war. As can be seen, the particle

-*wa(r)*-, here written in bold, is repeated in each reported clause, to indicate that the direct speech continues.

- (18) ANA ^MU^hha-LŪ=ma ^{LŪ}TEMU uīyanun nu=ši
 to Uhhaziti-but messenger send:PRET1SG CON-him:DAT
 hatranun ARAD^{MEŠ}=YA=**wa**=ta=kan kuēš anda
 write:PRET1SG servant:PL-MY-QUOT-you:DAT-PTC who:NOM.C.PL PREV
 uer nu=**war**=ašta EGIR-pa kuit wewakkinun
 come:PRET3PL CON-QUOT-PTC back because demand:PRET1SG
 nu=**war**=aš=mu EGIR-pa UL pāišta
 CON-QUOT-THEM:ACC-me:DAT back NEG give:PRET.2SG
 nu=**wa**=mu=za DUMU-lan halzeššešta
 CON-QUOT-me:ACC-REFL child:ACC call:PRET.2SG
 nu=**wa**=mu=za tepnuškit kinun=a=**wa** ehu nu=**wa**
 CON-QUOT-me:ACC-REFL belittle:PRET.2SG now-and-QUOT EXCL CON-QUOT
 zahhiyawaštati nu=**wa**=naš ^DU BELÍ=YA DINAM
 fight:PRES.MED1PL CON-QUOT-us:ACC Stormgod lord-my case
 hannāu
 decide:IMP3SG
 ‘But to Uhhaziti I sent a messenger, I wrote him: ‘My servants who came to you, because I now demanded (them) back, you have not given them back to me. And you have been calling me a child, and you have been belittling me! And now we will fight each other. And the Stormgod, my lord, decide our case.’ (Kbo.III.4)

A syntactic change from parataxis to embedded hypotactic structures is traditionally identified in Germanic, where utterance predicates select a complementizer, such as English *that*, German *dass*, Icelandic *að*, which goes back to an originally demonstrative pronoun (cf. Hopper & Traugott 1993: 185ff). Similarly to what we have seen in Section 2 for relative pronouns derived from demonstrative pronouns in Germanic, we may have a reanalysis in complementation as well. Originally, two paratactic clauses connected by asyndeton are reconstructed, e.g. [*The man said that.*] [*We have cold weather.*]. The demonstrative pronoun, direct object of the utterance predicate, refers cataphorically to the content of the subsequent clause. Later, the two clauses may be reinterpreted as a unitary complex sentence, with a shift in the clause boundary, i.e., [*The man said [that we have cold weather]*]. Here we have a complementizer with no argument function in the MC. This traditional reconstruction has been recently challenged by Axel-Tober

(2017), among others, who considered Germanic complementizers to be derived from relative pronouns, rather than from demonstrative pronouns. Originally, we would have a correlative construction, e.g., *The man said that, that we have cold weather*, where the subordinate clause is functionally complete. Later, the resumptive element in the MC would have been omitted. The relationship with the demonstrative source of the subordinator would be indirect in this case.¹¹ A similar analysis has been offered for the Hungarian complementizer *hogy* ‘that’ (É. Kiss 2023). However, as correlative constructions are adjoined and not embedded, this hypothesis does not substantially contradict the change from a loose clause linkage to embedding, or from less subordinate to more subordinate structures – subordination, as any other syntactic category, is often a matter of degree, rather than a dichotomy (cf. Haiman & Thompson 1984).

4.4 Deranking strategies of complement clauses

Unlike parataxis, embedding typically implies the use of shifters from deictic into anaphoric elements in indirect speech (first- or second-person pronouns > third-person pronouns; *here* > *there*; *now* > *then*; *yesterday* > *the day before*, cf. Jakobson 1971), as we have seen in (15a)-(15b). Also, in this case, as the difference between coordination and subordination is not discrete, we may have various hybrid structures. Classical Sanskrit, for example, attests utterance predicates selecting a complementizer (*yad*, from the stem of the RP) without shifters. A sentence such as *The man said to the woman_i that you_i are a good friend* is perfectly

¹¹The use of CCs derived from correlative clauses is not limited to the case of subordinators (more or less directly) derived from the stem of demonstrative pronouns, as in Germanic. Ancient Indo-Iranian, for example, attests “explicative clauses” (*Explicativsätze*, cf. Delbrück 1888: 573ff; 1900: III, 324ff) that go back to correlative constructions and still have a relative pronoun (**yó-*) that synchronically differs from demonstrative pronouns (**tó-* etc.). In this case, a predicate of utterance, propositional attitude, or knowledge in the MC presents a noun phrase which is further elaborated in the subsequent subordinate clause. E.g., *We know this deed of Indra, that he killed the dragon* or *This deed of yours is renowned, o Indra, that you killed the dragon*. Like a relative pronoun, the subordinator ‘that’ agrees here in gender, number, and case with a nominal constituent of the MC. Unlike a relative pronoun, however, and like a complementizer, this subordinator has no syntactic function in the subordinate clause. It is enough to omit the head noun in the MC to have embedding with predicates of utterance, propositional attitude, and knowledge – this is an easy step, since the head noun usually denotes a discourse topic that is recoverable from the context. E.g. *We know this deed of Indra, that he killed the dragon* > *We know that he killed the dragon*, and *This deed of yours is renowned, o Indra, that you killed the dragon* > *it is renowned, o Indra, that you killed the dragon*. For the development of complement clauses in Indo-Iranian, cf. Viti 2007b: 211ff and Lühr (2008, 2014). The latter hypothesizes that the structural change from relative pronoun to complementizer in Old Indic subordination is facilitated by the presence of null objects in the sentence.

