

## CLIPP

Christiani Lehmanni inedita, publicanda, publicata

titulus

9. Latin

huius textus situs retis mundialis

[christianlehmann.eu/publ/lehmann\\_latin.pdf](http://christianlehmann.eu/publ/lehmann_latin.pdf)

dies manuscripti postremum modificati

30.09.2022

occasio orationis habitae

-

volumen publicationem continens

Heinz, Matthias & Hinzelin, Marc (eds.), *Manual of Classification and Typology of the Romance languages*. Berlin: W. de Gruyter (Manuals of Romance linguistics/Manuels de linguistique romane)

annus publicationis

ignotus

paginae

ignotae

# 9 Latin

**Christian Lehmann**

## 1 Introduction

Linguistic typology does not classify languages but seeks to discover types of linguistic structure. While in principle any structural trait may be used to characterize a language, those features are most fruitful for typology which correlate with other features so that they may be understood as jointly following an underlying principle of linguistic structure. From its beginning, linguistic typology has focused on grammatical structure; and this has been the most important research field to this day. Besides, some work has been done on phonological and lexical typology. There have also been proposals for classifying languages by criteria outside the linguistic system, like sociolinguistic typologies or typologies of writing systems. These will not be considered here, the main reason being that they are classifications rather than typologies.

Doing the typology of one language requires comparing it with languages of a different structure in order to bring out its specificity. Here Latin will be contrasted with some other, chiefly Non-Indo-European, languages with respect to a set of properties which have proved fruitful in the history of typology. All quantitative statements are based on the data provided in Dryer & Haspelmath (eds.) 2013. henceforth referred to as WALS.

## 2 State of the art and current research questions

### 2.1 Phonological typology

Both typological work on Latin and typological work on phonology are scarce, and their intersection is well-nigh empty. Ballester 1998 gives a concise survey of some relevant facts. Lehmann 2005 focuses on phonotactic aspects.

The Latin phoneme system comprises a consonant system of fifteen phonemes, comprising voiceless stops: /p t k/, voiced stops: /b d g/, fricatives: /f s h/, nasals: /m n/, liquids: /r l/ and semivowels: /w j/, although /h/ does not survive the period of Old Latin in the spoken language. The typological extremes here are Rotokas (West Bougainville, Papua New Guinea) with six consonants and !Xóõ (Southern Khoisan, Botswana) with 122 consonants. The world-wide average ranges at 22 consonants (WALS ch. 1). Thus, the size of the Latin consonant system may be rated as moderately small. It makes balanced use of the available phonetic dimensions.

Latin has a five-vowel system: /a e i o u/, the most wide-spread system, which it shares with one third of the world's languages. The typological extremes, counting only short vowels without any secondary articulation, are two vowels differing in height, e.g. in Yimas (Lower Sepik-Ramu; Papua New Guinea), and 14 vowels in German (WALS ch. 2). Here again, the Latin system distributes the vowels equally over the available phonetic space. This typologically dominant system is, however, complicated by being

matched with five long vowels. The functional load of these phonemic contrasts is high; minimal pairs like *latus* ‘side’ vs. *lātus* ‘wide’, including even grammatical functions as in *uenit* ‘comes’ vs. *uēnit* ‘came’, are not rare. Nonetheless, long vowels survive in none of the daughter languages. Diphthongs, instead, come and go over the entire Latin-Romance history.

The syllable structure (Lehmann 2005) is moderately complex. With onsets as in *scribo* ‘I write’ and codas as in *siremps* ‘likewise’, it is slightly above typological average (WALS ch. 12) and gets simplified on the passage to Romance. While there are syllables without an onset at the lexical level, liaison provides these with an onset at the phonetic level.

Latin has a stress accent that is phonologically determined. The rule is based on the mora. A syllable ending in a short vowel counts one mora; all other syllables count two morae. Word stress is determined thus: The last syllable counts one mora; stress falls on the syllable containing the third-last mora. Thus, we have: /'indikem/ (acc. sg. of *index* ‘indicator’), but /in'di:kem/ (1. sing. pres. subj. act. of *indīco* ‘indicate’). Latin word prosody is, in this, simpler than its ancestor language, which was a tone language, but also simpler than the accent system of several Romance languages, which is not phonologically conditioned.

There are not too many phonological processes to complicate the mapping of lexical phonological structure onto phonetic structure. All in all, Latin phonology avoids the typological extremes on all relevant parameters. In comparison, it is simpler and more clearly structured than the phonology of many other languages.

## 2.2 Grammatical typology

### 2.2.1 Classical linguistic typology

Disregarding some forerunners, linguistic typology started essentially at the beginning of the 19<sup>th</sup> century with work by Friedrich Schlegel (1808), August Wilhelm Schlegel (1818) and Wilhelm von Humboldt (1836). The period of this and subsequent work in the same vein may be called classical typology and contrasted with linguistic typology as it has evolved since the middle of the 20<sup>th</sup> century, which may be called modern typology. The most important difference between these two research strands lies in the fact that adherents of classical typology considered it possible to characterize a language as a whole in terms of its language type. They counted with a small number of **holistic types** that every language could be assigned to. This was considered possible since a language formed a coherent system so that its properties could not vary independently of each other, as was put in an aperçu by Antoine Meillet (1934: 474f) saying “que chaque langue forme un système où tout se tient”. The last important typologists who entertained variants of such a conception are Vladimir Skalička (1966), Petr Sgall (1971) and Eugenio Coseriu (1968, 1980, 1998).

In modern typology, this proposition is no longer taken for granted. Instead, a modular view of the linguistic system prevails. As a consequence, we not only have phonological, morphological, syntactic and lexical typologies, but even inside each of these fields partial typologies, such as alignment typology or head vs. dependent marking typology as aspects of syntactic typology. The notion of a holistic type appears to survive only in the concept of **Standard Average European** (launched by B. L. Whorf;

Haspelmath 2001), used – often with disparaging implications – to characterize a type of language which comprises Latin and has provided an implicit or even explicit *tertium comparationis* in much linguistic work.

Given the linguistic education enjoyed by scholars active in classical typology, the classical languages Latin and Greek played a fundamental role in shaping their linguistic thinking, including their ideas about typology. As a consequence, all early typologies include Latin as a representative of one of the types, viz. the flexional type.<sup>1</sup> To the extent that both knowledge of the classical languages has receded and many more languages have become better known in contemporary linguistics, the role of Latin in linguistic typology has been going down. It has, e.g. not been incorporated in WALS, although there is doubtless richer linguistic knowledge about Latin than of most other languages. It may be generalized at this point that Latin has been playing a marginal role in the linguistic typology of the last half century, Christol 1998 and an article series by Lehmann (s. References) being among the exceptions.

Classical typology is often dubbed ‘morphological typology’. This is not accurate in view of the fact that the most eminent representative of 19<sup>th</sup> century linguistics, W. v. Humboldt (1836: 653f), stated clearly that the types were ones of ‘clause formation’. These are the **isolating**, **flexional**, **agglutinative** and **incorporating type**. The flexional type is opposed to the agglutinative type as follows (o.c. 490f; Lehmann 1979): While the agglutinative type – standardly represented by Turkish, but in fact widespread in the world – provides a separate expression unit for the lexeme and for each of the grammatical features born by it, the flexional type – standardly represented by an ancient Indo-European language, but also found in other linguistic families – merges all of these semantic components in one expression unit. More precisely, a flexional language signals grammatical categories on the stem by modifying it (this being the original intuition behind the term *inflection*). Thus, Turkish (Turkic, Anatolia, E1) has a morpheme for each of the notion ‘year’, the genitive and the plural. Each grammatical feature is coded when needed and otherwise absent. In the position of the subject or predicate nominal, an agglutinative language just uses the bare stem without any specification of number or case; thus Turkish ‘year’ is simply *yıl*.

E1	a.	<i>yıl-in</i>	(year-GEN)	‘of the year’
TURKISH	b.	<i>yıl-lar-in</i>	(year-PL-GEN)	‘of the years’
E2	a.	<i>ann-i</i>	(year(M)-GEN.SG)	‘of the year’
LATIN	b.	<i>ann-orum</i>	(year(M)-GEN.PL)	‘of the years’
	c.	<i>ann-us</i>	(year(M)-NOM.SG)	‘year’

By contrast, Latin (E2) has an ending for the genitive singular, another ending for the genitive plural and no word form corresponding to the bare lexeme, since neither the root (*ann-*) nor the stem (*anno-*) can be used as a word form in the construction of a sentence. Instead, another inflected form (E2c) appears in the syntactic functions mentioned. In 20<sup>th</sup> century morphology, this insight has been reframed as the distinction between an ‘item-and-arrangement’ and a ‘word-and-paradigm’ morphology (Matthews 1965): The Turkish word form can be profitably described as a sequential arrangement

<sup>1</sup> The original term *flektierender Typ* involved a concept of inflection which has to be distinguished from the 20<sup>th</sup> century concept and was therefore translated as *flexional* instead of *inflecting type* by scholars like Otto Jespersen.

of morphs each of which codes a value of a parameter. The Latin word form is more efficiently described as a (possibly structurally unanalyzed) member of a paradigm, as has, in fact, been overwhelmingly the case in the teaching of Latin over the past two millennia. A methodological consequence of this insight is that languages of different types may require different models of grammar for an appropriate description.

