

Note on Latin Grammar

Jinyuan Wu

July 5, 2024

Chapter 1

Introduction

1.1 The language and the speaker

1.1.1 Latin as a classical language

Latin was the language of the Romans and the official language of both the Roman Republic and the Roman Empire, and hence the official language of the Catholic Church, which was *the* church for the Western Roman Empire. The international nature of the Roman Empire made Latin the international language around the Mediterranean Sea at that time – indeed, *Mare Nostrum* ‘our sea’ in Latin, and its importance in science, arts, law, religion, and literature lent it more than one thousand of years of life as a common literary language and a sacred language in western Europe after the collapse of the Western Roman Empire and the emergence of the Romance language family.

As recently as the nineteenth century, Latin was still fluently used by scholars and in the Catholic Mass. A decline in the popularity of Latin was observed after that. The rapid development of English (at first, also French and German and sometimes Russian) as the language of science largely replaced the status of Latin as a scholar language. After Vatican allowed vernacular languages being used in liturgies, Latin also largely lose its position in the daily use in the Catholic church.

This note is about **Classical Latin** – the Latin of classical Latin writers – and **Ecclesiastical Latin** – the kind of Latin of the Catholic church. That’s to say Old Latin, vulgar Latin (with prototypes of Romance articles), etc. are not discussed in detail in this note. Still, some historical knowledge is important for us to understand why Latin is the way it is.

1.1.2 Latin in ancient Mediterranean world

The historical and contemporary importance of Latin of course doesn’t endorse it as a inherently superior language. Indeed, we only have a handful of Latin texts before 600BC; as a comparison, there are about 150 pre-600BC Etruscan texts. Even in the period between 600BC and 100BC, during which we have around 3000 Latin texts, we have about 9000 Etruscan texts, which is three times as many as their Latin counterparts.

1.2 Previous studies

1.3 Texts

TODO: classical writers

Chapter 2

Phonology and the writing system

2.1 Phonemes and the alphabet

Although the phoneme inventory of a language often is not accurately reflected by its preferred system, since Latin is a classical language and no ancient Roman is alive today, its phonology has to be inferred from known texts.

The most accepted writing system of Latin developed into what we call **Latin letters** – or the **Roman alphabet** – today, which is the most widely used writing system in the world. **Old Italic scripts**, used by Early Old Latin inscriptions as well as neighbor languages, show a larger degree of variation, which clearly derived from Greek letters. The standard Latin alphabet derived from old Italic scripts. Note that ancient Romans only used the big letters; the small letters was invented during the era of Charlemagne.

The letter *J* was not used by ancient Romans, although we sometimes see *I* appearing at the start of a word and therefore possibly represents the semivowel /j/. The letters *U* and *W* are also not used. Similar to the case of *I*, the letter *V* is used to represent what appears to be the semivowel /w/ as well as the vowel /u/. The letter *K* is an archaic one and only appears before *A* in a small number of words (Oniga, 2014, chap. 2). The letters *Y* and *Z* are used to spell Greek words that include *TODO*: *Y* and the voiced dental affricate, respectively.

The Latin consonant inventory is therefore given by Oniga (2014, Table 3.1). Note that the letter *X* is a double consonant: it means /ks/ or /gs/.

Two semivowels – /j/ and /w/ – can be recognized, which appear as *i* or *u*.

The vowels are given by Oniga (2014, Table 3.2). Each vowel has a long variety and a short variety.

2.2 Prosody

2.3 Morphophonological rules

Some phonological rules in Latin are sensitive to morpheme boundaries. We can therefore assert that at least in a historical stage, Latin speakers had a clear sense of morphemes as real phonological objects, instead of mere theoretical models.

2.3.1 Vowel deletion

Short vowels *a*, *o* and *e* become zero before a morpheme boundary or another vowel (Oniga, 2014, § 8.3). This rule is exemplified by the absence of the thematic vowel in both declension (TODO: rosis) and conjugation (TODO).

2.3.2 Vowel shortening

A long vowel before another vowel or morpheme boundary is not deleted, but shortened. Again this is exemplified (§ 7.3.1, TODO: ref)

Also, in the final syllable of a phonological word, a long vowel before a consonant except *s* is also shortened. Counterexamples when the vowel is not in the final syllable exist, like *bāris* ‘a type of flat-bottomed freighter used on the Nile in Ancient Egypt’.

A long vowel is generally shortened before a sequence containing a liquid or nasal and a following stop consonant, like *nt*. This rule comes from an older Indo-European sound law: the Osthoff’s Law (Oniga, 2014, p. 55).

2.3.3 Vowel weakening

When a short syllable is in a medial, open syllable, and a morpheme boundary occurs immediately before, within or after the syllable, it becomes *i* (Oniga 2014, p. 55; TODO: ref).

2.3.4 Vowel lengthening

A vowel is always *lengthened* before *nf* and *ns* (Oniga, 2014, p. 55).

Chapter 3

Grammatical overview

We carry out this report and this chapter round by round: first we distinguish words and morphemes, then we briefly describe the clause structure and the noun phrase structure, and then according to how roots fit into these environment, we define parts of speech. The following chapters each focus on a different topic touched in this chapter.

3.1 Morphological typology

Wordhood Latin is well known for its rich morphology, which enables a rather free – but still not completely arbitrary (chap. 10) – constituent order. The grammatical categories seen in nominal and verbal morphology already reflect the most salient grammatical categories in NPs and clauses; thus the concept of morphological wordhood can be defined according to the boundaries of inflectional templates; and the traditional wordhood definition largely agrees with this definition. Thus, in this note I follow the traditional definition of Latin wordhood, which is easily done using phonological criteria, or, to be more accurate, orthographical criteria: what was documented by ancient Romans as a word is recognized as a word. For words without inflectional morphology,¹ wordhood can still be defined according to phonological wordhood. The two criteria of wordhood also largely agrees with the traditional word.

Also we can talk about various purely syntactic ways to define wordhood. For example a noun phrase minus large complements or adjuncts should be recognized as a noun, or a clause minus large complements or adjuncts should be recognized as verb. In this regard Latin has split verbs in periphrastic conjugation, and in the preverb construction, the prefixed verb, a morphologically single word, contains the verb proper and an adverbial root that syntactically should be a part of the locative complement of the verb. This line of wordhood definition helps us understand the structure of Latin noun phrases or clauses but goes against any conventional definition of wordhood.

The only subtlety in wordhood definition seems to be due to clitics, the most important one being the coordination *-que*: they are not a part of inflectional morphology of any stem but have to be attached after another word. For them, the orthographical standard simply dictates that they are to be recognized as a part of the words they follow. In conclusion, we can say the traditional notion of Latin words is (expectedly)

¹The traditional name is *particles*; see below.

linguistically sound and should be kept in use.

Stems and endings in nominal and verbal morphology A stem can be well defined for both verbal and nominal morphology in Latin, enabling a clear derivation-inflection distinction: prototypical derivational processes that are considered to be a part of the head noun/verb and not the surrounding NP/clause are morphologically realized strictly before prototypical inflection processes, forming the stem (§ 5.2.1, § 7.1.1) to which inflection processes are applied, with possible contextual allomorphs (§ 2.3). Latin inflection is always suffixal, while derivation is predominantly prefixal. Despite its richness, a large portion of instances of Latin derivation are historical, with meanings of derived forms having significantly shifted and no longer regularly inferrable.

A general tendency in the stem structure is modification affixes usually are prefixes, while affixes bearing grammatical information (like change of part of speech) are usually suffixes (§ 5.1, TODO: ref). Apart from affixation, Latin does use compounding, as in *aequilibrium* (*aequ-i-libr-ium*, ‘even-balance-SG.NOM, equilibrium’; Smith 2016, § 92), but compounding was already less productive – if not completely obsolete – in the Classical period.

Non-concatenative morphology Concatenative morphology (affixation and compounding) is prominent but isn’t the only morphological device. Reduplication is attested in formation of the perfect stem (TODO: ref); this however is largely historical and is no longer productive. Dropping of first-conjugation stem-final vowel (§ 7.3.1) may be analyzed as subtraction, although it can be seen as due to morphophonological rules. The imperfect *-ba-* is sometimes said to be an infix (as well as its counterparts like *-bi-*), though it fits in a concatenative picture of verbal morphology.

3.2 Clausal syntax

3.2.1 Overview

Clauses in utterances We start our systematic discussion on Latin syntax from the verbal system, i.e. the clause. A clause may appear as a single utterance called a **sentence**² in this case, either declarative or interrogative, or as a complement of a verb (complement clause construction), or as an adverbial of another clause (i.e. modifier in the verbal system), and as a relative clause (i.e. modifier in the nominal system).

Box 3.1: Distribution of clauses

List where clauses can be found in other structures.

Clause types Clauses types in Latin can be roughly divided into finite and non-finite ones, which can be told from verbal morphology.

In sentential clauses, the category of speech force is marked by e.g. fronting of the interrogative pronoun. No other grammatical markers are dedicated to clause type: there are, for example, no sentence-final particles that can only be observed in *sentential* clauses and not embedded finite clauses.