possible in Classical Sanskrit, albeit less common than the strategy with direct speech. The latter is marked by the particle *īti*, lit. ‘thus’, used once at the end of the reported speech, cf. example (13). In Viti (2007a), I explained the synchronically aberrant cases of complementation without shifters in Old Indic as the result of a syntactic calque from Dravidian, where complementizers (also beyond the domain of utterance complement-taking predicates) commonly derive from a verb of saying. The complementizers of Tamil *en-ru*, Malayalam *en-na*, Telugu *an-i* are originally perfect participles of the Proto-Dravidian verb **aHn-* ‘say’ (cf. Krishnamurti 2003: 451). This development from verbs of saying to complementizers, which is uncommon in IE, is common in many other language families and areas of Asia, Africa, and Oceania (cf. Sauerland et al. 2020, for example, on the form *wa* ‘tell; that’ in Teiwa, a Papuan language spoken on the Pantar Island in eastern Indonesia). This represents another instance of the change from direct speech to indirect speech and, more generally, from parataxis to embedding.

In embedding, various morpho-syntactic strategies may be used to represent the tighter linkage between MCs and subordinate clauses. Agreement between the tenses or moods of the two clauses is one of the most common strategies for this purpose. Accordingly, the tense or mood of the subordinate clause is conditioned by the tense or mood of the main clause. This is not limited to complementation, but seems to be much more typical of CC than of other types of subordinates, according to my data. In Ancient Greek, for example, a historical tense in the MC (i.e., an imperfect, an aorist, or a pluperfect) typically triggers an optative mood in the subordinate, in the so-called use of the “oblique optative”. In Latin, agreement in tenses and moods (*consecutio temporum*) becomes even more elaborated, according to the distinction into two aspects, imperfective (*infectum*) and perfective (*perfectum*), and three tenses (present, past, and future), which characterizes the whole Latin verbal system. If, for example, the MC has a verb in the indicative, the subordinate will have the present subjunctive to express simultaneity between the two clauses, a subjunctive perfect to express anteriority to the MC, and a periphrastic structure (with *sim*, *sis*, etc.) for posteriority, cf. Hofmann & Szantyr (1965: 550) for details. In Latin, the subjunctive is associated with subordination, and its productivity has been inherited by the Romance languages, which attest the subjunctive in complementation to a much larger extent than most Germanic languages.¹² Usually, when complementation admits variation between the indicative and the subjunctive in a language, predicates expressing a report are associated with the indicative. The subjunctive is

¹² An exception to this may be seen in Icelandic, where the subjunctive is quite productive e.g. in expressions of belief (cf. Sigurðsson 2010, Gärtner & Eyþórsson 2020).

preferred for the expression of more subjective thoughts, doubts, or scarce commitment. Cf. Italian *Dico che viene* ‘I say that he comes’, with an indicative verb in the subordinate, vs. *Penso / dubito / non credo che venga* ‘I think / I doubt / I don’t believe that he comes’, with the subjunctive. Italian preserves the subjunctive more than other Romance languages. Agreement in tense or mood between MCs and complement clauses represents an innovation in IE. Most IE languages have no such agreement, and each clause presents the tense and mood that is more appropriate for its situation independently of the surrounding clauses. We can observe this in ancient Indo-Iranian, for example. Even nowadays, many IE languages, such as Lithuanian, do not have this kind of agreement. Thus, no agreement in tense or mood can be reconstructed for PIE. This may be because the common paratactic or correlative loose clause linkage of the earliest IE languages is not compatible with any deranking strategy that tends to present the subordinate clause as different from a corresponding independent clause.¹³

Embedding is especially evident when the structure of the subordinate clause is deranked, in the form of both a non-finite verb and oblique arguments, with respect to an independent sentence. This can be seen in English, where complement-taking predicates of propositional attitude, for example, may be structured in infinitival patterns, e.g. *I believe him to be sincere*. This is also an innovation, independently occurring in different IE branches, which cannot be reconstructed for PIE. It appears already, however, in some ancient languages such as Latin, where this use, called *accusativus cum infinitivo*, is found with many more types of complement-taking predicates than in English. The following passage, from Old Latin comedy, contains two examples of this construction with a predicate of utterance (*dicere* ‘to say’) and a predicate of propositional attitude (*negāre* ‘to deny’) at short distance. In particular, the MC subject says or denies in the present that something happened in the past. The anteriority of the subordinate situations is rendered by the infinitive perfect in Latin, which extends its typical agreement in tense to the domain of non-finite complement clauses.

¹³In some languages, the higher or lower commitment on the information conveyed by the subordinate clause is expressed by different subordinating conjunctions, rather than by different moods. This is a rarer strategy for mood alternation, but is already attested in some ancient IE languages, such as Ancient Greek, again by independent innovation. In Ancient Greek, predicates of utterance, propositional attitude, or knowledge tend to select the subordinator *hóti* ‘that’ (from the stem of the relative pronoun) when the MC subject or the speaker is confident that the situation of the complement clause has occurred or will occur. In case of doubt, instead, the cognate subordinator *hōs* is usually privileged.