Needless to say, 19<sup>th</sup> century typologists were keenly aware of the fact that there are no clear-cut borderlines between the types and that most languages are not pure representatives of a type. Latin was not, in fact, the prototype of a flexional language; this role was born by Sanskrit and, in somewhat lesser perfection, by Ancient Greek. Features of Latin that approximate it to the agglutinative type include the existence of a large set of nouns whose stem serves as the nominative (as in E27 below) and of some verbal categories like the imperfect which consist in an invariable morpheme added to the stem: add the morpheme *-ba-* (IMPF) to the stem *lauda-* ‘praise’, and you have the imperfect stem, which you then conjugate, e.g. *lauda-ba-s* (praise-IMPF.IND-2.SG) ‘you were praising’. Even in this case, though, one morpheme codes both tense and mood.

At a time where languages were evaluated against the standard set by ancient Indo-European languages, August Wilhelm Schlegel (1808) got aware of the fact that most modern Indo-European languages, including the modern Germanic and Romance languages, fall somewhat short of the flexional ideal. E3 translates E2 into French.

E3 a. *de l'année* (of<sub>DEF</sub> year) ‘of the year’  
 FRENCH b. *des années* (of:DEF.PL year:PL) ‘of the years’

In E3, the genitive is coded by a separate word. In E3b, this bears, in addition, the expression of number. Thus, the lexeme is not inflected; what is inflected is that additional word. The same morphological contrast is observed in conjugation.

E4 a. *cantavit* (sing:PRF:3.SG) ‘has sung’  
 LATIN b. *cantatur* (sing(PRS):PASS.3.SG) ‘is sung’

E5 a. *ha cantato* (have:3.SG sing:PTP.PRF) ‘has sung’  
 ITALIAN b. *è cantato* (be:3.SG sing:PTP.PRF) ‘is sung’

While in the Latin E4, all the conjugation categories are coded in one word form, in their Italian counterparts in E5, tense, person and number are coded by auxiliaries in periphrastic verb forms. Schlegel proposes the following conception: The modern Indo-European languages belong to the flexional type just as the ancient ones. However, the ancient languages show **synthetic morphology**, while the modern ones show **analytic morphology** (Ledgeway 2012, ch. 2).

### 2.2.2 Syntactic autonomy

These insights were carried further in different directions. The idea that the flexional type involves not merely a technique of constituting word forms, but also one of constituting syntactic constructions is pursued in Meillet 1948:145ff. No special segmental or suprasegmental means are necessary to compose a clause out of word forms because a word form is already provided with all the grammatical information necessary to indicate its function in the syntactic construction. This is the **syntactic autonomy** of the word in Latin. For an initial illustration, compare E6 and E7.

E6            eski şehir-de  
 TURKISH    old    town-LOC  
                  'in an old town'

E7    a.    in antiqu-ō oppid-ō  
 LATIN    in    old-ABL.SG    town(N)-ABL.SG  
           b.    in oppid-ō antiqu-ō  
           c.    antiquō in oppidō

The Turkish adjective attribute precedes its head noun immediately, but bears no sign of attribution. The case of the nominal group is indicated once at its end. By contrast, the Latin adjective attribute may precede or follow its head noun, as in E7a and b, and even be separated from it, as in c, which all mean the same as E6 (cf., however, E14). Its attributive function is not signaled by its syntagmatic position, but by its agreement with the head noun in gender, number and case. The palpable correlate of the Latin word's syntactic autonomy is its well-known versatility in the sentence. The next pair of examples, however, demonstrates the reverse of the coin.

E8            can u gönül-den  
 TURKISH    soul and heart-ABL  
                  'with heart and soul'

E9            anim-ā        et    cord-e  
 LATIN        soul(F)-ABL.SG    and heart(N)-ABL.SG

With the agglutinative technique, the syntactic function of a coordinate noun phrase is signaled once at its end (E6 and E8). The flexional technique redundantly signals it on each member of the coordinate construction (E9). Latin syntax does not form phrases. There is a level of the syntagma between the syntactic levels of the word form and the clause; but it is not held together by contiguity, but by dependency relations among word forms. This topic will be followed up in §2.2.4.

### 2.2.3 Grammaticalization

Meillet (1912) also takes up Schlegel's thinking concerning the synthetic vs. analytic technique. He shows that a periphrastic form evolves from a syntactic configuration of word forms by a general process of directed variation which he calls **grammaticalization**. For instance, the perfect auxiliary *ha* of the periphrastic form *ha cantato* of E5 is grammaticalized from the Latin *habet* 'has', which is yet a lexical verb. Although the grammatical forms of a language generally instantiate layers of degrees of grammaticalization, entire areas of its grammar may be characterized by their degree of grammaticalization.

This idea is rendered fruitful for typology by Eugenio Coseriu (1998). He subdivides the grammatical functions coded in morphology into internal and external determinations of a word. The internal ones concern the word in itself. They comprise such categories as gender and number in the nominal sphere and the simple past in the verbal sphere. The external determinations concern the relation of the word to its syntagmatic context. They comprise such categories as case in the nominal sphere and the compound past in the verbal sphere. With few exceptions, this distinction does not matter for the morphological technique in Latin since all inflection categories are coded synthetically.

Not so in the Romance languages. The Romance prototype obeys a principle: internal determinations are coded synthetically, external determinations are coded analytically. Thus, to the completely synthetic Latin *infantibus* (child<sub>(M)</sub>:DAT.PL) ‘to children’ corresponds Italian *a bambini* (to child<sub>(M)</sub>:PL), where gender and number are coded synthetically, but the dative case is coded analytically. Again, to the synthetic Latin perfect form *cantavit* (E4a) correspond two forms in Italian: the simple past *cantò* ‘he sang’ fixes the tense like an aorist as an inner determination, while the compound past *ha cantato* (E5a) relates it to topic time. In this way, the general trend of the typological change from Latin to the Romance languages, which gradually replaces synthetic forms by analytic forms, is articulated into functionally defined phases.

### 2.2.4 Word order

Linguistic typology got fresh momentum in the middle of the twentieth century when Roman Jakobson’s idea, first proposed in 1941 for phonology, that linguistic subsystems are structured by implicational relations between their components, was applied more widely. Joseph Greenberg (1963) showed that large parts of morphology and syntax, too, are subject to universal implicational generalizations. These are generalizations of the kind ‘if a language has postpositions (rather than prepositions), then it also has genitive-noun order in possessive constructions (rather than the reverse order)’. His **basic order typology** relies on the concept of the basic order: even if in a syntactic construction the relative order of its components is variable, one order may be identified as basic. It is claimed that the basic word-order patterns of a language may be grouped into types held together by implications. There are, indeed, several languages that follow a uniform **serialization** principle: They are either **left-branching**, with the dependent preceding its head, or **right-branching**, with the dependent following the head. The constructions themselves are presupposed to be identifiable across languages. Often the constructions listed in Table 1 are considered:

Table 1 Syntactic constructions of basic order typology

head		dependent	
verb	V	subject	S
verb	V	direct object	O
verb	V	adverb	Adv
head noun	N	possessive attribute	G
head noun	N	adjective attribute	Adj
head noun	N	relative clause	RC
auxiliary	Aux	full verb	V
adposition	Adp	complement	N
conjunction	Conj	dependent clause	Cl

Classical Arabic (Afro-Asiatic, Arabia) is a pure right-branching language: in all of the constructions, the components take the order assigned to the columns of Table 1. Japanese (isolate, Japan) is a pure left-branching language where all of the constructions show the opposite order (cf. E16 below). Beside such “consistently serializing” languages, there are many languages which serialize a subset of these constructions in one direction and the rest in the other direction. Frequently, for instance, the subject precedes the verb, while the direct object follows it. This produces the main constituent

order SVO, typical of the Romance and Germanic languages. Likewise, there are many languages where the basic order is impossible to determine in the first place, i.e. languages with free word order. Classical Latin allows variation in most syntactic constructions (cf. E13 below). Right-branching order is fixed only in the last two constructions of Table 1. The basic order for the relative clause is right-branching; alternative orders are not left-branching, but head-internal. The high degree of variation is described in a diachronic perspective in Adams 1976. Adams claims that basic word order was left-branching at some pre-Latin stage, that it changes to the right-branching order found in the Romance languages and that diachronic and synchronic variation obtains in the intermediate period. A comparison with other ancient Indo-European languages supports a reconstruction by which left-branching order was the default in several Proto-Indo-European constructions, although free word order prevailed similarly as in Ancient Greek and Classical Latin.