Non-finite constructions include INFINITIVES and PARTICIPLES, and the GERUND. These constructions can be distinguished by morphology only: in English we have

²Some authors (including some Latinists) use the term *sentence* as a synonym of *utterance*; it is not followed in this note.

infinitive clauses, but strictly speaking, there is no such thing as “infinitive verb”, as the head verb of an infinitive clause has exactly the same form of a non-third person singular present tense verb; in Latin however, the head verb of a infinitive clause in Latin indeed has a separate position in the paradigm.

Box 3.2: Implicit marking of speech force categories

Are there any e.g. obligatory fronting or contraction that may be related to the speech force of a clause?

We also need to investigate the licensing environment and the inside *wh*-forms if the clause is a relative clause.

Box 3.3: Clause-level subordination

- Coordination
- Conditional
- Concessive

Internal makeup of a clause A clause consists of a nucleus clause with possible information structure-related order alternations and/or clause-level subordination constructions. The flexible constituent orders of Latin lead someone to characterize it as a free-order language, but closer examination reveals that the order alternations can be explained by information packaging and/or prosody (chap. 10).

At the surface level, the nucleus clause contains a (omissible) subject (§ 3.2.2), a main verb, a possible auxiliary verb, and possible adjuncts (including peripheral arguments and Tense, aspect, mood (TAM)-adverbials) and internal arguments.³ The main verb shows regular agreement with the subject. The inflectional ending of the main verb (chap. 7), together with TAM adverbials and an optional auxiliary verb, marks TAM grammatical categories (§ 3.2.3). The distinction between what are known as peripheral arguments and core arguments is discussed in § 3.2.4; we expect most so-called peripheral arguments to in some sense have scopes over core arguments.

The core arguments – and to some extents the peripheral arguments – are ultimately determined by the **argument structure** of the main verb. Arguments have “deep” and “surface” positions. The deep positions are defined by properties like how reflexives work, mapping of semantics to argument slots,⁴ the semantics difference between S=A and S=P valency alternations (see below), etc. The mapping from deep

³In some works on English syntax, e.g. Huddleston and Pullum (2002), the term *complement* means a dependent position which, according to various standards, is more closely related to the lexical head; in the case of clauses, a complement is just an argument.

In traditional Latin grammar, however, *complement* means the copular complement. This note follows the terminology used in most descriptive grammars, so use the term *copular complement* to refer to the *predicative complement* in Huddleston and Pullum (2002).

A further confusion, as is seen in § 8.2, is due to the term *complement-taking verb*, i.e. a verb that take a complement clause or something semantically equivalent to a complement clause as one of its arguments.

⁴We should however note that two arguments being denoted by the same semantic role label (e.g. *stimulus*) do not mean they are truly semantically equivalent, and clearly they are necessarily syntactically comparable. In Dixon (2005, § 4.2), for example, what bears the label *target* is almost prototypically patient-like in verb frame I and is the goal of a directional construction in verb frame II, and of course there are syntactic as well as semantic differences between the two frames.

positions to surface positions (like *subject*) is influenced by the **alignment typology** of the language and also **voice** constructions.

Details of deep argument structures and their reflections in the surface argument structure are specified in possible **verb frames** in a language. The argument structure is often related to the *event structure* of the verb, and therefore the verb frame may also influence the TAM marking (chap. 8).

Regarding alignment typology, Latin is a nominative-accusative language with clear S=A neutralization and therefore a well-defined **subject** (§ 3.2.2).

Latin has a passive voice, which is marked on verbal morphology (chap. 7) and seems to work solely on the concept of **object**; no other highly regular voice constructions of such kind are observed (§ 3.2.6). It's still possible to have verb frame alternations that highly depends on the properties of the verb stem.

3.2.2 Subjecthood

Latin is a nominative-accusative language. This means we have a well-defined clause-level pivot which, in active constructions, can be identified with the structurally most agent-like argument⁵ in the argument structure. This pivot is called the **subject**.

Existence of a clause-level pivot can be clearly identified from the following facts.

- *Extraction in coordination.* When two clauses with a shared subject are coordinated, the subject usually appears at the beginning of the coordination construction (Pinkster, 2021, § 19.5). This demonstrates that the subject is somehow structurally promoted to a higher position in this kind of coordination.⁶
- *Case marking.* Subjects are always nominative for finite clauses, whenever case marking is available (i.e. whenever the subject is an NP or a gerund). On the other hand, nonfinite clauses are either subjectless or have an accusative clausal dependent functioning as the subject (§ 12.1.1.1): we may argue that by being nonfinite, they have deficient nucleus clause structures and cannot afford a canonical subject.
- *Agreement.* The number and person features on the subject leave marking on the verb complex. Latin does not have verbal agreement with arguments other than the agreement with the subject.
- *Constituent order.* In clauses with neutral information structure, the subject usually appears at the beginning of the clause. Violations to this generalization are countless but they can all be analyzed as focalization or topicalization of other constituents and postponing of the subject (§ 10.4).

Latin subjects also have a quirky features. A subject is an NP or a complement clause (§ 12.1), usually an infinitive but never a gerund (§ 7.4.2.1). This constraint isn't seen in any other clausal complement types.

In Latin we can also recognize an argument structure-level pivot, or in other words, an “agent-like” argument in each clause (§ 3.2.5). The fact that in active clauses,

⁵What it takes to be agent-like is discussed below.

⁶Omission of the object in coordination, on the other hand, shows more diversity in constituent orders (Pinkster, 2021, § 19.6), suggesting that the exact structural analysis may also be more complicated.

an agent-like argument is always the subject means Latin is indeed a nominative-accusative language.

Box 3.4: What else to be done for the subject

- Passivization
- Psych-verb (Box 8.2)
- Describe phenomena in the last paragraph.

Box 3.5: Clause template

When all information packaging devices are removed, we have SOV order, and the position of O varies depending on its information structure.

Box 3.6: Core arguments: what are they?

- Structural affinity to the verb: anaphora in space of verb phrase (VP)
- Obligatoriness: finiteness

3.2.3 TAM, and so on

Tense, aspect and modality is most evidently marked by verbal inflection in Latin (chap. 7). Other TAM categories however are regularly marked by adverbials.

Box 3.7: Latin nonfinite constructions

Nonfinite constructions lack *morphological TAM marking*, but are TAME adverbials possible for them? Is nonfiniteness a morphological or syntactic property?

3.2.4 Peripheral arguments

The term *peripheral arguments* is used in Dixon (2009) to refer to what others call (non-TAM) adjuncts and specify information like the spatial and temporal location of the event, the instrument, the manner, etc. Parameters traditionally taken as core-peripheral distinctions involve structural closeness of the clausal dependent to the main verb, obligatoriness of the argument, and the internal makeup of the argument.

3.2.5 Core arguments

Prototypical transitive construction In Latin

The sole argument of intransitive verbs The sole argument of intransitive verbs – referred to as S in the typological literature – may have behaviors comparable to A or P in prototypical transitive constructions: the intransitive verbs are correspondingly known as *unergative* and *unaccusative* verbs, although the names are misleading as we are discussing about the deep argument structure and not alignment yet. We can also observe S=A and S=P labile valency alternation.

In Latin, it seems that unaccusative verbs are all deponent (Oniga, 2014, pp. 308-309)

Benefactive constructions Verbs meaning ‘giving’ in Latin

Directional constructions Some Latin verb frames seem to have a directional small clause (§ 3.2.9) hidden in them. In the surface form, this results in a preverb (usually the *path* in the directional construction) appearing before the verb root, and the *location* appears as an argument of the verb. The case marking of the location argument may reflect what type of location it is (and hence is dictated by the preverb, as if the preverb is used as a standalone preposition and the location is governed by it), or in may appear in accusative. Passivization of accusative location arguments is sometimes possible (Pinkster, 2015, § 4.22). These phenomena seem to be comparable to English preposition verbs and verb-particle constructions, where the preposition may be somehow defunct and the location argument it governs syntactically is the object of the verb.

It is also possible for the location argument to appear in dative (Pinkster, 2015, § 4.25). In this case

Oblique arguments and peripheral arguments Beside the subject, various types of objects, and copular complements, there are more clause dependents corresponding to less frequently seen semantic roles like purpose, direction, location, etc.; similar semantics can be realized by serial verb constructions, which are however absent in Latin. These more or less peripheral dependents may be licensed or even required by the verb and in this case are often still recognized as core arguments (**oblique argument**, corresponding to the E argument in Dixon (2010, p. 116)), or they may be modifying the whole clause and therefore are usually optional (**peripheral argument**, or “**adjuncts**”). Besides clauses and NPs (with or without prepositions), their categories also include adverb phrase (AdvP)s, a majority of the latter historically originating from fossilized case forms.

3.2.6 Voice

3.2.7 Information packaging

Due to the extreme complexity of Latin clausal constituent order, we delay our detailed discussions to chap. 10.

3.2.8 Clause combining

In Latin there is no serial verb constructions. Subordination strategies can be neatly summarized into complement clauses, relative clauses and adverbial clauses.

3.2.9 “Small clauses”

Some constructions may be analyzed as “small clauses”, which are smaller than usual clauses and lack real TAM marking; they are therefore not “real” clauses, but still appear frequently in Latin texts. (1) is a formula used frequently in Catholic Church and is an example of a small clause as a single and complete utterance. The fact that *gratias* is not in the nominative or the vocative case means (1) is neither a finite construction nor a single-NP utterance.