(19) AMPHITRUO:

Tun **me** heri **advenisse** dicis?
 you me:ACC yesterday come:IF.PRF say:PRES.IND2SG

ALCMENA:

Tun **te** **abiisse** hodie hinc **negas**?
 you:NOM.INT you:ACC go.away:IF.PRF today from.here deny:PRES.IND2SG

‘AMPHITRUO: Do you say that I came yesterday?’

ALCMENA: Do you deny that you went away from here today?’ (Plautus, *Amphitruo* 758)

The accusative with infinitive is used in several ancient and modern languages, such as Latin, Ancient Greek, English, etc., also with complement-taking predicates other than utterance, propositional attitude, and knowledge. It is even more common with desiderative, manipulative, phasal, and modal predicates. In example (20), again from Old Latin, the desiderative predicate *volō* ‘I want’ in the MC selects an infinitive complement. The MC subject and the subordinate subject differ, and the latter is coded in the accusative (*adfinitem* ‘alliance’). The Latin structure corresponds quite well to the English translation *I want this alliance to continue* etc. (but not to its Romance continuations, e.g. Italian *voglio che questa alleanza continui*, with a finite subordinate marked by a complementizer, see below).

- (20) **manere** **adfinitem** **hanc** inter nos
 continue:IF.PRES alliance(F):ACC.SG this:ACC.F.SG between us:ACC
 volo, si ullo modo est ut possit,
 want:PRES.IND1SG if any way:ABL be:PRES.IND3SG that may:PRES.SUBJ3SG
 quod spero fore
 which hope:PRES.IND1SG be:IF.FUT
 ‘I want this alliance between us to continue, if by any way it possibly
 may, which I hope will be the case.’ (Terence, *Hecyra* 723–724)

Manipulative predicates meaning ‘make’ or ‘let’ also usually select an infinitive complement, e.g. Latin *facio te canere* ‘I make you sing’, *sino te canere* ‘I let you sing’, where the primary argument of the subordinate situation, if explicit, is coded by an oblique case, mainly the accusative. This syntactic construction remains in Romance, e.g. Italian *ti faccio cantare*, Spanish *te hago cantar*, French *je te fais chanter* (apart from the gradual change from second-position clitics to verb-adjacent clitics). We can find plenty of parallels of infinitival causative predicates

in other IE branches.¹⁴ Similarly, an infinitive is usually required by modal predicates, expressing possibility or necessity, with either deontic or epistemic functions. Auxiliaries of possibilities are originally verbs meaning ‘be strong, have power’, such as Latin *possum* ‘I can’ from **potis sum* ‘I am the master’. Auxiliaries of necessity are expressions of lacking or taking away, such as Latin *debeō* ‘I must’ from **dē-habeō* ‘I hold away from someone, I withhold’. In this case as well, we have a substantial syntactic continuity from Latin *facere possum* ‘I can do’, *facere debeo* ‘I must do’ and the Romance reflexes, such as French *je peux faire*, *je dois faire* (apart from the change in word order). Phasal predicates, as anticipated, imply the inception, continuation or completion of a process. The infinitive is the most common strategy used for this, e.g. French *je commence à / continue à chanter* ‘I start / continue to sing’, *j’arrête de / termine de chanter* ‘I stop / finish singing’. Other IE languages, such as English, use the participle in this case, another non-finite and deranked construction.

Deranking was even more pronounced for these predicates at a previous stage of IE. Originally, in PIE, desiderative predicates having the same subject as the subordinate situation (*I want to do something*) could be expressed by means of morphological strategies, such as a series of suffixes. We can reconstruct a desiderative formation with the suffix *-se-* added to an accented reduplication in *i* and to a root in the zero grade. E.g., from the PIE root *tek^w* ‘run’, we have PIE **tí-tk^w-so-nti* ‘they want to run’ > Young Avestan *tixšənti* ‘they want to run’. Similarly, suffixation was originally used in PIE to express causative predicates (*I make you do something*). In this case, a root in the full grade *o* could be added with an accented suffix *-éye-*. E.g., from the PIE root **wert-* ‘turn’, we have PIE **wort-éye-ti* ‘it makes turn’ > Gothic *fra-wardeiþ* ‘it corrupts’.

Morphological strategies of prefixation, suffixation, or inflection could also serve to represent modal predicates and phasal predicates. Modal functions were originally coded by inflection in different grammatical moods. Besides the indicative, expressing realis, and the imperative, used for commands, PIE also had an optative and a subjunctive. The forms of these two moods still coexist in Ancient Indo-Iranian and in Ancient Greek, despite some functional overlaps.¹⁵ The op-

¹⁴This does not imply that infinitive complement clauses are the only expression of causative predicates. We can also have syntactically heavier constructions with subordinator and a finite verb, such as Latin *efficio ut*, *curo ut* + subjunctive. These constructions often have a connotation of less direct causation with respect to their more integrated infinitival correspondents. A similar situation holds in Romance. In Italian, for example, the sentence *faccio in modo che tu canti* (which can also be translated as ‘I let you sing’, or ‘I make sure that you sing’) iconically implies a less direct causation than the more deranked sentence *ti faccio cantare*.