Variation in word order and grammaticalization are intimately connected. The most common source of case suffixes are grammaticalized postpositions. Not only does the grammaticalization of postpositional phrases create a paradigm of case declension in a language system; an existent case paradigm can be constantly enlarged and renewed by the grammaticalization of further postpositions. This process can be observed in the history of many languages. Turkish uses postpositions rather than prepositions and has a rich case system inherited from Proto-Turkic. More recently, it has been acquiring an instrumental-comitative case suffix *-le* by grammaticalizing the postposition *ile* 'with'. For example, the construction of E10b may be used as a colloquial alternative to #a.

- E10 a. Elif ile görüş-tü.  
 TURKISH [Elif with] meet-PST(3.SG)  
 'He/she met with Elif.'
- b. Elif-le görüş-tü.  
 Elif-INSTR meet-PST(3.SG)  
 'He/she met Elif.'

It is commonly said that the Romance languages, lacking a case system, use prepositions to fulfill the functions of the Latin cases. In principle, a case system can be protected against loss by the recruitment of further postpositions as demonstrated by E10. However, the ambivalent adpositional adverbs inherited from Proto-Indo-European had been fixed as prepositions already at a Pre-Latin stage. There were therefore, in Latin, no postpositions by which the case system could have been renewed; the source had dried out. Thus, ongoing grammaticalization led simultaneously to the loss of the case system and to the recruitment of prepositions to mark case functions like the genitive and the dative (Lehmann 1985[L]). This is illustrated by E11f.

- E11 a. Paul-i (Paul(M)-GEN.SG) 'of Paul'  
 LATIN b. Paul-o (Paul(M)-DAT.SG) 'to Paul'
- E12 a. de Pablo (of Paul)  
 SPANISH b. a Pablo (to Paul)

As a consequence of high-school teaching, the free word order of Latin has been a commonplace. It should, however, be clear at the outset that this statement must be restricted in several respects. First, word order is much freer in poetry than in prose. E13 is an extreme example.



E13      At uolucres      patulis      residentes      dulcisa      ramis  
but bird(F):NOM.PL wide:ABL.PL.M sit:PTP:NOM.PL.F sweet:ACC.PL.N twig(M):ABL.PL  
carmina      per uarios      edunt      resonantia      cantus.  
song(N):ACC.PL by various:ACC.PL.M utter(PRS):3.PL resound:PTP:ACC.PL.N song(M):ACC.PL  
‘But birds sitting on wide twigs utter sweet songs resounding with varied  
melodies.’ (Appendix Vergiliana, *Culex* 146f)

It is true that the inflectional marking leaves no doubt about the obtaining syntactic relations; s. §2.2.2 on the autonomy of the word. However, as far as Latin is concerned, E13 is rather a poetic exercise than a reflection of common first century speech. Really free word order in common speech is reported from various languages of Australia, including Dyirbal (Pama-Nyungan). The word order in E25f below has been normalized for didactic reasons; actually any permutation of the words produces a grammatical and acceptable sentence. Second, there is, in Latin, a set of syntactic relations whose components have fixed order. This includes the preposition, which always precedes its complement, as it does in E13, and likewise the conjunction, which precedes its clause, although it may be enclitic to its first constituent.

Third, while scrambling as in E13 may look like free variation, in actual speech the order of components of the clause and of lower-level syntagmas follows principles of information structure. This may be observed in the position of the adjective attribute relative to its head (Spevak 2010, ch. 6.2): Postnominal position is the default, including the position appropriate for an attribute with a classificatory function. Thus, ‘from old oil’ would generally be *ex oleo vetere* (Celsus *med.* 3, 23). Prenominal position is chosen for emphasis and contrast, as in E14.

E14      ...      ex      vetere      oleo ...      ex      quolibet      oleo ...  
LATIN      out      old:ABL.SG.M      oil(M):ABL.SG      out      INDF:ABL.SG.M:at.Will      oil(M):ABL.SG  
‘... from old oil ... from any oil’ (Celsus *med.* 3, 11)

Likewise there is a default order inside several others of the binary syntagmas of Table 1. Topicalization then allows for sentence-initial position, while focusing licenses various placements, both following the topic and following the respective head at a distance (§2.2.10).

As the syntax is “formalized” in the course of the language history (Kurzová 1993), this order freedom shrinks. This can be described as the introduction of phrase structure into a grammar which used to abide by syntagmas constituted by relations of dependency and sociation (Danckaert 2017). In this way, Latin acquires, on its way to Romance, a verb phrase in the order verb-object and a noun phrase introduced by an article (Ledgeway 2012, ch. 4f).

By a general principle of grammatical structure, syntagmatic variability of meaningful elements is the freer the higher the level of grammatical structure (Greenberg 1963: 82). For example, a prepositional phrase as a whole enjoys more order freedom in its clause than the complement enjoys inside the prepositional phrase (normally, none). The most catchy formulation of this is Ross’s (1973:397) **Penthouse Principle**: “More goes on upstairs than downstairs.” The diachronic counterpart of this synchronic principle is that once the order at some level of structure has become invariable, it cannot change in the same construction. The only way for a language to change the order of the components in such a low-level construction is to renew the entire construction, typi-

cally by grammaticalizing a higher-level construction which allows for order variation and, thus, for a different order (Lehmann 1992). As a special case of the general principle, consider main constituent order: It is freer in main clauses than in subordinate clauses. At a Pre-Latin stage, the default order was SOV. It was freer in main clauses than in subordinate clauses. In Classical Latin, subordinate clauses have much more regular verb-final order than main clauses, where the verb is often followed by other material including the subject. Main clause order changed to SVO in Vulgar Latin. SOV order remained more stable in subordinate clauses. It changed there, too, by remodeling the entire constitution of subordinate clauses: the inherited conjunctions were abolished (Meillet 1915), and new ones were combined with the constituent order meanwhile in use in the main clause.

### 2.2.5 Modification and government

Yet another insight by Meillet (1934: 357-359, 440) proved fruitful. It presupposes a division of dependency relations into government and modification. In **government**, the head (the *regens*) determines the syntactic function of the dependent (the *rectum*) and its properties. In **modification**, the relation of the dependent to the head is brought about by the former, the modifier, and its properties are freely chosen in consonance with the semantics of the construction. The dependents of the predicate verb of a clause may be governed by it, which means that verbal **valency** determines the shape of the construction (Lehmann 2002). Alternatively, the dependent may not be comprised by the verb's valency; then it acts as a modifier and is typically provided with a case relator which codes the particular relation that the dependent carries to the verb.

In E15 from Abkhaz (North-West Caucasian, Abkhazia), the verb has prefix slots reserved for its valency relations. The morphemes inserted there point to the dependents by indicating their grammatical categories. The dependents themselves are not case-marked in any way; their syntactic relations emanate from the verb form.

E15            a-x°əč'-k°à    a-š°q°'-k°à    Ø-rə-s-to-yt'.  
 ABKHAZ        ART-child-PL    ART-book-PL    ABS.3-DAT.3.PL-ERG.1.SG-give.DYN-FIN  
                   'I give the children the books.' (Hewitt 1979: 105)

In E16, the verb has the semantic relationality inherent in its concept just as in Abkhaz. However, this is not grammaticalized in the form of valency. Each of its dependents displays its own semantic role vis-à-vis the verb by a case clitic.

E16            Watasi=ga kodomo=ni hon=o yar-u  
 JAPANESE     I=NOM            child=DAT        book=ACC give-PRS  
                   'I give the children the books.'

Meillet assumes that Proto-Indo-European was like Japanese in that the verb exerted no government on its dependents and instead these were added freely as modifiers, marked by the case fitting the particular situation. Its daughters gradually acquired valency. In the Latin version E17, the valency of the verb comprises a subject, a direct and an indirect object as in Abkhaz. The two objects are still case-marked, but the subject is marked on the verb as in Abkhaz.