- (1) De-o grati-as
God-SG.DAT thank-PL.ACC
'(May) thanks be to God.'

Since English – and a lot of other languages – bans small clauses appearing as complete utterances, translation of these constructions may map them into full clauses in the target language.

Box 3.8: Why small clauses

Alternatively, small clause constructions can be analyzed as a full clause (for example, a copular clause) with the verb deleted. The main problem of this analysis is if this is true, we need to explain why prototypical transitive clauses never see their verbs omitted. It's therefore better to say that a copular clause is a small clause *plus* the copular, and that a giving or receiving ditransitive verb takes a small clause as its internal complement, which doesn't involve a questionable verb deleting process and also captures the intuition that a small clause is a full clause minus the verb.

3.3 Nouns, pronouns and the nominal system

3.3.1 Overview

Internal makeup of noun phrases Similar to the case in clause structure (chap. 10), the internal constituent order of NPs are strongly influenced by information structure. The default constituent order is attributive-before-head noun (Allen and Greenough, 1903, p. 396).

3.3.2 Case endings and prepositions

3.4 Adjectives and adverbs

3.5 Coordination and subordination

3.6 Parts of speech division

3.6.1 Theoretical caveats

Strictly speaking, terms like *verb* or *noun* above are about *syntactic positions* of roots; they however do not tell us correlations between these syntactic positions and realizational aspects of these roots. It is, for example, in principle possible to have two classes of roots that can both head adjectival constructions but behave differently in nominalization have different inflectional morphology. In a word, after sketching grammatical constructions, we still need to set up labels in the *lexicon* – **part of speech labels** – to guide how roots fits into the grammatical constructions. In other words, we need to know how the roots are *categorized* in the lexicon.

Another thing to keep in mind is that the lexicon contains construction with different sizes. For example, the meaning, argument selectional rules and phonological form of a verb root is to be recorded in the lexicon, but larger constructions, like the verb

with a directional prefix, or even a whole verb phrase (e.g. *kick the bucket* in English) may have developed their own meanings, and they are to be recorded in the lexicon as well. A syntactic unit with a fixed meaning recorded in the lexicon is *lexicalized*. This sometimes may lead to unit to be *fossilized*: certain syntactic operations may hinder its non-compositional interpretation and hence be unacceptable. This kind of fossilization may eventually lead to *syntactic fossilization*, in which the internal structure of the syntactic unit collapses, altering its behaviors (a typical example being some English prepositional verbs allowing passivization), and *grammaticalization*, where the syntactic unit becomes a part of the grammar (e.g. complement-taking verbs become auxiliary verbs, which appear in parts of the conjugational paradigm and have to appear in a given order).

Grammaticalized syntactic objects strictly speaking should not be included as a part of the lexicon: they are a part of the grammar by definition. But frequently they are for convenience. Putting together the degree of grammaticalization, the openness of a part of speech, and the syntactic position of the units, we have the following classification of parts of speech:

1. prototypically lexical (open to new members, large in size, being able to head noun phrase, etc.), or
2. closed but still lexical (a class with no new member allowed but still large in size, being able to head noun phrases, clauses, etc.),
3. a group of pronoun-like things, which appear in syntactic environments similar to those of nouns, verbs, etc., or
4. a finite list of members that are pure markers of grammatical categories.

The first two types are usually considered lexical. The first three types are usually considered to have “real” part of speech labels: a member of the fourth type is always attached to a construction headed by something in the first three types of parts of speech and doesn’t really need its own part of speech label and is prototypically functional, not lexical.

The distinctions between the four type of course are not categorical. The most important issue is that some so-called closed classes still admit new members albeit slowly. Also, grammaticalization from lexical words to functional words is not synchronic among speakers. The distinction between the third and the fourth types may be problematic for, say, demonstratives, which can be a determinative marker as well as a whole NP on its own. Finally, as is mentioned above, prototypical lexical words can be seen as roots plus “categorizers,” and in languages with lots of etymologically related nouns and verbs, we can even argue that the noun and verb endings are of the third type. This however is merely a terminological issue.

3.6.2 Lexical categories with inflectional paradigms

Some Latin word classes can be defined easily via morphology and these classes prove to have uniform morphosyntactic behaviors. Words with large inflection paradigms can be divided into two large classes: those with similar morphology of prototypical nouns (i.e. *DECLENSION*) are *NOMINALS*, while words with similar morphology of prototypical verbs (i.e. *CONJUGATION*) form a uniform class rightfully called *VERBS*.

Nominals include NOUNS and ADJECTIVES, the distinction between the two can also be defined morphologically.

Latin nouns, verbs, and adjectives are all prototypical lexical categories; there seems to be no closed lexical categories (like Japanese verbal adjectives) in Latin.

Nouns Nouns prototypically head noun phrases. They are declined for case and number (§ 5.2), and the features also spread to other nominals in the NP by agreement.

According to their meanings and ability to license NP dependents, descriptive parameters of Latin nouns include inflection class, valency, compatibility with single or plural number, and TODO

Verbs Classification parameters of the Latin verb include the conjugation class (§ 7.3), the syntactic and semantic argument structure (§ 9.4), the event structure (TODO: Compatibility with TAM),

3.6.3 Latin “particles”

Parts of speech without much inflectional morphology include PREPOSITIONS, ADVERBS, INTERJECTIONS, and CONJUNCTIONS, traditional known as PARTICLES, although in modern usage the term *particle* usually refer to grammatical items, and therefore at least some adverbs should be excluded from that class; also, some adverbs have inflection with respect to *grade* (comparative and superlative). It’s often said that adverb class and the preposition class have a large overlap: often a preposition has an intransitive counterpart, which is similar to a prototypical adverb – but see the discussion below about the lexical/functional distinction in prepositions. Conjunctions may be seen as “prepositions for clauses”. The functions and etymologies of particles are highly diverse.

The preposition class usually shows great diversity both within a language and among different languages, and what are known as prepositions may be lexical words or function words (e.g. Garzonio et al. 2021). In morphologically more analytic languages, we may find preposition that has strong interaction with the rest of the grammar and hence may be seen as a case particle (e.g. English *of* used to introduce the object in nominalization and *by* used to introduce the deep A argument in passivization). Since Latin has a morphological case system, such prepositions are expectedly largely missing. There is one use of *ab* that may fall in this category: personal agents in passive clauses are always preceded by *ab*.

Some prepositions may be treated as realizations of inherent cases and are therefore still functional particles instead of lexical words. A very typical sign for such type of non-lexical prepositions is stacking of case markers and/or prepositions with similar roles in a fixed order and with a finite number of possibilities, which is the case in Finnish or, to a lesser degree, in English (Svenonius 2010; consider English *from behind the wall*). In Latin there is no clear example of this type. Semantically, *ex* encodes a path going away from a given region, and *cum* resembles a comitative marker; the fact that they dictate the case of the noun phrases following them therefore resembles the fact that English perfect auxiliary *has* selects a past participle or the Japanese verbal conjugation suffixes add CONTINUATIVE or GERUND endings to the verbal complex before them, which can be explained by diachronic reasons; the possibility of the past being the present however cannot be ruled out: these prepositions may also be instances of the lexical prepositions below. The distinction between lexical and functional directional prepositions in this case is hard; making this distinction needs native

speakers' judgments on subtle cases and/or neurolinguistic and psycholinguistic data (in acquisition, agrammatism, etc.), which is of course not available to Latinists; besides, we do note that grammaticalization is not synchronic among all speakers and even if we do have access to the relevant data, what we find will likely be author-specific and do not lead to any generalization to Latin as a whole.

Some Latin prepositions are likely as lexical as other lexical words. The morpheme *circum* means a (static) location in space or a dynamic path depending on whether a path is implied by the verb; the behaviors are very unusual if the word is understood as a purely grammatical particle. In Mare (2018) *circum* is analyzed as a root surrounded by noun phrases expressing the GROUND and the FIGURE in a spatial relation, resulting in a locational expression. It however is also possible that a preposition is a modifier of an existing prepositional construction (Botwinik-Rotem and Terzi 2008; c.f. *up* in English *up in the tree*).

Many prepositions can also be used without any complements and may thus classified as an adverb; they can also appear as preverbs.

3.6.4 Latin pro-forms

For categories in the third type, Latin has pronouns and pro-adverbs; depending whether we treat prepositions as heads: though prepositions are often said to be markers of a periphrastic case system, the semantics carried by certain Latin prepositions are too complicated for a case system, and these prepositions can be understood as adverbs selecting a complement. TODO: adverb modifying PP This is also the case of adverbs: some adverbs seem to be periphrastic markers of TAM categories and therefore may be considered as a part of the grammar, while others seem to carry “real” meanings.

3.6.5 Diagram of Latin word classes

Fig. 3.1 is a visualization of the classification of Latin word classes.

Parts of speech absent in Latin Articles (English *a* or *the*), despite prevalent in other Indo-European languages, are missing in Latin. This, together with the fact that Classical Sanskrit and Old Persian didn't have articles and the Slavic languages still don't, is a strong indicator that proto-Indo-European (PIE) didn't have articles. Note that the fact that Latin lacks articles doesn't mean the determiner syntactic function doesn't exist: there are evidences suggesting certain aspects of the behavior of Latin NPs are just like English (Giusti and Iovino, 2014).