¹⁵From the functional point of view, instead, the term “optative” has been less properly extended to some uses of the subjunctive, as in Latin (cf. van der Auwera & Schalley 2004 for such terminological problems).

tative (formed with the PIE suffix **-yéh₁/ih₁-* and with historical endings) especially expressed the speaker's wish, as well as the possibility that the denoted situation might take place, e.g., from the PIE root **h₁es-* 'be', we have PIE **h₁s-yéh₁-t* 'it may be' (Vedic *syát*). The subjunctive, a thematic formation instead, was mainly associated with expectation and with situations that shall or should happen in the future (this is the reason why several IE languages form their future tense, nonexistent in PIE, on the basis of the subjunctive), e.g. PIE **h₁és-h₁e-t* 'he shall be, let him be, he will be' (Vedic *ásat*). In the domain of phasal predicates, durative and completed events were originally coded by the imperfective and the perfective verbal stem, respectively. Alternatively, continuation could be expressed by various iterative formations, such as an iterative suffix **-éye-* associated with the *o* grade of the root. E.g., from the PIE root **g^{wh}ed^h-* 'beg' we have the PIE injunctive **g^{wh}od^h-éye-t* 'he continued begging' > Homeric Greek *póthei* 'he missed, longed for'. Inception could be coded by an inchoative suffix **-sk-*, e.g. Latin *florēscō* 'to begin to blossom'. Moreover, originally adverbial particles with a locative or directional meaning may also be prefixed and acquire various actional or aspectual values. We can see this especially in Slavic, e.g. Russian *bezhát* 'to run' vs. *pobezhát* 'to start running'. For reasons of space, I cannot describe in detail all these phasal formations, which develop in different ways in different IE languages, and which reveal phenomena of interaction between aspect and actionality.

The integrated constructions used for desiderative, causative, phasal, and modal predicates can be explained by the semantics of these predicates, which imply shared arguments and determinate time reference for the subordinate situation. Modal predicates and phasal predicates imply the same subject for the MC and the subordinate situation. In the sentences *He can sing*, *He must sing*, the person who can or must sing is also the person who sings. In the sentences *He starts / continues / stops / finishes singing*, the person who starts, continues, stops, or finishes singing is also the person who sings. The time reference of the subordinate situation, in this case, the action of singing, is simultaneous with the time reference of the MC. In complement clauses headed by causative predicates, the object of the MC is the subject of the subordinate clause. In the sentence *I let him sing*, the form *him* indicates the shared argument. Moreover, the time reference is the same, at least in the case of immediate causation, which is the unmarked situation for causative predicates. Desiderative predicates have a determinate time reference as well, in this case following the MC situation. Their arguments are also shared. In sentences such as *I want to sing*, the MC and the subordinate share the same subject. In sentences such as *I want him to sing*,

the subordinate subject is the MC object. The former case is more common in usage, as people prefer to express their own desires rather than someone else's.

Morphological strategies of complementation are quite productive in ancient Indo-Iranian. In Old Indic, for example, morphological desiderative and causative predicates are regular (e.g., *pípāsāmi* 'I want to drink', *dhārāyāmi* 'I make you hold'). In most other languages, instead, they decline and have just left residues that are no longer synchronically transparent. This is due, on the one hand, to the decay of verbal inflection, e.g. to the loss of some tenses or moods in the individual IE languages, and on the other hand to the increasing opacity of derivational affixes. Morphological formations of desiderative, causative, modal, and phasal predicates are replaced by periphrastic constructions, according to a common drift from synthesis to analysis. In this case as well, the competition between more and less integrated syntactic constructions is sensitive to semantic factors of shared information. If the MC and the subordinate clause have the same subject, we often have the infinitive, as we have seen. If they have different subjects, finite clauses marked by a subordinator prevail. In French, for example, constructions with desiderative predicates such as *je voudrais savoir* vs. *je voudrais que tu saches* present the same pattern as structures of propositional attitude such as *je pense d'avoir eu de la chance* vs. *je pense que tu as eu de la chance*. In case of a different subject, the accusative with infinitive tends to be lost.¹⁶

4.5 Increasingly specific complement-taking predicates

The category of complement-taking predicates widely differs in extension across different languages. While English allows embedding with a large variety of complement-taking predicates expressing utterance, propositional attitude, knowledge, pretence, fearing, etc., we have seen that in ancient IE languages such as Old Indic or Hittite, complementation is much more restricted. Remarkable differences may emerge even within the same class of complement-taking predicates. Utterance complement-taking predicates, for example, represent in English a large class including predicates that express specific manners of verbal communication, such as *shout* or *whisper*. This, however, is not possible in many other languages. Manfred Krifka (p.c.) observes that in Daakie (an Oceanic language spoken in Ambrym Island, Vanuatu) one cannot say a sentence such as "He shouted that there is a whale out there". One would rather say, "He shouted, he said that there is a whale out there". A study of the diachronic development

¹⁶Note that the loss of the accusative with infinitive in Romance is not due to the downfall of the ancient case system. Cases remain in Romance pronouns (cf. French *je* 'I' vs. *moi, me* 'me'), and could be in principle employed with the infinitive similarly to what happens in English.

of complement-taking predicates in languages would require a precise corpus analysis, which to my knowledge is still a desideratum, and which is beyond the scope of this paper. It may be argued that semantically unspecific utterance predicates, such as *say*, which do not express the manner of verbal transfer, are associated first with embedding, and that other predicates may follow, according to whether subordination is more or less productive in the language at issue. Similarly, among desiderative predicates, generic predicates such as *want* better establish morphological devices of verbal suffixation with respect to predicates that specify the manner of desire, such as *hope*, *wish*, *crave*, etc. A grammaticalization process starts with forms characterized by a generic semantics.