E17	Liberis	libros	do.
LATIN	child(M):DAT.PL	book(M):ACC.PL	give(PRS):1.SG
	'I give the children the books.'		

In the further development, several Romance languages will acquire pronominal indexes on the verb which cross-reference its dependents, while the latter bear no case. Latin is in the middle of this development. This is an aspect of the “formalization” of its grammar (§2.2.4) because the semantic relations inherent in the verbal lexeme are grammaticalized to become structural relations.

### 2.2.6 Head marking and dependent marking

This approach, again, is further developed in Milewski 1950. He uses the same sort of example as E15f to establish two types of syntactic structure, with particular respect to the marking of grammatical relations. In the **concentric type**, these relations are marked on the head of the construction. In a verbal clause, this is the verb, as in E15. However, the same principle may be operative in at least two other constructions of Table 1: If possessive attribution is a concentric construction, the head noun bears a pronominal index cross-referencing the possessive attribute, which latter may be left unmarked (E18); and likewise in the adpositional phrase, the adposition cross-references its dependent, which is a bare noun phrase (E19). Contrariwise, in the **eccentric type**, grammatical relations are marked on the dependent of the construction. In a verbal clause, the dependents are the main nominal constituents, case-marked or converted into adpositional expressions, as in E16. In an eccentric construction of possessive attribution, the possessor is marked by the genitive, while the possessum needs no mark (E20). Likewise in an adpositional phrase, a bare adposition governs a cased NP (E21).

E18	u	ne'	péek'
YUCATEC	POSS.3	tail	dog
	'the dog's tail'		

E19	a-jèyas	a-q'+nè
ABKHAZ	ART-river	OBL.3.SG.NHUM-at
	'by the river' (Hewitt 1979: 103)	

E20	canis	cauda
LATIN	dog(M):GEN.SG	tail(F):NOM.SG
	'the dog's tail'	

E21	apud	fluvium
LATIN	by	river(M):ACC.SG
	'by the river'	

Nichols (1986) develops this typology further and calls the two types **head-marking** and **dependent-marking**, resp. In the series E18 – E21, Latin illustrates the dependent-marking construction type. In comparison with thoroughly head-marking languages like Abkhaz and Yucatec (Mayan, Mexico), on the one hand, and with thoroughly dependent-marking languages like Japanese, on the other, it gets evident that Latin is predominantly dependent-marking; it is head-marking only in that one relation, viz. the subject relation, in which most languages are head-marking (E17). As already noted, several of the

Romance languages carry this incipient development further by cross-referencing verbal objects and occasionally even the possessive attribute on the head (Ledgeway 2012, ch. 6).

### 2.2.7 Alignment

Most languages have grammaticalized the semantic roles associated with verbs in the form of syntactic relations formalized in valency patterns (Christol 1998). Their structural correlates include obligatoriness/optionality of the actants, their syntagmatic position (as in English), case marking (as in E17 and E25f) and pronominal indexes on the verb (as in E15 and E22 – E24). Among bivalent verbs, the configuration in which an actor controls an undergoer is the most important one and founds the **transitive valency frame** in many languages. The only actant of a monovalent verb can be an actor or an undergoer, too. This distinction, however, is made in grammatical structure only in a minority of languages. The Lakhota (Sioux, U.S.A.; Van Valin 1977: 8f) verb cross-references actor (A) and undergoer (U) by prefixes and infixes as illustrated by E22 – E24.

- E22 a. ma-hã'ske (u.1.sg-tall) 'I am tall'  
 LAKHOTA b. ni-hã'ske (u.2-tall) 'you are tall'
- E23 a. wa-hi' (A.1.sg-arrive) 'I arrive'  
 LAKHOTA b. ya-hi' (A.2-arrive) 'you arrive'
- E24 a. wã<chi>yãnke (see<u.2&A.1.sg>) 'I see you'  
 LAKHOTA b. wã<ma'-ya>lake (see<u.1.sg-A.2>) 'you see me'

E22 and E23 show a stative and an active monovalent predicate, resp., with prefixes of the undergoer and actor series, resp. E24 shows a two-place predicate with infixes both for the undergoer and the actor. While the person combination in E24a is expressed by a portmanteau morph, the person combination in #b is coded as would be predicted on the basis of E22a and E23b. Concerning the **alignment** of valency relations of intransitive and transitive verbs, the undergoer of the intransitive verb is, thus, assigned the same syntactic function as the undergoer of the transitive verb, and likewise for the actor of both verb types. This grammatical arrangement of the two macro-roles is called the active type in Klimov 1977, but has been called variously split intransitivity, split-S<sup>2</sup> or **active-stative alignment** in subsequent typology.

In most languages, the actor and undergoer roles are not distinguished with intransitive verbs. Instead, the actant of the intransitive verb is aligned with either the actor or the undergoer of the transitive verb in one syntactic function. The second most frequent type is **ergative alignment**. It provides one syntactic function, called absolutive, for the actant of the intransitive verb and the undergoer of the transitive verb, and a different syntactic function, called ergative, for the actor of the transitive verb. E25f from Dyirbal (Pama-Nyungan, Australia; Dixon 1972: 59) illustrate this.

- E25 a. bayi yaɾa bani-ɟu  
 DYIRBAL DET(ABS):M man(ABS) come-REAL  
 'man is coming'

<sup>2</sup> Since S means 'subject', this term is rather eurocentric, as these languages lack a subject that they could split.

- b. bala-n djugumbil bani-ɲu  
 DET(ABS)-F woman(ABS) come-REAL  
 ‘woman is coming’
- E26 a. bayi yaɾa ba-ɲgu-n djugumbiɾu balga-n  
 DYIRBAL DET(ABS):M man(ABS) DET-ERG-F woman:ERG hit-REAL  
 ‘woman is hitting man’
- b. bala-n djugumbil ba-ɲgu-l yaɾa-ɲgu balga-n  
 DET(ABS)-F woman(ABS) DET-ERG-M man-ERG hit-REAL  
 ‘man is hitting woman’

The verb in E25 is intransitive, and its actant is in the absolutive case, zero in most ergative systems (though not in E31 below). The verb in E26 is transitive. The undergoer in the #a and #b versions bears the same absolutive function as in the corresponding versions of E25, while the actor is marked by the ergative.

A translation of the latter example series into Latin produces E27f.

- E27 a. uir ueni-t  
 LATIN man(M.NOM.SG) come(PRS)-3.SG  
 ‘man is coming’
- b. mulier ueni-t  
 woman(F.NOM.SG) come(PRS)-3.SG  
 ‘woman is coming’
- E28 a. uir-um mulier feri-t  
 LATIN man(M)-ACC.SG woman(F.NOM.SG) hit(PRS)-3.SG  
 ‘woman is hitting man’
- b. mulier-em uir feri-t  
 woman(F)-ACC.SG man(M.NOM.SG) hit(PRS)-3.SG  
 ‘man is hitting woman’

The actant of the intransitive verb in E27 is in the nominative, which is not always a zero case in Latin (cf. E2c), but is in these nouns, as it is regularly in many other languages of this alignment type. The actor of the transitive verb in E28a and b is marked by the same case as the actant in the corresponding examples of E27; it bears the subject function. The undergoer is marked by the accusative and bears the direct object function. This is therefore called the **accusative alignment**. It is the most frequent one world-wide.

These are the three major alignment types (WALS ch. 98-100). There are a few minor patterns, known from very few languages. Latin here represents the majority type (Ledgeway 2012, ch. 7). However, the alignment types enjoy variable preference depending on the type of marking of syntactic relations by the criteria of §2.2.6. Active-inactive alignment is almost as frequent as ergative alignment if the constructions are concentric, while it is rare if they are eccentric. Again, ergative alignment is relatively frequent in eccentric constructions, but less so in concentric ones, where accusative alignment is even stronger than in eccentric ones.

If the marking of syntactic relations in the main clause shows a certain alignment type, this does not mean that the entire syntax of the language works by the same principle. Different syntactic constructions favor different alignments (Lehmann 1985[E]). For instance, imperative constructions universally show active-stative alignment in the sense

that they are limited to agentive verbs or at least can be used as true commands only with these. Latin has a class of stative verbs in the *ē*-conjugation. Table 2 contains some examples and also comprises some stative verbs of the consonantal conjugation. Their imperative, e.g. *latē* ‘hide’, would be quite regular morphologically, but is not documented in Old or Classical Latin.