3.7 Advice for reading Latin

The morphological richness (and the scrambled constituent order) makes Latin hard to read especially for people whose first languages are, say, English or Mandarin. Whenever unsure about a sentence, do the follows:

1. Skim over the words and label the stems that can be easily recognized.
2. Skim over and circle uncontroversial grammatical items, like inflectional endings and prepositions. It's OK to be unable to interpret them immediately (and we need the steps below).

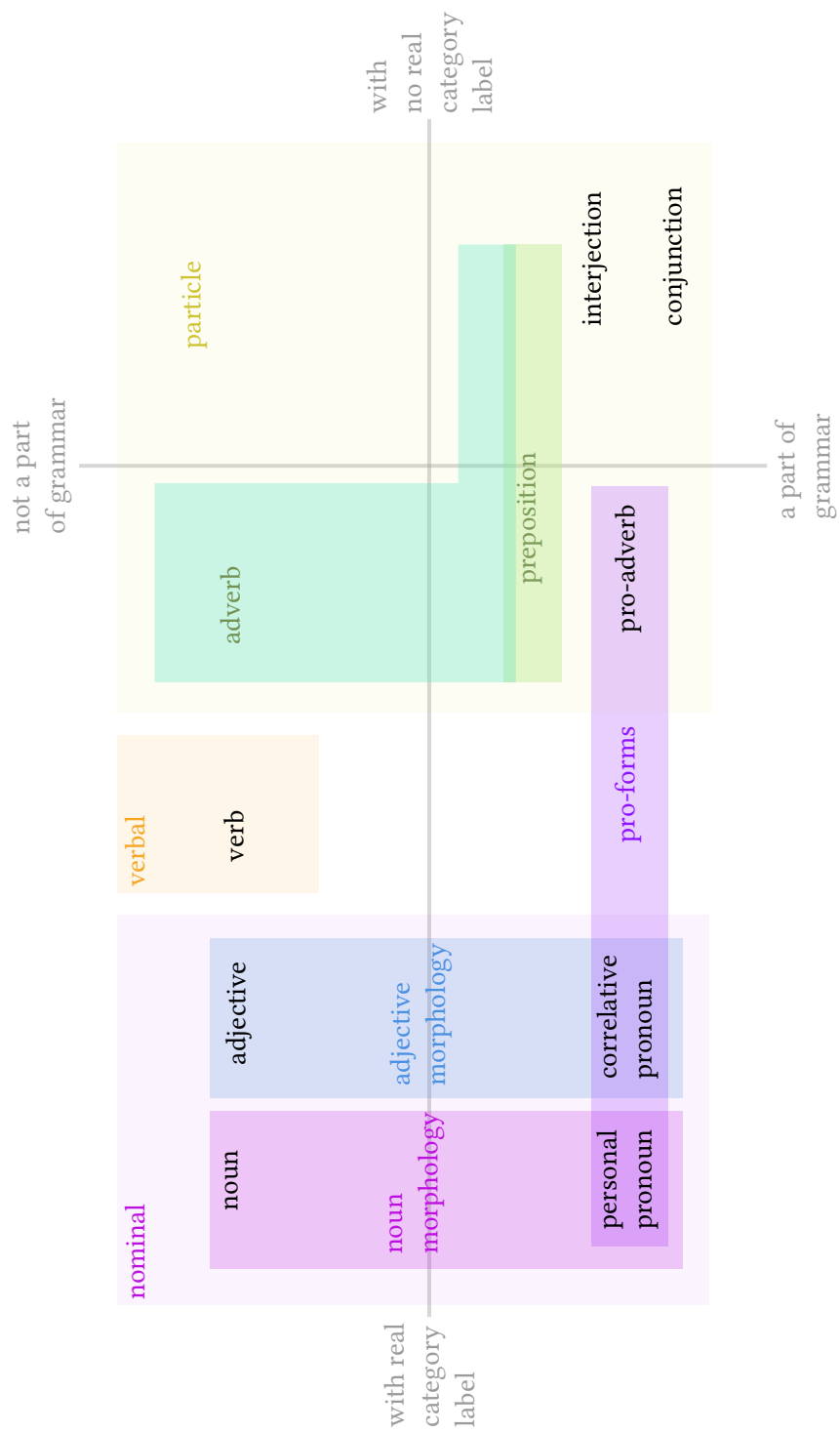


Figure 3.1: Latin word classes

3. According to these grammatical items, segment the sentence into large parts: noun phrases, relative or complement or adverbial clauses, verbal system, etc.
4. Choose a grammatical item and tentatively give a list of possible features it carries. For example, seeing *-v-* in a verb usually means it's based on the perfect stem (§ 7.2.3); *-um* may be second declension accusative, but there are other possibilities (Table 5.1).
5. Use constraints like “the preposition *in* licenses the accusative case or the ablative case” to narrow the possibilities identified above.
6. Draw unfinished dependency arrows: for a verb, draw arrows pointing to the subject and/or the object; for a nominative adjective, draw an arrow pointing to the modified head noun. But note that it's possible that the subject is dropped, or there is no head noun (compare English *the poor*). Then try to pair the arrows.

Repeat the above procedure and finally the sentence can be understood. This procedure is demonstrated in § 15.2.1.

Chapter 4

Pro-forms

4.1 Overview

Pro-forms can be divided into pronouns and correlatives.

4.2 Personal pronouns

Latin pronouns are complete NP themselves: no attributives should be attached to them. Pronouns are declined for case, gender and number, and they also can be governed by prepositions.

It should be noted that what counts as third person pronouns is not uncontroversial among reference books regarding Latin; some place the reflexive *sui* series into the third person column in the table of personal pronouns, citing the fact that the *is*, *ea*, *id* series mark the gender category, while the first and second person pronouns do not.

4.3 Reflexive pronouns

It can be seen that *se* lacks nominative forms. In some circumstances, the reflexive *se* looks like the subject; but these cases are more appropriately analyzed as reflexive usages of transitive verbs (1).

- (1) ...necessarium est primo investigare de ipsa sacra doctrina, qualis sit, et ad
quae [se extendat]

‘It is necessary to first investigate into this sacred doctrine, of what kind it would be, and what it would extend itself to (i.e. its nature and extent).’ (*Summa*, I q. 1 pr.)

4.4 The “basic” correlatives

Latin correlatives can be classified in Table 4.1; the rows correspond to their immediate roles in the clause (“basic” i.e. NP heads, place, source, time, etc. that are otherwise expressed by prepositional phrases), while the columns correspond to their meaning

and/or relation with a more precedent in a higher position, like whether the word is a demonstrative or maybe a relative pronoun.

The form of correlatives in the “basic” can be summarized as the follows:

- *Indefinite*. They may be homonyms of *qui* or *quis*, or they may be formed by adding *ali-* or *-dam* to *qui* or *quis* (with possible internal alternations), or maybe they are formed in the format of *ali-quis-X-quam*.
- *Indefinite relative*. Formed by reduplication of *quis* forms.
- *Other*. Always take a form similar to *alter* or *alius*.

Here is a guide for quickly determining the meaning of a “basic” pro-form:

- What starts with *ill-* is a distal demonstrative.
- What starts with *ist-* is a medial demonstrative.
- What starts with *h-* is a proximal demonstrative.
- What starts with *s-* is a reflexive pronoun, or maybe a reflexive possessive determiner like *suus*.
- What starts with *qu-* is an interrogative pronoun or a relative pronoun or an indefinite pronoun or an indefinite relative.
- What starts with *null-* is a negative pronoun.
- What starts with *omn-* is a collective pronoun.
- What starts with *ips-* is an identity pronoun (‘this exactly ...’).
- What starts with *alter-* is an “other” pronoun.
- What starts with *ali-* may be an “other” pronoun, or an indefinite pronoun.
- What ends with *-dem* and starts with a third person pronoun is an identity pronoun; sound changes are possible. For example *īdem* or *eandem*.
- What ends with *-dam* and starts with *q-* is an indefinite pronoun.

Some forms can be easily confused with correlatives. They include:

- *quod* used to introduce a finite complement clause.
- *quia* as a conjunction of a reason clause.

4.4.1 Demonstratives

Latin has proximate, medial, and distal demonstratives: the first refers to something close to the speaker, the last refers to something far away, and the second refers to something in between, probably near the listener.

4.4.2 Interrogative and relative pronouns

Table 4.1: Classification of Latin correlatives

	“question”			demonstrative		
	interrogative	relative	indefinite	proximal	medial	distal
nominal head						identity
						indefinite
						collective
						negative

Chapter 5

Noun and noun phrase

Noun phrases are constructed surrounding a head noun, serving mostly as arguments in the clause. Some minor types of NPs contain only one word – a pronoun – and *can't* accept any modification or complementation. In this chapter we consider the more prototypical noun phrase.

5.1 The structure of noun

Latin nouns are declined for case and number, which is agreed upon by other nominal words in the NP. The structure of the Latin noun is just the stem plus inflectional ending with the case and number categories fused into one suffix (§ 5.2). There are five declension classes in Latin.