5 Conclusion

In this paper, I have discussed some common developments of embedded SCs in the history of IE. We have seen that embedding, which is so typical of subordination in English and most modern languages of Europe, is often secondary from a diachronic point of view. On the one hand, embedded subordinates tend to develop from originally paratactic constructions (usually by reanalysis of demonstrative pronouns or adverbs into subordinators), from originally adjoined correlative structures (by deletion of the anaphoric resumptive in the MC), or more generally, from looser to tighter clause linkage. On the other hand, implicit embedded subordinating structures, represented by verbal adjectives or participles and by verbal nouns or infinitives, tend to be replaced by explicit embedding with a finite verb and a subordinator. We have also seen that the attainment of explicit strategies of subordination often follows a path of relative cognitive complexity in languages, as semantically more complex subordinating relations, such as concession, receive a specific subordinator later than simpler relations, such as those expressed by temporal clauses or conditional clauses. Of course, components of semantic complexity must be independently identified and motivated by specific cognitive factors – there is a rich linguistic literature on that, based on both synchronic and diachronic evidence, as well as on first and second language acquisition and pidgin studies, cf. Dahl (2004); Miestamo et al. (2008), etc. A subordinate clause that implies a semantic component of negation, for example, is more complex than a subordinate clause without negation. Moreover, within embedding, we may have more or fewer deranked constructions, according to whether the MC and the SC share pieces of information, such as arguments or time reference.

We can also observe that RCs are diachronically primary with respect to ACs and to CCs, that is, RCs have been expressed by specialized constructions earlier

than ACs and CCs, as far as we can reconstruct by comparing the subordinating system of the different IE languages. ACs and CCs are often marked by subordinators based on the stem of the relative pronoun, for example. While RCs longer retain the ancient adjoined pattern of the correlative diptych, embedding turns out to develop for ACs first. Most ACs do not present any anaphoric resumptive element in the matrix clause. This may be explained by the fact that ACs refer to an event or state, and are therefore intrinsically less referential than relative constructions, which refer instead to an object or person and hence more easily trigger the use of an anaphoric pronoun. We have seen that the old correlative pattern is retained only in a few ACs, such as clauses expressing space, time, and manner (which can be functionally equivalent to RCs), as well as consecutive clauses. The latter presents a cause-and-effect sequence of events as a balanced structure. CCs turn out to be the last acquisition in the subordinating system of the IE languages. They often originate from paratactic or adjoined constructions, which may leave traces in the morphology of the complementizer or the clause structure. Consider the following construction, drawn from the Old Persian story of Darius killing the usurper Smerdis, where a complement-taking predicate of fearing is rendered by parataxis: *kāra-šim hacā daršam atarsa kāram vasiy avā-janiyā hya paranam Bardiyam adānā* ‘The people feared greatly that he could slay at will the people who previously had known Smerdis’, lit. ‘The people feared him greatly. (They thought that) he would slay’, etc. (Darius inscription at Behistun, 1 column, 50–51). Within complementation, we have seen that the earliest embedded pattern is represented by the participle with predicates of immediate perception. It may be interesting to investigate to what extent these diachronic tendencies, based on the IE languages, may be applied to other language families.

All this does not imply that a change from implicit to explicit subordination is necessary. In Mandarin Chinese, for example, subordinating relations are often left implicit and must be inferred from the context. Participles have always been the main strategy to express RCs in Dravidian, and nominalizations have always been the main pattern of RCs in Turkic. In the same way, we do not imply that paratactic patterns necessarily develop into subordination and embedding. In Arabic, subordination has always been less common than coordination in clause linkage. One of the first things one learns in doing translations from Arabic into English is that the Arabic coordinator *wa* is much more frequent than the English coordinator *and*, and can often be rendered in English by means of subordinating conjunctions as *when*, *even though*, etc. We rather imply that, *if* a change occurs in the subordinating system of a language, implicit strategies of subordination are much more commonly replaced by explicit strategies than the opposite, and that the path from parataxis to embedding, from correlation to embedding, and

more generally from a looser to a tighter clause linkage is much more common than the opposite.

Some potential criticisms must be briefly addressed. Hendery (2012), for example, talks of an alleged “ideological investment” to an “earlier ‘received view’ of adjunction and parataxis as primitive and embedding and hypotaxis as sophisticated, modern devices” (2012: 165). She contests the idea of a general change from juxtaposition to syntactic bondedness – an idea which has originated in the past (e.g. by the Neogrammarians) and is often still supported nowadays (cf. Hopper & Traugott 1993, Givón 2001, etc.). Now, we do not have to confound ideology with philology. The Neogrammarians may be corrected in some uses of the terminology – their statements must be understood in the framework of the positivistic culture of their time, and currently, scholars are no longer opposing primitive to mature languages in IE studies or in historical linguistics. The Neogrammarians, however, had a profound mastery of the ancient IE languages – especially Delbrück (1900) identified syntactic phenomena and mechanisms that have been proven to be valid also in other language families and to stand to the scrutiny of modern linguistic theory (cf. Cotticelli-Kurras & de Decker 2023). The data of the IE languages substantially *confirm* the idea that – in case a syntactic change occurs – clause linkage usually proceeds from a looser to a tighter connection, where syntactic bondedness can be measured on the basis of precise syntactic criteria, such as the use of more or less deranked forms for the predicate and for the arguments (cf. Lehmann 1988). Whoever has some familiarity with Vedic or Hittite finds this absolutely clear. But also in other IE languages – even languages with dominant embedding, such as Latin and Greek – traces of ancient paratactic or adjoined structures appear, especially if we consider the earliest texts.