Table 2 Stative Latin verbs

<i>ē</i> -stem		other	
infinitive	meaning	infinitive	meaning
<i>carēre</i>	lack	<i>fluere</i>	flow
<i>feruēre</i>	seethe	<i>senescere</i>	grow old
<i>latēre</i>	hide		
<i>liuēre</i>	be blue		
<i>patēre</i>	stand open		

The same holds for passivization. While Latin is one of the relatively few languages which passivize intransitive verbs, this is possible only for agentive verbs, e.g. *curritur* (run:(PRS)PASS.3.SG) ‘one is running’, lit. ‘it is being run’. The passive of the verbs in Table 2, e.g. *latētur* (hide:(PRS)PASS.3.SG) ‘one is hiding’, would be morphologically quite regular, but does not occur in the corpus. Thus, the actant which is demoted by the Latin passive is always an actor, whether the verb is transitive or intransitive.

Latin (like other languages) shows ergative alignment in preverbatation. A preverb like *ex* is relational, it means ‘A [moves] out of B’. B is some local syntagma in the clause; in E29, it is *urbe*. A is an actant of the verb provided with the preverb.

E29	a.	Caesar	ex-it	urbe.	
LATIN		Caesar(M):NOM.SG	out-go(PRS):3.SG	city(F):ABL.SG	
		‘Caesar goes out of the city.’			
	b.	Caesar	legiones	e-ducit	urbe.
		Caesar(M):NOM.SG	legion(F):ACC.PL	out-lead(PRS):3.SG	city(F):ABL.SG
		‘Caesar leads the legions out of the city.’			

In E29a, the entity A whose locomotion is implicit in the preverb is the subject of the verb. In #b, this role is taken by the undergoer of the verb. Thus, A is identified with the actant of an intransitive verb and with the undergoer of a transitive verb. This is so with all Latin preverbs which mean ‘A is localized with respect to B’; they show ergative alignment.

The idea of alignment typology has been extended to **ditransitive verbs**, i.e. verbs whose valency includes an actor, a transferred entity and an entity which is prototypically a recipient, thus verbs whose prototype is ‘give’ (Haspelmath 2005, WALS ch. 105). Criteria to define alignment types are applied in a way analogous to the analysis of transitive constructions. There are, again, three alternative constructions. In **indirective alignment**, the transferred entity is aligned with the undergoer of the transitive construction, while a distinct syntactic relation, viz. the indirect object, is provided for the recipient, as in E30.

E30	a.	corbem	texuit
LATIN		basket(M):ACC.SG	weave:PRF:3.SG
		‘he wove the basket’	

- b. corbem            infanti            dedit  
 basket(M):ACC.SG   child(M):DAT.SG   PRF:give:3.SG  
 'he gave the basket to the child'

In **secundative alignment**, it is the recipient which is aligned with the transitive undergoer in a function called primary object, while the transferred entity is marked as a secondary object, as in E31 from Chamorro (Austronesian, Mariana Islands; adapted from WALS ch. 105).

- E31 a. Ha    tuge' i    kannastra.  
 CHAMORRO   he.ERG   weave   ABS   basket  
 'He wove the basket.'
- b. Ha    na'i i    patgon    ni    kannastra.  
       he.ERG   give   ABS   child    OBL   basket  
 'He gave the basket to the child.'

In **neutral alignment**, both the recipient and the transferred entity are marked like the transitive undergoer, with the result of a double object construction, as in E32 from Panyjima (Pama-Nyungan, Australia; WALS ch. 105).

- E32 a. Ngunha    parnka    ngarna-rta    mantu-yu.  
 PANYJIMA   that        lizard        eat-FUT        meat-ACC  
 'That lizard will eat the meat.'
- b. Ngatha    yukurru-ku    mantu-yu    yinya-nha.  
       I.NOM        dog-ACC        meat-ACC        give-PST  
 'I gave the dog meat.'

With the alignment of E30, Latin sides with half of the world's languages. However, while this is cross-linguistically the majority valency frame just for the verb 'give', there is wide variation in the size of the set of verbs which share it in a given language. In interlingual comparison, Latin is one of relatively few languages which generalize the indirective ditransitive pattern over a large set of verbs.

### 2.2.8 Word classes and categoriality

Just as most languages, Latin possesses major (productive) and minor (closed) word classes. The former include nouns, verbs, adjectives and adverbs. The latter are relatively few. Among the classes of grammatical formatives which are widespread in the world, but lacking from Latin, are numeral classifiers, possessive classifiers and articles, whether definite, indefinite or specific, light verbs and auxiliaries with the exception of *esse* in some passive forms.

Languages differ in the grammatical level at which expressions are classified rigorously as to part of speech (Lehmann 2008). This is only really necessary at the level of the syntagma – if only by word order and prosody –, but may already be done at lower levels. In some languages, even most of the roots are categorized for word class; German is an example. In other languages, only a stem derived from a root belongs to a word class; this is the case of Latin. In yet other languages, even a lexeme of this structural level has yet flexible class-membership (Hengeveld 1992); and every once in a while a language is claimed to lack word classes altogether (Evans & Osada 2005). In Latin, the difference in categoriality between roots and stems is striking. A root like *tim-* 'fear' can-

not be inflected and belongs to no category. The stem *time-* can only be conjugated and thus is a verb. The stem *timor* can only be declined in one gender and therefore is a noun. The stem *timido-* ‘fearful’ can only be declined through the three genders and therefore is an adjective. This distance between root and stem shrinks in Romance languages as roots are more category-specific while stems are less categorial than in Latin.

### 2.2.9 Complex sentences

While all languages form complex sentences by parataxis, they differ considerably in their strategies of hypotaxis (Lehmann 1989). The main distinction here lies in the finiteness of the verb of the subordinate clause. Many languages are like Turkish in chiefly building subordinate clauses on the basis of non-finite verb forms, so these display various symptoms of desententialization. Latin does have non-finite verb forms, viz. the infinitive, the gerund and gerundive, the participles and supines. In the literary style, participles may be expanded into syntactic constructions known as *participium coniunctum* and *ablativus absolutus*. These do not survive into Romance, but are partially replaced by gerundial constructions. The infinitive may be expanded into a *nominativus* or *accusativus cum infinitivo*, which functions as a complement clause. Such infinitival complements are more wide-spread in the world’s languages. What is special about Latin complex sentence formation is that all kinds of subordinate clauses are finite; they are introduced by a large set of subordinative conjunctions and show few or no symptoms of desententialization. The formation of highly complex sentences, with embedding at any desired depth – the “Ciceronian period” – is a hallmark of the classical style and provided the ideal for European prose far into the 20th century. It is very rare world-wide. Instead, many languages use clause chaining, where semi-finite clauses are concatenated and only the last clause is fully finite (Dooley 2010).

The Latin relative clause, being postnominal, follows a majority model, but is in a distinct minority with its relative pronoun. Almost all languages which use a relative pronoun are SAE languages (WALS ch. 122, 123); and many of them do this on the basis of a calque from Latin.

### 2.2.10 Information structure

Information structure involves the marking of components of a sentence as topic and focus. Apart from prosody, this may be achieved by assigning them positions in the **sentence topology**. If a language has strict word order, it transgresses the confines of the clause in order to attain a level of order freedom (cf. §2.2.4). The most important complex constructions then resorted to are left-dislocation and cleft-sentences. Left-dislocation is used in the translation of E33.

E33    signa        nostra    et    Hermeraclas,                    ut    scribis,  
LATIN   statue:ACC.PLOUF:ACC.PL and Mercury:Hercules:ACC.PL as    write:(PRS)2.SG

      cum commodissime poteris uelim                    imponas ...  
      when convenient:SUPL:ADV can:FUT:2.SG want:SBJ.PRS:1.SG    upload:SBJ.PRS.2.SG

‘As for our statues and bust of Mercury and Hercules, I would want you, as you say, to send them at your earliest convenience ... (Cic. *Att.* I, 10, 3)



The sculptures form the third of a list of topics treated in this letter. Their syntagma is moved into initial position, its function as the direct object in an object complement clause notwithstanding. Romance languages use a left-dislocation construction with an anaphor (Span. *cuanto a ...*) to render this.

Many languages use a cleft-sentence to mark contrastive focus. Latin only has a non-grammaticalized variant of it, viz. a relative construction. E34 contrasts the present focus with two alternative foci treated in the following text.

E34 Interdum populus est quem timere debeamus ...  
 LATIN sometimes people:NOM.SG is REL:ACC.SG.M fear:INF must:SBJ.PRS.1.PL  
 ‘Sometimes it is the people that we have to fear ...’ (Sen. *Epist.* 14, 7)

This is the strategy which is further expanded and grammaticalized in the Romance languages. Latin still has the alternative possibility of just moving the focus into front position (E35).