Inside the stem we find a list of

5.2 Declension of regular nouns

5.2.1 The paradigm

There exists a discrepancy in recognizing the stem of a noun. One approach is to find the common part of all case forms of one nominal lexeme; thus *rosam* 'rose-SG.ACC' is analyzed as *ros-am*, with *ros-* being the stem. Allen and Greenough (1903, p. 17) documents the paradigms of all the five declension classes in this approach. The full list of attested noun endings is Table 5.1. The list is still not the full picture of Latin nominal inflection: stem alternation is seen in the third declension (§ 5.2.2), and this alternation is frequently directly exposed at the end of a nominative and accusative noun.

Table 5.1: Declension endings; Roman numerals are declension classes

ending	declension
-a	I, SG.NOM, SG.VOC; IIN, PL.NOM, PL.ACC, PL.VOC; IIIN, PL.NOM, PL.ACC, PL.VOC
-ā	I, SG.ABS
-ae	I, SG.GEN, SG.DAT, PL.NOM, PL.VOC;
-am	I, SG.ACC
-ārum	I, PL.GEN
-ās	I, PL.ACC
-e	IIM, SG.VOC; IIIFMN, SG.ABS
-ē	V, SG.ABS
-ei/-ēi	V, SG.GEN, SG.DAT
-em	IIIFM, SG.ACC; V, SG.ACC
-ēbus	V, PL.DAT, PL.ABS
-ērum	V, PL.GEN
-ēs	IIIFM, PL.NOM, PL.ACC, PL.VOC; V, SG.NOM, SG.VOC, PL.NOM, PL.ACC, PL.VOC
-ī	IIM, SG.GEN, SG.VOC, PL.NOM, PL.VOC; IIN, SG.GEN; IIIFMN, SG.DAT, SG.ABL
-ibus	IIIFMN, PL.DAT, PL.ABS; IVFMN, PL.DAT, PL.ABS
-is	IIIFMN, SG.GEN
-īs	I, PL.DAT, PL.ABS; IIMN, PL.DAT, PL.ABS
-ō	IIMN, SG.DAT, SG.ABS
-ōs	IIM, PL.ACC
-ōrum	IIMN, PL.GEN
-r	IIM, SG.NOM, SG.VOC
-ū	IVFM, SG.ABS; IVN, SG.NOM, SG.DAT, SG.ACC, SG.ABS, SG.VOC
-ua	IVN, PL.NOM, PL.ACC, PL.VOC
-uī	IVFM, SG.DAT
-um	IIMN, SG.ACC; IIN, SG.NOM, SG.VOC; IIIFMN, PL.GEN; IVFM, SG.ACC
-us	IIM, SG.NOM; IVFM, SG.NOM, SG.VOC
-ūs	IVFM, SG.GEN, PL.NOM, PL.ACC, PL.VOC; IVN, SG.GEN
-uum	IVFMN, PL.GEN

There are several possible cases where something seeming to be a noun case ending turns out to be something else. To name a view:

- The ending *-as* is not always the first declension plural accusative ending; *-a-* could be a part of the stem of a third declension noun. Consider *vērītās*.
- The ending sequence *-io* can be found in Table 5.1 and we may hurry to the conclusion that it's the third declension abstract noun ending *-io* in the nominative or accusative case. Not necessarily – it can also be *-ium* in the dative or ablative case (when the macron symbol for long vowels are not used).
- A word ending with *-us* could be an adverb: *-tus* is an adverb suffix.

Another approach – informed by historical comparison with other Indo-European languages – make use of the thematic vowel in declension endings that can also be found in Table 5.1 (Oniga, 2014, pp. 45, 63). Thus, the first conjugation singular accusative ending *-am* is *-a-m*, with *-a-* being the thematic vowel; we may also attach

the thematic vowel back to the stem in the narrow sense *ros-* and redefine the complex *ros-a-* as the stem of *rosam*. The table is shown

5.2.2 The third declension

The third declension is a big tent containing several subclasses. Nominative singular endings attested in the third declension include *-s*, *-t*, *-x* (i.e. *-cs*) (Allen and Greenough, 1903, § 53).

5.3 The structure of the noun phrase

Although Latin lacks the article

5.3.1 Attributives

This section only discusses adjective or numeral attributives in detail. For in-depth discussion of relative clauses, see chap. 13.

5.3.2 The possessive construction

5.3.3 Numerals in the noun phrase

5.4 Case and preposition

After discussing the internal structure of NPs, we investigate the interaction of NPs with their syntactic functions. This is done by two systems: the case system and the preposition system.

5.4.1 Distribution of the cases

The roles of the five cases are not symmetric. Being nominative simply means being the subject in a finite clause or something agreeing to it and nothing else: the subject may be in a passive clause and is not agentive at all. The nominative case and the accusative case received by the direct object are *structural* cases: they are purely decided by the syntactic environment and don't have much semantic significance.

On the other hand, the rest cases are *inherent* cases: they have direct semantic interpretations – “source” or “target” or ... – themselves, and once an inherent case is assigned to an NP, the latter is “sealed” just like a prepositional phrase: the change of the outside syntactic environment doesn't change anything inside. Sometimes inherent cases are realized as prepositions.

5.4.1.1 The nominative

5.4.1.2 The accusative

The accusative case has both structural and inherent uses (Oniga, 2014, p. 238):

5.4.1.3 The genitive

Although in most teaching materials, the genitive case is considered to mark the possessor (and is thus an inherent case), a complete overview of its distribution reveals that it's actually a structural case representing *any* type of nominal dependency and not just possession and also sometimes assigned by certain types of verbs; some of the nominal dependencies can be metaphorically understood as possession, but others can't (Oniga, 2014, p. 244): it's hard to explain genitive of quality in terms of possession. This is comparable to English *of* in *a man of great talent* and *it's so kind of you*, which seems to be a purely structural phenomenon and is not based on metaphor; and indeed what is considered as prototypical possession also has strong cross-linguistic variation (Dixon, 2010, pp. 262-263).

5.4.1.4 The dative

The dative case is an inherent case assigned to the benefactive, the experiencer, and the purpose (Oniga, 2014, p. 251); thus it's the case assigned to the indirect object (§ 9.1.2) as well as many adjuncts TODO

5.4.1.5 The ablative

The ablative case is an inherent case assigned to the source and the instrument; so its distribution is similar to the dative case: it appears in source and instrument indirect objects (TODO) as well as various adjuncts. A source argument can be the position from which something moves (**ablative of source**) or the source in a separation event (**ablative of separation**: 'remove', 'deprive'), or the place where something comes into being (**ablative of material**, 'birth', 'origin'), or the cause of something ("the source of the event", **ablative of cause**); the agent in the passive voice possibly comes from one of the figurative use of the ablative as well.

5.4.2 Prepositions

Chapter 6

Adjectives and adverbs

The adjective class and the adverb class are linked together by several factors: the adjective phrase and the adverb phrase are both prototypical modifiers, often with parallel structures; they both have the category of degree; adverbs can be formed regularly from adjectives.

6.1 Declension of regular adjectives

Peripheral arguments may also be regarded as adverbials. This chapter, however, is mainly about mean, TODO

6.2 Arguments of adjectives

6.2.1 Dative

- (1) ... consubstantialem Patri ...
consubstantial-SG.ACC Father-SG.DAT
'consubstantial with Father' (Nicene Creed)

6.3 Comparative construction

tam quam

TODO: structure of some prepositions

Chapter 7

Verb morphology

7.1 The structure of the paradigm

7.1.1 The verb template of the finite paradigm

For a heavily inflected language like Latin it's an appealing idea to start the discussion on clause structure with a surface-orientated analysis of verb structure. The morphological template is shown in Fig. 7.1; some inflected Latin verbs and their parts are shown in Table 7.1. Traditionally, the verb is divided into the **stem** and the **ending**. Derivation in Latin is predominantly preverbal, and hence the conjugation is mostly about the final lexical morpheme in the verb stem, which is represented as the root in Fig. 7.1. There may be a perfect suffix after the root. Components of the verb ending include the **tense and mood suffix** (also known as the **tense suffix**), and the person, number and voice marker, which is called the **personal ending** here, following the terminology in Allen and Greenough (1903, § 165).

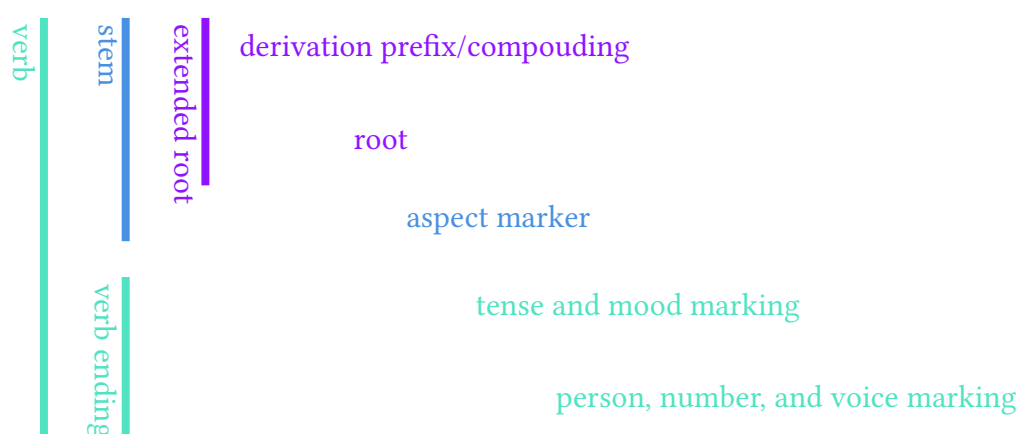


Figure 7.1: The template of Latin verbs. Indentation means linear order and not necessarily constituency structure.