The fact that adverbial subordinators and complementizers may originate from relative pronouns does not impinge upon this diachronic tendency. Firstly, this makes the process just less direct, as relative pronouns usually originate from demonstrative or interrogative-indefinite pronouns, which had a syntactic function in independent, paratactic, or adjoined sentences. Secondly, the morphology of a subordinator is one thing, and the clause structure is another. Again, correlative structures are *not* embedded – this is worth stressing since correlation is often wrongly presented as a piece of evidence against the idea of a relatively late established hypotactic system. And again, the difference between parataxis and hypotaxis must be seen as a matter of degree. The gradual development of embedding can be explained by syntactic theories of grammaticalization, as we have seen, as well as by reanalysis and extension (cf. Harris & Campbell 1995 on the latter two mechanisms).

Neither is this diachronic tendency contravened by the change from one kind of embedding to another, rather than directly from paratactic constructions. As not all relations, in any linguistic domain, are grammatically coded at the same time, it is also expected that some subordinating relations are associated with embedding earlier than others. We have seen that some embedded constructions, especially in the domain of complementation, emerge relatively late as compared to relative clauses and to adverbial clauses. While the change from relative clauses to complement clauses prevails in the history of IE, cases of complement clauses derived from the syntactic pattern of adverbial clauses are also quite common. In the Romance languages, for example, complementizers such as Italian *come*, Spanish *como*, French *comment*, Romanian *cum* ‘how’, all derived from Latin *quō-modo* ‘in what manner’) may introduce complement clauses of indirect questions (e.g. French *je voudrais savoir comment tu as pu faire cela*), and clearly derive from manner clauses, which belong to adverbial subordination. Beyond IE, a similar change is widely attested in Semitic, where the conjunction *k-* expressing manner and time (cf. Ugaritic *k-*, *ky*, Hebrew *kī*, Phoenician *k-*, Aramaic *ky*, Arabic *ka-*, *kay-*, etc.) passed to adverbial relations of cause (Old Akkadian and Assyro-Babylonian *kī*, *kīma*, Phoenician *k-šm* ‘*ql*’), etc.) and later to complement relations, especially after verbs of command and perception (cf. Hebrew *kī*, Middle Babylonian *kī*, Ge’ez *kama*, Tigre *kə*, etc.), cf. Brockelmann (1913: 606; 6013ff); von Soden (1995: 275ff); Lipiński (2000: 473-474; 538-539; 546), etc. Deutscher (2000) has studied an instance of this change in detail, that is, the use of the Akkadian comparative and temporal conjunction *kīma* ‘as, when’ first as a causal subordinator meaning ‘because’ and later as a complementizer meaning ‘that’. In his view, this change may seem “unusual” and “very problematic” (2000: 37). In fact, this change is not problematic at all. There is no reason to believe that finite sentential complements develop *straightforwardly* from paratactic constructions, and no historical linguist, to my knowledge, has ever claimed the regularity of such a direct step. Again, we speak of a tendency from a looser to a tighter type of clause linkage, to be measured according to precise criteria of sharing of argument, time reference, control, etc. All these changes considered so far do *not* represent the opposite change from tighter to looser connection, or from subordination to parataxis. Incidentally, consider that Akkadian *kīma*, as all the above-mentioned subordinators, ultimately originates from a deictic and asseverative particle *k-* (Lipiński 2000: 473), that is, not from a subordinating conjunction.

The claim that the syntactic change from loose clause linkage to tight clause linkage may be as common as the opposite one, or that embedding may be equally common in ancient as in modern languages, relies on a wrong interpretation of

uniformitarianism. Nobody is contesting the fact that what is observed in the present can also be proposed in the past. This, however, does not imply that the textual expressions of ancient languages, associated with a fundamental oral mode of communication, have to follow the same pragmatic principles as the texts of modern languages, spoken in societies where the written mode is dominant, especially in urban cultures. In a functional view of language change, as adopted here, discourse pragmatics may also affect syntax, as different contexts and different formal or informal registers may be related to different constructions. In historical linguistics, by definition, syntactic constructions must be contextualized in their history, society, and culture. On the one hand, although the ancient languages are documented by written texts, these texts were usually composed for an oral performance, and often were recorded much later than their original oral form. On the other hand, it is well-known that the use of subordination is much more common in the written language; Givón (2001), for example, has provided important evidence on this. To argue that no particular directionality may be identified in the diachrony of embedding and subordination in general assumes a comparison among syntactic systems that are completely extrapolated from their socio-cultural context.