E35 Consulem ego tum quaerebam ...  
 LATIN consul:ACC.SG 1.SG then seek:IMPF:1.SG  
 ‘It was a consul that I was then looking for ...’ (Cic. *Pis.* 19, 4)

On balance, Latin makes limited use of complex constructions in information structure, relying mainly on the syntactic autonomy of the word (Spevak 2010, ch. 2).

### 2.2.11 Other grammatical properties

In many respects, the structure of Latin is quite unremarkable in interlingual comparison. With each of the following properties, Latin joins the majority of the languages of the world:

- The same pronominal forms function both as substitutes for an NP and as determiners. I.o.w., demonstratives, indefinites, interrogatives etc. have both pronominal and adnominal uses.
- Indefinite pronouns are based on interrogative pronouns.
- There is no politeness distinction in second person pronouns.
- Head marking (or double marking) is used for the subject relation, but dependent marking for all the other syntactic and semantic relations (§2.2.6).
- Inflectional morphology is predominantly suffixing, to the detriment of other affix positions.
- Case markers are suffixes. Excluding the vocative, the paradigm comprises six cases. If the locative, of limited applicability, is discounted, too, it has just average size (WALS ch. 49).
- Adverbial subordinators are words preceding the subordinate clause.
- Polar interrogatives are marked by an interrogative particle (WALS ch. 116).

On the other hand, Latin is in the minority with one positive property: It has a passive construction based on a synthetic conjugation category. This is among the first conjugation categories lost on the way to Romance. Apart from the lack of certain classes of grammatical formatives (§2.2.8), Latin lacks the following categories, otherwise shared by sizable portions of the world’s languages:

- There is no inclusive/exclusive distinction in pronominal elements.
- There are no morphological distinctions in possessive relations, no alienability, control or temporal features.

- There is no evidentiality, i.e. indication, at the level of grammar, of the source of evidence on which some statement is based.
- There is neither a synthetic nor a periphrastic causative construction which would be productive (§2.3.1).

These are the most important domains in which Latin grammar is poor in interlingual comparison.

## 2.3 Lexical typology

Morphological structure is introduced into the lexicon by word formation. Although there is a modest set of compound nouns like *aequ-i+libr-i-um* (equal- $\emptyset$ +balance- $\emptyset$ -NOM.SG) ‘equilibrium’ and *sign-i+fer* (sign- $\emptyset$ +bear) ‘standard-bearer’, Latin lacks a productive process of compounding. Contrariwise, derivation is highly developed. The system is especially rich in processes of abstraction, producing nouns like *natio* ‘nation’, *amplitudo* ‘amplitude’, *curiositas* ‘curiosity’ *avaritia* ‘avarice’, many of which have been borrowed into languages with poorer derivational morphology.

From a semantic point of view, the lexicon may be considered structured at two levels. At the upper level, the question is in which syntactic category, i.o.w. in which word class or subclass, concepts of a certain conceptual domain are lexicalized. For instance, states can be lexicalized as adjectives or as (stative) verbs (Table 2). At the lower level, the question is about the internal structure of such a conceptual domain, viz. about the semantic features which distinguish its members.

As examples of the upper level of lexical structure, base transitivity (§2.3.1) and spatial regions (§2.3.2) will be considered. As an example of the lower level, color and kinship terminology are briefly discussed in §2.3.3.

### 2.3.1 Base transitivity

There is a set of dynamic relational concepts (those which are preferably lexicalized as verbs), including, e.g., ‘burn’, which involve an undergoer, but are equally compatible with the presence (like ‘A burns U’) and the absence (like ‘U burns’) of an actor. These are typically concepts of position (e.g. Latin *sedeo* ‘I sit’), motion (e.g. *uenio* ‘come’), physical change (e.g. *ardeo* ‘burn’), emotion (e.g. *timeo* ‘fear’) and phase (e.g. *incipio* ‘begin’). In the lexicalization of such concepts as verbs, there are various possibilities of reflecting the paradigmatic relationship of the intransitive and the transitive version in lexical structure (Nichols et al. 2004). The concept may be primarily coded as an intransitive verb, and a process of **causativization** may form its transitive counterpart. This is the case in E36, where the sentence in #b is a causative version of #a.

- E36 a. Mary=ga odorok-u.  
 JAPANESE Mary=NOM be.scared-PRS  
 ‘Mary is scared.’
- b. John=ga Mary=o odorok-as-u.  
 John=NOM Mary=ACC be.scared-CAUS-PRS  
 ‘John scares Mary.’

- E37 a. Maria terre-tur.  
 LATIN Mary(F):NOM.SG scare(PRS)-PASS.3.SG  
 'Mary is scared.'
- b. Johannes Mariam terre-t.  
 John(M):NOM.SG Mary(F):ACC.SG scare(PRS)-3.SG  
 'John scares Mary.'

The mirror-image to this strategy is the primary lexicalization of the concept as a transitive verb, as in E37b. The intransitive version is then formed by some operation of **deagentivization**. In E37a, this is passivization. Both languages possess, in addition, less transparent patterns of the lexical relation in such transitive-intransitive pairs. Among other things, Latin has such morphologically equipollent pairs as *expergefacio* – *expergiscor* for transitive and intransitive 'wake up' and such morphologically unrelated pairs as *doceo* 'I teach' – *disco* 'I learn'. However, while Japanese codes most of those concepts as in E36, Latin codes most as in E37; further examples analogous to E37 include *uro* – *uror* 'burn', *rumpo* – *rumpor* 'break' and *moueo* – *moueor* 'move'. Morphologically marked transitive counterparts, as in *pateo* 'I am open' – *patefacio* 'I open [sth.]', are a distinct minority. Labile verbs like the English translations of the examples are absent from Latin.

To the extent that a language follows a consistent strategy, one may speak of its base transitivity: What prevails in Latin is **base transitivity**, while **base intransitivity** prevails in Japanese. Base transitivity is a minority strategy – not a particularly efficient one, by the way – found mainly in European languages. One typological correlate to this property of Latin is the weak role played by causativization. The formation of the type *patefacio* involves compounding; but compounding is disfavored in Latin. Moreover, using the passive for deagentivization enhances vagueness: *moueor* means both 'I move' and 'I am moved'. The same goes for Ancient Greek, in which base transitivity is even more pronounced than in Latin.

### 2.3.2 Spatial regions

Spatial regions are geometrical aspects of a physical object defined by its dimensions and its topology. Most languages lexicalize these notions as **relational nouns**, like English *top*, *bottom*, *front*, *back* etc. Latin here is in a distinct minority by lacking most of such nouns and instead categorizing such notions primarily as adverbs and prepositions. Thus, the notions just enumerated are designated by the adverbs and prepositions *supra*, *infra*, *ante* and *post*. In order to designate a spatial region nominally, one first derives an adjective from such an adverb or preposition, which may be in the comparative, like *superior* 'upper', *inferior* 'lower', *anterior* 'front', *posterior* 'back', resp. The relational noun *pars* 'part' may be modified by such an adjective, e.g. *pars superior* 'top', forming thus a complex designation of a spatial region. If a spatial region of a particular physical object is to be designated, the noun designating it is modified by the superlative of such an adjective (Lehmann 1998).

For instance, the Japanese E38 uses a basic relational noun to designate the spatial region 'top' and converts this into a complex postposition by using it in the locative. Latin, instead, modifies the noun designating the reference object by a superlative adjective designating the spatial region (E39).

E38            yama=no      ue=de  
 JAPANESE    mountain=GEN top=LOC  
                  'on (top of) the mountain'

E39            in summ-o          mont-e  
 LATIN        in highest-ABL.SG.M    mountain(M)-ABL. SG  
                  'on top of the mountain' or 'on the highest mountain'

The resulting Latin construction is ambiguous. The superlative presupposes a set of objects among which the one possessing the scalar property to the highest degree is selected. This set could be either the parts of the reference object (first reading of E39) or other reference objects of the same kind (second reading).

The concept of the spatial region is lexicalized as a relational word in both languages. In Japanese, it is a relational noun which governs the noun of the reference object, while in Latin, it is an adjective which modifies this noun. This is another symptom of the general preference that Latin gives to modification over government (§2.2.5).