The core stem assumes semi-regular alternations. The perfect marker is often but not always *-v-* (§ 7.2.3). Although the aspect marker is a part of the TAM marking (§ 7.1.3), it somehow is much more closely attached to the root in the morphophonological realization of the verbal system: how it is realized is not completely predictable, and therefore the two finite stem forms – the present stem and the perfect stem – are

both required for complete characterization of the paradigm of a regular verb. To fully decide the paradigm we still need a further stem, the supine stem (§ 7.2.1).

Contextual allomorphs also exist for other morphemes in Fig. 7.1 (Embick and Halle, 2005, p. 11), as can be observed in Table 7.1. The most salient change is that an alternating vowel may appear between the root and the aspect suffix (if any), which is also known as the **thematic vowel**. It is the residue of the PIE ablaut and labels the conjugation class (§ 7.2.4), and is subject to morphophonological contextual alternation (§ 7.3.1). The tense affix is determined by “tense” in the sense of traditional Latin grammar, i.e. both tense and aspect (Table 7.3), and has dependence on the conjugation class (TODO: ref). The PERFECT tense also has its own personal endings (Table 7.4).

Box 7.1: More than one thematic vowel?

Phenomena similar to the thematic vowel – vowel alternation with no morphosyntactic significance between two morphemes – can be observed between the tense suffix and the personal ending (as in *amābuntur*; § 7.3.3) as well. We may go further and say that the *-imus* ending seen in first person active indicative perfect, as opposed to *-mus* observed in other cases, contains a thematic vowel *i* after the perfect marker *-v-*. Some generalize the concept of thematic vowel and stipulate a thematic vowel position after *every* morpheme Embick and Halle (2003). This is also motivated by evidences from some descendants of Latin (Oltra Massuet, 1999; Oltra-Massuet and Arregi, 2005). Whether this is needed for describing Latin however is

Below I discuss the subsystems in Fig. 7.1.

7.1.2 Voice

Latin doesn’t have rich valency changing devices: there is only one clause-wide valency decreasing device – passivization – and there is no valency increasing device. Causative constructions are realized by complement clauses, not any change in the argument structure. A verb (and hence the clause headed by it) is therefore either in **active voice**, or in **passive voice**. Some verbs are deponent (§ 7.5), which means they only have morphological passive forms.

It seems that the Latin passive is a catch-all realization of any argument structure and voice construction that lacks a prototypical agent: thus we have so-called *impersonal passives* (TODO: ref).

7.1.3 TAM categories

The categories of primary tense (past or not; future or not), secondary tense (perfect tense – or aspect – or not), imperfective-perfective distinction and realis-irrealis distinction are marked by the Latin tense and mood suffix in Fig. 7.1. The parameter of primary tense has three values: present, past and future; the past future tense seen in English is absent in Latin. The imperfective-perfective distinction and the secondary tense are fused into one category, with three possible values: plain, imperfect, and perfect. The “progressive perfect” category (*have being doing* in English) is missing in Latin. TODO: Some say the imperfective/perfective distinction (for example the *be doing* construction in English) is not syntactically coded in Latin? What’s perfect, then? The realis-irrealis distinction is discussed below as *mood*.

Table 7.1: Examples of Latin finite verbs

verb form	stem			tense and mood	personal ending
	extended root	extended root	aspect marker		
	extended root	thematic vowel			
<i>amō</i>	<i>am</i>				<i>ō</i>
<i>laudāmus</i>	<i>laud</i>	<i>ā</i>			<i>mus</i>
<i>olēvimus</i>	<i>ol</i>	<i>ē</i>	<i>v</i>		<i>imus</i>
<i>amāveris</i>	<i>am</i>	<i>ā</i>	<i>v</i>	<i>eri</i>	<i>s</i>

The composition of present, past and future tenses and plain, imperfect and perfect “aspects” gives nine options, but in Latin, the past simple and the present perfect are identified with each other, for obvious semantic closeness; the imperfect and simple “aspects” are fused when the primary tense is present or future. Thus there are only six morphologically distinguished options, as is shown in Table 7.2. In the Latin grammatical tradition, the term **tense** usually refers to these six options, instead of the past/present/future system.

Table 7.2: Latin tenses

	past	present	future
imperfect	IMPERFECT	PRESENT	FUTURE
simple	PERFECT		
perfect	PLUPERFECT	PERFECT	FUTURE PERFECT

Box 7.2: Mismatch between TAM constructions and fine-grained categories

Note that semantic TAM features are different from syntactic TAM features, and the two are in turn different from packaged TAM marking constructions that can be easily identified in surface-orientated analyses. This is illustrated in Table 7.2. Following the example in Grimm (2021), in this note, I use small capitals for the names of attested surface realizations of TAM and the default font for TAM values. (Some other grammars, like Jacques (2021); Friesen (2017), use initial capitals for the former.)

Similar fusion between categories is shown in the category of **mood**.¹ It’s the fusion of morphologically marked clause type (declarative and imperative) and morphologically marked modality. The verb morphology of interrogative clauses is exactly the same as declarative clauses: the interrogative clause type is marked by the existence of interrogative *pro*-forms. Thus, there are three moods in finite clauses in Latin: INDICATIVE, SUBJUNCTIVE, and IMPERATIVE. The INDICATIVE is the fusion of the declarative/interrogative clause type and the realis modality. The SUBJUNCTIVE mood is the fusion of the declarative/interrogative clause type and the irrealis modality. The IMPERATIVE is basically the imperative clause type: it doesn’t allow modality marking. Sometimes people say the infinitive is the fourth mood, though it’s a non-finite clause.

7.1.4 Agreement

Latin is a typical nominative-accusative language, both morphologically and syntactically. In finite clauses, there is subject-verb agreement: the number and person of the subject is marked on the main verb. In the case of periphrastic conjugation, the features are marked on the copula.

¹Dixon (2009) only calls the category of clause type *mood*. Huddleston and Pullum (2002), on the other hand, calls syntactic modality mood and uses the term *modality* for pure semantics. Different linguists use the term *mood* and *modality* in radically different ways. In this note I just focus on the common practice in Latin grammar study. In terms of general linguistics, the Latin *mood* is a mixture of clause type (the real mood) and modality.

7.1.5 Compatability of categories

There is no FUTURE tense and FUTURE PERFECT tense in subjunctive clauses, probably for the semantic reason that the future tense already contains certain sense of modality (an event predicted to happen), and thus is not compatible with the SUBJUNCTIVE mood. The IMPERATIVE mood is not compatible with other TAM markings except the PRESENT tense and the FUTURE tense. It's still compatible with the voice category, and allowed persons are second person singular/plural with the PRESENT tense, and second/third person singular/plural with the FUTURE tense. The absence of first person is also probably from semantic origin.

In conclusion, the categories involved in the finite verb paradigm of Latin are shown in Fig. 7.2. Here mood and tense are realized in one morpheme, and voice, person and number are realized in one morpheme. The paradigm is realized synthetically in all circumstances except in passive voice and perfect tense. In that case, the verb conjugation is realized like the English passive, i.e. via a copula and the perfect passive participle.

indicative	present/ imperfect/ future/ perfect/ pluperfect/ future perfect	active/ passive	1/ 2/ 3	single/ plural
subjunctive	present/ imperfect/ perfect/ pluperfect			
imperative	present		2	
	future		2/3	
mood	tense	voice	person	number

Figure 7.2: Categories in the finite paradigm

7.1.6 The non-finite paradigm

According to the morphology, Latin non-finite verb forms can be classified into the infinitives (§ 7.4.1) and the nominal forms (§ 7.4.2), the latter having noun-like or adjective-like morphology. Non-finite verb forms don't agree with the subjects they take, so there is no number or person category marked on them in the same way as Fig. 7.2, though for nominal verb forms there are number and person categories marked in the same way as the nominal morphology.

The infinitives include present active, present passive, perfect active, perfect pas-

sive, future active, and future passive infinitives. The latter three are realized periphrastically (Fig. 7.3).

The nominal verb forms include the **simple active**, the **perfect passive** (often just called the perfect participle), and the **future active** participles, the **gerund**, the **gerundive** which is also known as the **future passive** participle, and two supine forms. The **first supine** is identical in the form to the singular neutral accusative perfect participle, without any reference to the number category of any argument it takes. The **second supine** is identical to the singular neutral ablative or dative past participle, also with no inflection with respect to the number category of any argument it takes. The gerund is morphologically the singular neutral of the gerundive (§ 7.4.2.5).