Genuine exceptions to the change from loose to tight linkage occur, of course. It is well-known that clauses introduced by the conditional coordinator ‘if’, for example, can be used as independent clauses, as in English *If you’ll just come next door* or in French *Oh, si j’étais riche!* (cf. Patard 2014). These are instances of what Evans & Watanabe (2016) calls “insubordination”, that is, the formation of independent clauses from subordinating structures, which seems to be a well-attested phenomenon cross-linguistically (cf. also Evans & Watanabe 2016, Beijering et al. 2019). Moreover, we may also have constructions that diachronically acquire more paratactic functional features. Converbs, for example, traditionally called “gerunds” in IE studies, are typically subordinating and adverbial in Old Indic (especially with the function of an *after*-clause), but are mainly coordinating in New Indic. These cases are rarer, however, as compared to the change from parataxis or correlation to embedding. The very fact that insubordination has been only recently identified is revealing. These exceptional cases deserve to be explained, but they do not impinge upon the prevalent directionality towards embedding. For Indic converbs, contact factors may have played a role, as coordinating converbs are common in Dravidian, as well as in numerous other Asian languages.

Acknowledgments

I would like to thank the editors Hans-Martin Gärtner and Manfred Krifka, as well as an anonymous reviewer, for their kindness and precious feedbacks, which helped me to improve the first draft of this paper.

Abbreviations

AC	adverbial clause	MED	middle voice
C	(genus) commune	OPT	optative
CON	connective	PIE	Proto-Indo-European
CONJ	conjunction	PRET	preterit
EXCL	exclamative	PREV	preverb
GV	gerundive	PTC	particle
IE	Indo-European	RC	relative clause
INDEF	indefinite	RCON	relative construction
INT	interrogative	RP	relative pronoun
IPF	imperfect	SC	subordinate clause
MC	matrix clause		

References

- Ambrasas, Vytautas. 1997. *Lithuanian grammar*. Vilnius: Baltos Lankos.
- Andrews, Avery D. 2007. Relative clauses. In Timothy Shopen (ed.), *Language typology and syntactic description: Complex constructions*, 2nd edn., vol. 2, 206–236. Cambridge: Cambridge University Press.
- Axel-Tober, Katrin. 2017. The development of the declarative complementizer in German. *Language* 93. 29–65.
- Axel-Tober, Katrin. 2023. Adverbial resumption in German from a synchronic and diachronic perspective. In Karen De Clercq, Liliane Haegeman, Terje Lohndal & Christine Meklenborg (eds.), *Adverbial resumption in verb second languages*, 167–194. Oxford: Oxford University Press.
- Bacskai-Atkari, Julia. 2020. English relative clauses in a cross-Germanic perspective. *Nordlyd* 44. 93–115.
- Beijering, Karin, Gunther Kaltenböck & María Sol Sansiñena (eds.). 2019. *Insubordination: Theoretical and empirical issues*. Berlin: de Gruyter.
- Bhatt, Rajesh. 2003. Locality in correlatives. *Natural Language and Linguistic Theory* 21. 485–541.

- Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: de Gruyter.
- Brockelmann, Carl. 1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen, Band II, Syntax*. Berlin: Reuther & Reichard.
- Chafe, William. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge: MIT Press.
- Cotticelli-Kurras, Paola & Filip Johannes de Decker (eds.). 2023. *Berthold Delbrück, historical and comparative Indo-European syntax 1922–2022*. Wiesbaden: Reichert.
- Coulmas, Florian (ed.). 1986. *Direct and indirect speech*. Berlin: de Gruyter.
- Cristofaro, Sonia. 2005. *Subordination*. Oxford: Oxford University Press.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.
- Dayal, Veneeta. 1996. *Locality in wh-quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer.
- de Vries, Mark. 2002. *The syntax of relativization*. Universiteit van Amsterdam. (Doctoral dissertation).
- Delbrück, Berthold. 1888. *Altindische Syntax*. Halle: Buchhandlung des Waisenhauses.
- Delbrück, Berthold. 1900. *Vergleichende Syntax der indogermanischen Sprachen III*. Strassburg: Trübner.
- Deutscher, Guy. 2000. *Syntactic change in Akkadian: The evolution of sentential complementation*. Oxford: Oxford University Press.
- Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology*, 1–48. Oxford: Oxford University Press.
- Dryer, Matthew S. 2013. Order of adjective and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. (May 2019). Leipzig: Max-Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/87>.
- Evans, Nicholas & Honoré Watanabe (eds.). 2016. *Insubordination*. Amsterdam: John Benjamins.
- Friedrich, Johannes. 1959. *Die hethitischen Gesetze*. Leiden: Brill.
- Friedrich, Johannes. 1960. *Hethitisches Elementarbuch*. Heidelberg: Winter.
- Gärtner, Hans-Martin. 2001. Are there V2 relative clauses in German? *Journal of Comparative Germanic Linguistics* 3. 97–141.