### 2.3.3 Lexical fields

In most languages including Latin, the primary lexicalization of color concepts is in terms of adjectives. Visual perception of colors involves a set of parameters: hue is based on wave length, saturation is the absence of white, brightness is the absence of black; moreover, there are parameters like color temperature and gloss. These are relevant to different degrees in the color terminology of different languages. In many languages, the negative poles of saturation, i.e. white, and of brightness, i.e. black, are projected onto one axis with the hues, producing the basic color terminology hierarchy of Berlin & Kay 1969. Latin differentiates minimal saturation and minimal brightness by gloss: *candidus* is 'brilliant white', while *albus* is just 'white'; and again *niger* is 'brilliant black', while *ater* is just 'black' (Viti 2020, ch. 3). On the axis of hue, the same distinction applies to some of the hues. Thus, *flauus* is 'brilliant yellow', while *giluus* is 'dull yellow'; *caeruleus* is 'brilliant blue', while *liuidus* is 'dull blue'. Internal lexical structure is only loosely connected with grammar, but more so with culture. Thus, lexical typology has chances to build a bridge between language and culture.

In kinship terminology, Latin distinguishes mother's siblings from father's siblings. Mother's brother is *auunculus*, her sister is *matertera*; father's brother is *patruus*, his sister is *amita*. English merges the parents' brothers and sisters in the lexemes *uncle* and *aunt*, resp. Latin kinship terminology has been studied repeatedly (e.g. Szemerényi 1977) and is an example of a neutral system which provides a simple kin term for most individual relationships in the system.

## 3 Outlook

Since Latin linguistics essentially emerged from Latin philology, there was initially little contact with general linguistics. To this day, much of Latin linguistics continues to be philology tackling linguistic problems. However, since the last third of the 20<sup>th</sup> century, an increasing number of Latin linguists have a background both in classical philology and in general linguistics. As a consequence, descriptive approaches have changed thoroughly in the past decades, in several respects.

While traditional work almost always takes the **semasiological** perspective and typically focuses on a Latin word or morpheme whose uses or fate in the corpus are documented and described, more recent work often takes an **onomasiological** approach. It starts from a conception of cognitive and communicative functions fulfilled by languages, selects a functional domain and investigates its structure in Latin. Typical examples of this include Fedriani 2014, which takes an onomasiological and partly typological approach, starting from a conception of the experiential situation and its components, sketching the different patterns that render it and describing the Latin strategies on this background. Another such study is Baldi & Nuti 2010, which starts from a comprehensive theoretical base explicating the concept of possession and then pursues its manifestations in the Latin corpus.

Still as a consequence of its philological origins, work on Latin linguistics used to be monolingual. If Latin was ever compared with any other language, it was Greek, not for any systematic reason, but because of their cultural contact. This has, again, changed. If taken seriously, the onomasiological approach requires the consideration of very different languages; otherwise the peculiar nature of the Latin constructions cannot be detected and insightfully described. The articles by Lehmann (1985, 1989, 1998, 2005, 2016), which investigate subdomains of Latin grammar “in typological perspective”, are a case in point. Certainly the most weighty representative of the new approach is the four-volume set by Baldi & Cuzzolin (eds.) 2009-2011, which explicitly aims at describing Latin from a functional-typological perspective.

Although Latin is one of the most thoroughly and extensively studied languages in the world, some aspects are yet understudied, especially because the onomasiological perspective has only been taken of late. Three fields may be mentioned which require deeper analysis. The first of these is the nature of **syntactic relations**. Government is based on the semantic relationality of the *regens*, which selects the *rectum*. Modification is based on a phoric relation between the modifier and the modified, which attributes the meaning of both to the same designatum. In the history of Latin, both of these types of relation start out at a low degree of grammaticalization. Thus, semantic relationality can work at a syntagmatic distance, and a syntactic combination of the *regens* and the *rectum* in one phrase is not required. Likewise the phoric relation starts out as a distance relation, so a syntactic combination of modifier and modified in one phrase – e.g. in adjective attribution – is not presupposed. The “formalization” of Latin grammar throughout its diachrony leads to the effect that these relations become ones of syntagmatic adjacency. This has never been investigated in typological perspective.

Second, Latin shares company with one third of the languages which have neither definite nor indefinite articles (WALS ch. 37f). It has been recognized for some time that some of the functions of referentiality that are fulfilled by articles in other languages may be fulfilled by word order in Latin. In particular, word order codes information structure (§2.2.10); and information structure is at least an aspect of what is coded by articles. On the other hand, Latin has an elaborate system of pronouns and determiners which code manifold shades of (in-)definiteness and (non-)specificity. These paradigms are less grammaticalized than pure articles. In this perspective, the lack of articles in Latin appears less as a defect and more as a facet of a system which combines a rich and weakly grammaticalized system of pronouns and determiners with a non-segmental system of information structure. The interplay of these two systems has seldom been studied, let alone at the typological level.

Third, Latin is one of the languages which possess a grammaticalized **presentative construction**, illustrated by E40.

E40	Ecce nos tibi oboedientes!
LATIN	PRSV 1.PL.ACC 2.SG.DAT obey:PTP:ACC.PL.M
	'Here we are, obeying you!' (Pl. <i>Mil.</i> 611)

Presentative (or ostensive) particles are found in Romance languages, including *inter alia* Italian *ecco* and French *voici/voilà*, in other Indo-European and in Non-Indo-European languages. Other languages, including English and German, lack them completely. They played an important role in Latin discourse, to judge not only from the Plautinian texts, but also from their use in the formation of Romance demonstratives: Vulgar Latin *\*eccu istu* and *\*eccu illu* produce Italian *questo* 'this' and *quello* 'that', among many other such formations. There is no descriptive study of presentatives in Latin, let alone a typological study.

More generally, several of the domains of grammar discussed above have been felt to be of typological relevance by their analysts. However, in most cases, comparative studies limit themselves to a comparison of a Latin subsystem with isofunctional subsystems of other languages. Systematic correlations among one such subsystem and the rest of the language system remain to be discovered.

### Abbreviations

1, 2, 3	1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> person	INSTR	instrumental
A	actor	LOC	locative
ABL	ablative	M	masculine
ABS	absolutive	N	neuter
ACC	accusative	NHUM	non-human
ART	article	NOM	nominative
CAUS	causative	OBL	oblique
DAT	dative	PASS	passive
DEF	definite	PL	plural
DET	determiner	POSS	possessive
DYN	dynamic	PRF	perfect
ERG	ergative	PRS	present
F	feminine	PRSV	presentative
FIN	finite	PTP	participle
FUT	future	PST	past
GEN	genitive	REAL	realis (non-future)
IMPF	imperfect	SG	singular
IND	indicative	U	undergoer
INDF	indefinite		

### References

- Adams, James N. 1976, "A typological approach to Latin word order." *Indogermanische Forschungen* 81:70-100.
- Baldi, Philip & Nuti, Andrea 2010, "Possession". Baldi & Cuzzolin (eds.) 2009-2011, vol. 3: 239-387.