I keep the traditional notion that the two supines are better recognized as separate morphological inflection forms, instead of two specific usages of the perfect participle and the past participle: although the two supines are identical to two other inflection forms in the surface form, that they don't have number agreement or case inflection means they are products of morphological devices that differ from the participle constructions.²

In Classical Latin, the gerund and participle forms are significantly more noun-like than their counterparts in English, and this also justifies the term *nominal form*, because they are not far from prototypically nominalization: although they are still modified by adverbs, they are unable to take arguments. In Ecclesiastical Latin, the so-called nominal forms are more verb-like (TODO: ref), being able to take arguments, and are therefore no longer “nominal”.

7.1.7 The two periphrastic conjugations

There are two additional periphrastic constructions beside the periphrastic forms in the perfect, passive part of the paradigm. In the first construction, the copula *sum* with arbitrary person, number and TENSE inflection (*including* the perfect TENSES) is placed together with the future active participle, to express the meaning of ‘I/you/he/she/it/they are going to do something’, TODO: a inchoative feeling? This construction is always active in voice. In the second construction the future active participle is replaced by the future passive participle; the meaning is ‘something should be done’.

The present active participle is almost never used in periphrastic conjugation. It can be seen that in periphrastic conjugation, the roles of future active participle and future passive participle and the role of passive perfect participle are not comparable: the latter dictates the secondary tense, i.e. it's perfect, as well as the voice; the former dictate voice and a “lower” aspect/modality category (inchoative or obligation), but not the secondary tense; and in particular, the term *future* is kind of (although not really) misleading: the two future participles carry aspect and modality values with them, which are semantically related to the concept of future; but the future primary tense is not involved.

²Similarly, the infinitive (or the “plain form”, since the infinitive is actually a label of clauses) is recognized as a form independent from the present form in English in Huddleston and Pullum (2002, p. 74).

7.2 Formation of stems

7.2.1 The three verb stems

Latin shows stem alternation that is not completely predictable. All verb forms can be obtained by three stems (Allen and Greenough, 1903, § 164), if the verb is regular:

- The **present stem**, which, after attached with proper endings, forms
 - The PRESENT, IMPERFECT, and future forms, regardless of whether they are indicative or subjunctive, active or passive. (There is no future or future perfect subjunctive).
 - All the imperatives.
 - The present infinitives, active and passive.
 - The present participle, the gerundive, and the gerund.
- The **perfect stem**, which, after attached with proper endings, forms
 - The perfect, pluperfect, and future perfect active, indicative or subjunctive. Again, there is no future or future perfect subjunctive. Note that the passives are *not* formed by the perfect stem.
 - The perfect active infinitive. (Or the perfective infinitive active, since infinitive is considered as a mood by some people.)

Note that the perfect passive participle is *not* obtained from the perfect stem.

- The **supine stem**, which, after attached with proper endings or used together with proper forms of *sum*, forms
 - The perfect passive participle, which, by being used with proper forms of *sum*, forms
 - * The perfect, pluperfect, and future perfect passive forms, indicative or subjunctive. Again, there is no future or future perfect subjunctive. This is periphrastic conjugation: it is done by using proper forms of *sum* with the perfect passive participle.
 - * The perfect infinitive passive.
 - The future active participle, which, used together with *esse*, makes the future active infinitive.
 - The future passive infinitive, by being used together with *īrī*.

This process is summarized in Fig. 7.3.

In a dictionary, typically the stems are not directly given – which are given are representative verb forms, from which the stems and the conjugation class can be inferred. The reasons are the follows. First, for fluent users, recording actually attested word forms is easier compared with the morpheme-based “anatomized” approach. Second, Latin has four conjugation types, and hence the three stems themselves aren’t sufficient to decide how to conjugate the verb: more information is needed, and by storing already conjugated verb forms, the conjugation class can be decided by observing the endings. What are stored are the following **principal forms**, from which the three stems and the conjugation class can be solved out (Allen and Greenough, 1903, § 172):

1. *The first-person present active indicative*: formed from the present stem.
2. *The present infinitive*: formed from the present stem. By observing its ending, the conjugation class can be decided, and by comparing with the first principal form, the present stem is obtained.
3. *The first-person perfect active indicative*: showing the perfect stem.
4. *The neutral accusative past participle*, i.e. the form of supine: showing the supine stem.

The ways to obtain the stems from the principal forms are:

- *The present stem* can be found by dropping *-re* in the PRESENT INFINITIVE (Allen and Greenough, 1903, § 175).
- *The perfect stem* can be found from the third principal part: just remove *-ī*.
- *The supine stem* can be found by dropping *-um* in the supine i.e. the fourth principal form (Allen and Greenough, 1903, § 178).

Note that in Medieval Latin, often, instead of *iri* plus the first supine, *fore* plus the perfect participle is used to form the future passive infinitive. TODO: find a reference <https://www.nationalarchives.gov.uk/latin/stage-2-latin/lessons/lesson-24-infinitives-accusative-and-infinitive-clause/>

7.2.2 Formation of the present stem

7.2.3 Formation of the perfect stem

7.2.4 Conjugation classes

Depending on the way realization of the paradigm for a verb, Latin verbs are traditionally divided into four conjugations classes according to the thematic vowel of the stem: if the stem ends in *a*, it's a first conjugation verb (§ 7.3.3); if it ends in *e* then we have a second conjugation verb (§ 7.3.4); TODO. The conjugation classes however can't be reduced to the thematic vowel, since there are some contextual allomorphs that can't be explained synchronically in this way (§ 7.3.1, Table 7.3). There are a handful of irregular verbs (§ 7.6) that can't be inflected using rules pertaining to the four regular conjugations.

Another aspect of the inflectional behavior of the verb is whether it's deponent; this also has some implications on the argument structure (§ 7.5).

7.3 The finite paradigm

Due to morphophonological rules, morpheme division inevitably involves controversies. This section follows the analysis in Oniga (2014, chap. 14).

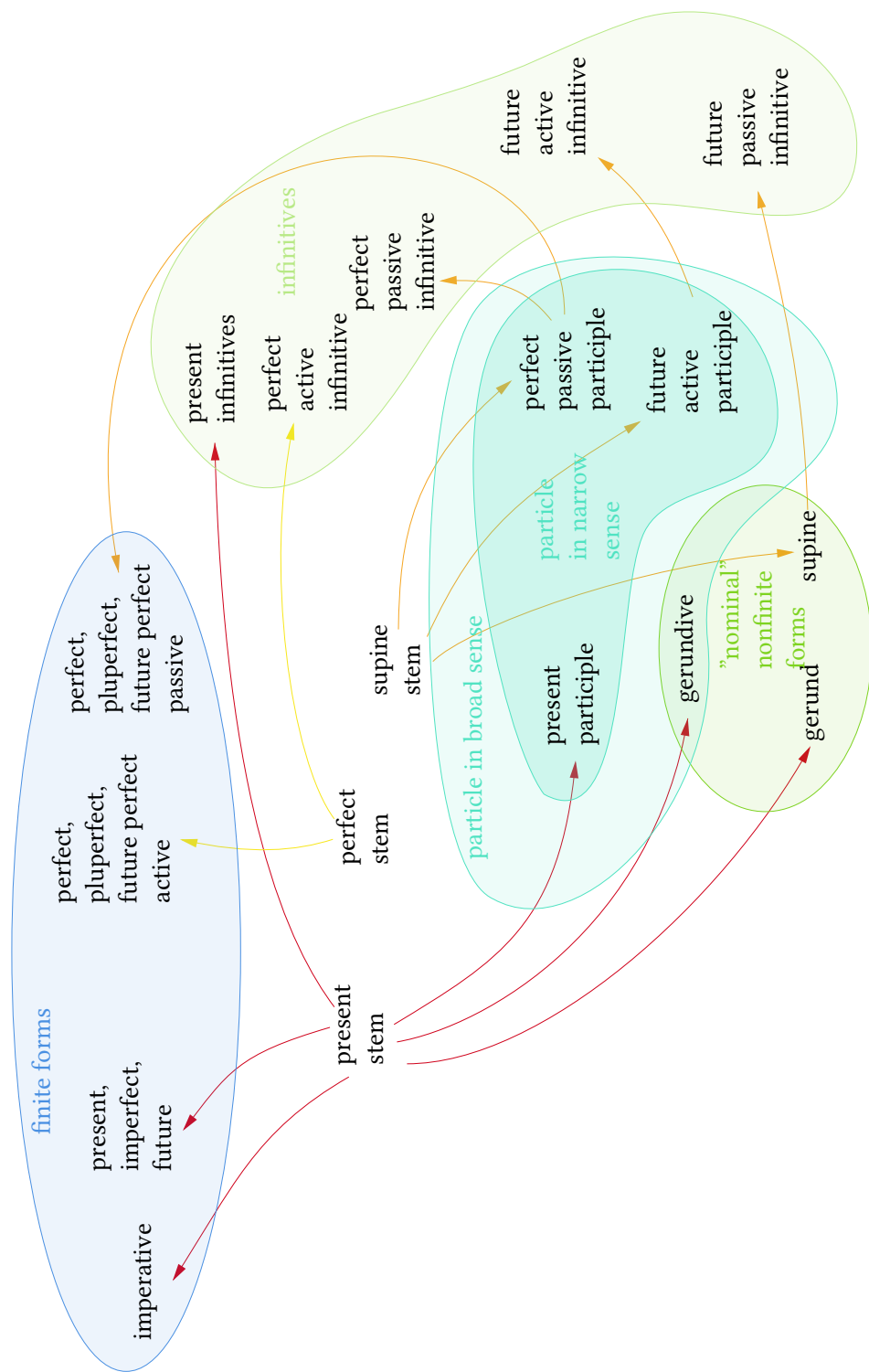


Figure 7.3: How to get all conjugation forms from the three stems

7.3.1 Marking of tense and mood

The contextual alternation of the tense and mood marker is listed in Table 7.3. They are subject to phonological rules and have allomorphs in different conjugation classes. A long vowel after the suffix, if any, is shortened before *-m*, *-r*, *-t*, and *-nt* due to vowel shortening in the final syllable before a consonant, and also before *-ntur* due to Osthoff's Law (§ 2.3.2; Oniga 2014, p. 130). The present indicative doesn't add a suffix after the stem, so the thematic vowel at the end of the stem is directly exposed to the personal ending, and vowel changes in § 2.3 happen.