- Gärtner, Hans-Martin & Þórhallur Eypórsson. 2020. Varieties of dependent V2 and verbal mood: A view from Icelandic. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 208–239. Oxford: Oxford University Press.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy. 2001. *Syntax: An introduction*, vol. 2. Amsterdam: John Benjamins.
- Haiman, John & Sandra A. Thompson. 1984. 'Subordination' in universal grammar. In Claudia Brugman & Monica Macaulay (eds.), *Proceedings of BLS 10*, 510–523. Berkeley: Berkeley Linguistics Society.
- Harbert, Wayne. 2007. *The Germanic languages*. Cambridge: Cambridge University Press.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haudry, Jean. 1973. Parataxe, hypotaxe et corrélation dans la phrase latine. *Bulletin de la Société de Linguistique de Paris* 68. 147–186.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2003. Contact-induced grammaticalization. *Studies in Language* 27. 529–572.
- Helgander, John. 1971. *The relative clause in English and other Germanic languages: A historical and analytical survey*. University of Gothenburg. (Doctoral dissertation).
- Hendery, Rachel. 2012. *Relative clauses in time and space*. Amsterdam: John Benjamins.
- Hoffner, Harry Angier. 1997. *The laws of the Hittites: A critical edition*. Leiden: Brill.
- Hofmann, Johann B. & Anton Szantyr. 1965. *Lateinische Syntax und Stilistik*. München: Beck.
- Hopper, Paul J. & Elisabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Jakobson, Roman. 1971. Shifters, verbal categories, and the Russian verb. In *Selected writings II*, 130–147. The Hague: Mouton.
- Keenan, Edward L. 1985. Relative clauses. In Timothy Shopen (ed.), *Language typology and syntactic description: Complex constructions*, vol. 2, 141–170. Cambridge: Cambridge University Press.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.
- É. Kiss, Katalin. 2023. From relative proadverb to declarative complementizer: The evolution of the Hungarian *hogy* 'that'. *The Linguistic Review* 40. 107–130.

- König, Ekkehard. 2006. Concessive clauses. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, 679–681. Amsterdam: Elsevier.
- Krifka, Manfred. 2010. How to interpret ‘expletive’ negation under *bevor* in German. In Thomas Hanneforth & Gisbert Fanselow (eds.), *Language and logos: Studies in theoretical and computational linguistics*, 214–236. Berlin: de Gruyter.
- Krishnamurti, Bhadriraju. 2003. *The Dravidian languages*. Cambridge: Cambridge University Press.
- Lehmann, Christian. 1984. *Der Relativsatz. Typologie seiner Strukturen. Theorie seiner Funktionen. Kompendium seiner Grammatik*. Tübingen: Narr.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–226. Amsterdam: John Benjamins.
- Lipiński, Edward. 2000. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.
- Lipták, Anikó (ed.). 2009a. *Correlatives cross-linguistically*. Amsterdam: John Benjamins.
- Lipták, Anikó. 2009b. The landscape of correlatives: An empirical and analytical survey. In Anikó Lipták (ed.), *Correlatives cross-linguistically*, 1–46. Amsterdam: John Benjamins.
- Lühr, Rosemarie. 2008. Competitive Indo-European syntax. In Gisella Ferraresi & Maria Goldbach (eds.), *Principles of syntactic reconstruction*, 121–159. Amsterdam: John Benjamins.
- Lühr, Rosemarie. 2014. Komplementsätze im Indoiranischen. *Historische Sprachforschung* 125. 227–241.
- Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson (eds.). 2008. *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins.
- Minard, Armand. 1936. *La subordination dans la prose védique: Études sur le Śatapatha-Brāhmaṇa*. Paris: Les Belles Lettres.
- Neu, Erich. 1974. *Der Anitta-Text*. Wiesbaden: Otto Harrassowitz.
- Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description: Complex constructions*, vol. 2, 42–140. Cambridge: Cambridge University Press.
- Patard, Adeline. 2014. Réflexions sur l’origine de l’insubordination: Le cas de trois insubordonnées hypothétiques du français. *Langages* 196. 109–130.
- Sauerland, Uli, Bart Hollebrandse & František Kratochvíl. 2020. When hypotaxis looks like parataxis: Embedding and complementizer agreement in Teiwa. *Glossa* 5(1). 89. DOI: 10.5334/gjgl.931.
- Schwyzer, Eduard. 1950. *Griechische Grammatik 2. Band, Syntax und syntaktische Stilistik*. München: Beck.

- Sigurðsson, Halldór. 2010. Mood in Icelandic. In Björn Rothstein & Rolf Thieroff (eds.), *Mood in the languages of Europe*, 33–55. Amsterdam: John Benjamin.
- Smits, Reinier J.C. 1989. *Eurogrammar. The relative and cleft constructions of the Germanic and Romance languages*. Dordrecht: Foris.
- Srivastav, Veneeta. 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9. 637–686.
- Thompson, Sandra A. 1985. Grammar and written discourse: Initial vs. final purpose clauses in English. *Text* 5. 55–84.
- van der Auwera, Johan & Ewa Schalley. 2004. From optative and subjunctive to irrealis. In Frank Brisard, Michael Meeuwis & Bart Vandenabeele (eds.), *Seduction, community, speech: A festschrift for Herman Parret*, 87–96. Amsterdam: John Benjamins.
- Vidyabhavan, Ramchandra Jha (ed.) 1991. *Pañcatantra*. Varanasi: Chowkhamba.
- Viitso, Tiit-Rein. 1998. Fennic. In Daniel Abondolo (ed.), *The Uralic languages*, 96–114. London: Routledge.
- Viti, Carlotta. 2007a. Ibridismo nell'espressione del discorso riportato in area indiana. *Studi e Saggi Linguistici* 45. 117–139.
- Viti, Carlotta. 2007b. *Strategies of subordination in Vedic*. Milano: Angeli.
- Viti, Carlotta. 2015. *Variation und Wandel in der Syntax der alten indogermanischen Sprachen*. Tübingen: Narr.
- von Soden, Wolfram. 1995. *Grundriss der akkadischen Grammatik*. 3rd. Roma: Editrice Pontificio Istituto Biblico.
- Zifonun, Gisela. 2001. *Grammatik des Deutschen im europäischen Vergleich. Der Relativsatz*. Mannheim: Amades.