- Baldi, Philip & Cuzzolin, Pierluigi (eds.) 2009-2011, *New perspectives on historical Latin syntax*. Volume 1: *Syntax of the sentence*; volume 2: *Constituent syntax: Adverbial phrases, adverbs, mood, tense*; volume 3: *Constituent syntax: quantification, numerals, possession, anaphora*; volume 4. : *Complex sentences, grammaticalization, typology*. Berlin: Mouton de Gruyter.
- Ballester, Xaverio 1998, „La tipología y la fonología latina.“ *Actas del Congreso Español de Estudios Clásicos* 9: 3-19.
- Berlin, Brent & Kay, Paul 1969, *Basic color terms. Their universality and evolution*. Berkeley & Los Angeles: University of California Press.
- Christol, Alain 1998, „Typologie des langues classiques (latin, grec).“ Feuillet, Jack (ed.), *Actance et valence dans les langues de l'Europe*. Berlin & New York: Mouton de Gruyter (Empirical Approaches to Language Typology, EURO-TYP, 20-2); 753-767.
- Coseriu, Eugenio 1968, "Sincronía, diacronía y tipología." *Actas del Congreso Internacional de Lingüística y Filología Románica* 11:269-281.
- Coseriu, Eugenio 1980, "Der Sinn der Sprachtypologie." *Travaux du Cercle Linguistique de Copenhague* 20:157-170.
- Coseriu, Eugenio 1998, "Typologie des langues romanes." Rabassa, Lídia & Roché, Michel (eds.), *Variation linguistique et enseignement des langues. Langue parlée, langue écrite*. Toulouse: Université de Toulouse II - Le Mirail (*Cahiers d'Études Romanes* 9, 1997); 217-251.
- Danckaert, Lieven 2017, *The development of Latin clause structure. A study of the extended verb phrase*. Oxford & New York: Oxford University Press.
- Dixon, Robert M.W. 1972, *The Dyirbal language of North Queensland*. Cambridge & London: Cambridge University Press (Cambridge Studies in Linguistics, 9).
- Dooley, Robert A. 2010, Exploring clause chaining. (SIL Electronic Working Papers 2010-001). Dallas, TX: SIL International.
- Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013, *The World Atlas of Language Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info>, accessed on 2020-07-08).
- Evans, Nicholas & Osada, Toshiki 2005, "Mundari: The myth of a language without word classes." *Linguistic Typology* 9:351-390.
- Fedriani, Chiara 2014, *Experiential constructions in Latin*. Leiden etc.: E.J. Brill (Brill's studies in historical linguistics, 3).
- Greenberg, Joseph H. 1963, „Some universals of grammar with particular reference to the order of meaningful elements.“ Greenberg, Joseph H. (ed.), *Universals of language. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*. Cambridge, MA: MIT Press; 58-90.
- Haspelmath, Martin 2001, "The European linguistic area: Standard Average European." Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.), *Language typology and language universals. An international handbook*. 2 volumes. Berlin: W. de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 20); 1492-1510.
- Haspelmath, Martin 2005, "Ditransitive constructions". *Annual Review of Linguistics* 1(1):19-41.
- Hengeveld, Kees 1992, "Parts of speech." Fortescue, Michael & Harder, Peter & Kristoffersen, Lars (eds.), *Layered structure and reference in a functional perspective. Papers from the functional grammar conference, Copenhagen, 1990*. Amsterdam & Philadelphia: J. Benjamins (Pragmatics & Beyond, New Series, 23); 29-55.
- Hewitt, B. George 1979, *Abkhaz*. Amsterdam: North-Holland (Lingua Descriptive Studies, 2).
- Humboldt, Wilhelm von 1836, *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechtes (= Über die Kawi-Sprache auf der Insel Java. Einleitung)*. Berlin: Königlich-Preussische Akademie der Wissenschaften; in Kommission: Bonn etc.: F. Dümmler. Repr.: *Schriften zur Sprachphilosophie [= Werke in fünf Bänden]*, ed. by A. Flitner und K. Giel, Bd.III]. Darmstadt: Wissenschaftliche Buchgesellschaft 1963: 368-756.
- Jakobson, Roman 1941, *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist &

Wiksell.

- Klimov, Georgij A. 1977, *Tipologija jazykov aktivnogo stroja*. Moskva: Izd. „Nauka“.
- Kurzová, Helena 1993, *From Indo-European to Latin. The evolution of a morphosyntactic type*. Amsterdam & Philadelphia: J. Benjamins (Current Issues in Linguistic Theory, 104).
- Ledgeway, Adam 2012, *From Latin to Romance. Morphosyntactic typology and change*. Oxford etc.: Oxford University Press (Oxford Studies in Diachronic and Historical Linguistics).
- Lehmann, Christian 1979, „Zur Typologie des Lateinischen.“ *Glotta* 57: 237-253.
- Lehmann, Christian 1985, „Ergative and active traits in Latin.“ Plank, Frans (ed.), *Relational typology*. Berlin etc.: Mouton (Trends in Linguistics, Studies and Monographs, 28); 243-255.
- Lehmann, Christian 1985, “Latin case relations in typological perspective”. Touratier, Christian (ed.), *Syntaxe et latin*. Actes du II<sup>me</sup> Congrès International de Linguistique Latine, Aix-en-Provence, 28-31 mars 1983. Aix-en-Provence: Université de Provence; 81-104.
- Lehmann, Christian 1989, “Latin subordination in typological perspective”. Calboli, Gualtiero (ed.), *Subordination and other topics in Latin. Proceedings of the Third Colloquium on Latin Linguistics, Bologna, 1-5 April 1985*. Amsterdam: J. Benjamins (SLCS, 17); 153-179.
- Lehmann, Christian 1992, “Word order change by grammaticalization”. Gerritsen, Marinel & Stein, Dieter (eds.), *Internal and external factors in syntactic change*. The Hague & Berlin: Mouton-de Gruyter; 395-416.
- Lehmann, Christian 1998, “Regiones espaciales en perspectiva tipológica”. García-Hernández, Benjamin (ed.), *Estudios de lingüística latina. Actas del IX Coloquio Internacional de Lingüística Latina, Universidad Autónoma de Madrid, 14-18 de abril de 1997*. Madrid: Ed. Clásicas; 455-466.
- Lehmann, Christian 2002, “Latin valency in typological perspective”. Kroon, Caroline & Pinkster, Harm & Rammelink, Wim & Risselada, Rodie (eds.), *Theory and description in Latin linguistics. Selected papers from the 11th International Colloquium on Latin Linguistics*. Amsterdam & Philadelphia: J. Benjamins; 161-181.
- Lehmann, Christian 2005, “La structure de la syllabe latine”. Touratier, Christian (ed.), *Essais de phonologie latine. Actes de l'atelier d'Aix-en-Provence 12-13 avril 2002 (avec le soutien financier du CNRS)*. Aix-en-Provence: Publications de l'Université de Provence (Langues et langage, 11); 157-206.
- Lehmann, Christian 2005, "Latin syllable structure in typological perspective." Calboli, Gualtiero (ed.), *Latina lingua! Proceedings of the Twelfth International Colloquium on Latin Linguistics (Bologna, 9-14 June 2003)*. 2 vols. Roma: Herder (Papers on Grammar, 9); 127-147.
- Lehmann, Christian 2008, “Roots, stems and word classes”. *Studies in Language* (Special Issue) 32:546-567.
- Lehmann, Christian 2016, “Latin causativization in typological perspective”. Poccetti, Paolo (ed.) 2016, *Latinitatis rationes. Proceedings of the 17th International Colloquium on Latin Linguistics, Rome, 2013*. Berlin: W. de Gruyter; 917-941.
- Matthews, Peter H. 1965, “The inflectional component of a word-and-paradigm grammar.” *Journal of Linguistics* 1: 139–171.
- Meillet, Antoine 1912, "L'évolution des formes grammaticales." *Scientia* 12(26):384-400.
- Meillet, Antoine 1915, „Le renouvellement des conjonctions.“ *Annuaire de l'Ecole Pratique des Hautes Etudes, section historique et philologique* 1915/6: 1-28.
- Meillet, Antoine 1934, *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette. 7e éd.
- Meillet, Antoine 1948. *Esquisse d'une histoire de la langue latine*. Paris: Hachette. 5. éd.
- Milewski, Tadeusz 1950, ‘La structure de la phrase dans les langues indigènes de l'Amérique du Nord’. *Lingua Posnaniensis* 2: 162-207.
- Nichols, Johanna 1986. ‘Head-marking and dependent-marking grammar’. *Language* 62, 56-119.
- Nichols, Johanna & Peterson, David A. & Barnes, Jonathan 2004, “Transitivizing and detransitivizing languages.” *Linguistic Typology* 8: 149-211.



- Ross, John R. 1973, "The penthouse principle and the order of constituents." Corum et al. (eds.), *You take the high node and I'll take the low node. Papers from the Comparative Syntax Festival*. Chicago: CLS; 397-422.
- Sgall, Petr 1971, "On the notion 'type of language'". *Travaux linguistiques de Prague* 4:75-87.
- Schlegel, August Wilhelm von 1818, *Observations sur la langue et la littérature provençales*. Paris: Librairie grecque-latine-allemande.
- Schlegel, Friedrich 1808, *Ueber die Sprache und Weisheit der Indier. Ein Beitrag zur Begründung der Alterthumskunde. Nebst metrischen Übersetzungen indischer Gedichte*. Heidelberg: Mohr & Zimmer (Repr.: Amsterdam: J. Benjamins (ASTHoLS, Ser.1, vol. 1), 1977).
- Skalička, Vladimír 1966, "Ein 'typologisches Konstrukt'". *Travaux linguistiques de Prague* 2:157-163.
- Spevak, Olga 2010, *Constituent order in Classical Latin prose*. Amsterdam & Philadelphia: J. Benjamins (Studies in Language Companion Series, 117).
- Szemerényi, Oswald 1977, *Studies in the kinship terminology of the Indo-European languages, with special reference to Indian, Iranian, Greek, and Latin*. Leiden ; Brill (Acta Iranica, 16).
- Van Valin, Robert D., Jr. 1977, *Aspects of Lakhota syntax. A study of Lakhota (Teton Dakota) syntax and its implications for universal grammar*. Berkeley: University of California PhD diss.
- Viti, Carlotta 2020, *Les dénomination des couleurs dans les anciennes langues indo-européennes entre universalisme et relativisme. Une étude de sémantique comparative*. Manuscrit. Paris: École Pratique des Hautes Études.