The perfect and pluperfect subjunctive suffixes are only used for the active voice. For the passive voice, periphrastic conjugation with the perfect passive participle is used.

Table 7.3: Tense and mood suffixes

Suffix	Mood and tense	Note
∅	Present indicative	With change on thematic vowel
<i>-bā-</i> , <i>-ebā-</i>	Imperfect indicative	Possibly shortened
<i>-be-</i>	Future indicative	For 1c, 2c; Allomorphs: <i>-b-</i> , <i>-bi-</i> , <i>-bu-</i>
∅	Perfect indicative	With its own personal endings
<i>-erā-</i>	Pluperfect indicative	Possibly shortened
<i>-eri-</i>	Future perfect indicative	Allomorph: <i>-er-</i> (1sg)
<i>-ē-</i>	Present subjunctive	Allomorphs: <i>-a-</i> , <i>-ā-</i> , <i>-e-</i>
<i>-rē-</i>	Imperfect subjunctive	Possibly shortened
<i>-eri-</i>	Perfect subjunctive	For active only
<i>-issē-</i>	Pluperfect subjunctive	For active only; possibly shortened

Box 7.3: Shortening or lengthening?

If we just restrict ourselves to the verbal paradigm, it may be attempting as well to consider *-ba-* as the indicative imperfect suffix, since *-bā-* does not outnumber it. But prolonging is rare in phonology: shortening, due to physiological motion control factors (some sound sequences are easier to pronounce), is more frequent. The same line of argumentation can be applied to justify the status of *-rē-* as the somehow canonical subjunctive imperfect suffix.

Perfect indicative active verb forms have their own set of personal endings (Table 7.4). The plural part of the personal endings still contains the usual *-mus*, *-tis*, *-nt* endings, and thus alternatively, we may analyze the verb endings for the PERFECT tense as the follows: the tense suffix is *-ī-* for first-person singular, third-person singular and first-person plural, *-is-* for second-person singular and second-person plural, and *-ēru-* for third-person plural; the first person singular ending is empty, and the second person singular ending is *-tī*, and the rest of personal endings are the same as those of other tenses.

Table 7.4: Personal endings for the PERFECT

Suffix	Person and number
<i>-ī</i>	1SG
<i>-istī</i>	2SG
<i>-it</i>	3SG
<i>-imus</i>	1PL
<i>-istis</i>	2PL
<i>-ērunt</i>	3PL

- The indicative:
 - Future:
 - * For first and second conjugation verbs, the tense-mood morpheme is *-bi-*, except for first-person singular (which is *-b-*) and third-person plural (which is *-bu-*).
 - * For third and fourth conjugation verbs, change stem-final stem.
 - Future perfect:
 - * *-eri-* for all cases except first person singular.
 - * *-er-* for first-person singular.
- The subjunctive:
 - Present: no suffixation, but there is regular change on the stem-final vowel:
 - * For first conjugation verbs, .
 - * For second conjugation verbs, $\rightarrow e\bar{a}$ -.
 - * For third conjugation verbs, $\rightarrow \bar{a}$ -.
 - * For fourth conjugation verbs, $\rightarrow i\bar{a}$ -.

These alternations apply for both active and passive verbs, so they have nothing to do with polarity, and this is why I put them in this section.
 - Imperfect: *-rē-*, possibly shortened.

7.3.2 The personal ending

Possible personal endings are listed in Table 7.5. In the active voice, *-ō* is used with present indicative, future indicative (first and second conjugations only), future perfect indicative, and *-m* is used with imperfect indicative, future indicative (third and fourth conjugations only), pluperfect indicative, and the subjunctive mood regardless of tense. The PERFECT tense has its own personal endings: in the alternative analysis outlined in the last section, the first person perfect indicative ends with *-ī*, and the second person perfect indicative ends with *-tī*.

The ending *-re* is alternative form of second-person singular compatible with all non-periphrastic tenses and moods. If this personal ending is used, then the tense and mood marking is none. Note that the resulting verb form is the same as the infinitive principal part (§ 7.2.1).

The *-or* version of the passive first person ending is seen in present indicative and future indicative in the first and second conjugations; the latter has the deep form *-be-or*, then then is realized as *-bor* because of vowel deletion (§ 2.3.1).

Table 7.5: Personal endings

Suffix	Person and voice	Note
<i>-ō, -m</i>	active, 1SG	Depend on tense
<i>-s</i>	active, 2SG	
<i>-t</i>	active, 3SG	
<i>-mus</i>	active, 1PL	
<i>-tis</i>	active, 2PL	
<i>-nt</i>	active, 3PL	
<i>-r, -or</i>	passive, 1SG	<i>-or</i> for present and future indicative only
<i>-ris</i>	passive, 2SG	Allomorph: <i>-re</i> (without tense suffix)
<i>-tur</i>	passive, 3SG	Non-perfect tenses
<i>-mur</i>	passive, 1PL	Non-perfect tenses
<i>-mini</i>	passive, 2PL	Non-perfect tenses
<i>-ntur</i>	passive, 3PL	Non-perfect tenses

- The passive:
 - First-person singular:
 - * *-r*: compatible with all tenses and moods, except the present indicative.
 - * *-or*: present indicative. Also, note that the future indicative (first and second conjugations only) ending is *-bor*, which may be analyzed as .
 - Second-person singular:
 - * *-ris*: compatible with all non-periphrastic tenses and moods.
 - *

Box 7.4: Parsing a finite verb

Follow these steps to parse a finite verb:

- First see whether the aspect is perfect (by looking at the stem part) and the personal ending.
- Then TODO

7.3.3 The first conjugation

Attested thematic vowel alternations include:

- In first person singular present forms, the final *ā* is dropped, because the personal ending is *-ō* or *-or* and the thematic vowel is subject to vowel deletion.
- In active subjunctive *ā-* → *ē-*.

7.3.4 The second conjugation

Attested thematic vowel alternations include:

- In active indicative present forms, $\bar{e} \rightarrow e$.

7.3.5 The third conjugation

The third conjugation has two subclasses: the thematic vowel-less case, and the *i*-thematic case.

7.3.6 Periphrastic conjugations

An auxiliary verb construction is a structure that contains one or more auxiliaries apart from the main verb, and yet is mono-clausal and is not a complement clause construction, and the auxiliaries are realizations of the verbal system surrounding the main verb, and not independent lexical verbs. Auxiliary verb constructions realizing grammatical categories that are usually realized by inflectional endings in a paradigm should be seen as a part of that paradigm, and therefore are known as periphrastic conjugations.

7.4 Non-finite forms

7.4.1 The infinitives

7.4.2 The gerund and participles

7.4.2.1 The gerund

The gerund is morphologically a neutral singular noun. The stem is formed by adding *-nd-* to the present stem; in other words, the accusative form of the gerund of the verb is obtained by removing the final *s* of the present active participle and adding *dum*. Note that the nominative case is missing – when a non-finite clause is required in the subject position, it's always an infinitive.

7.4.2.2 The present active participle

The present active participle (i.e. the present participle) is morphologically a third declension adjective (TODO: gender). The stem of the present active participle is obtained by adding *-nt* to the present stem. Equivalently, the nominative singular form – the citation form – is obtained by replace the *-re* ending of the present active infinitive by *-ns* (or in other words, add *-ns* to the present stem).

7.4.2.3 The perfect passive participle

The perfect passive participle (i.e. the perfect participle or the past participle) can be found by declining the neutral accusative past participle, i.e. the fourth principal part.

7.4.2.4 The future active participle

To get the future active participle (i.e. the future participle), add *-urus* to the supine stem.

7.4.2.5 The future passive participle

The future passive participle, or in other words the gerundive, can be obtained by declining the gerund as if it's an adjective (TODO: declension class); in other words the gerund is morphologically the singular neutral of the gerundive.

7.5 Deponent verbs

Deponent verbs are verbs that are in the passive voice realizationally, either with the corresponding morphological marking (1) or periphrastic marking (2), but still have active meaning. An observed tendency is that so-called “unaccusative” verbs seem to be deponent verbs in Latin, which means the passive voice essentially is the “non-agentive” voice (§ 9.4.1).

- (1) [Confiteor] unum baptisma in
confess-IND.PRES.PASS.1SG one-SG.ACC baptism-SG.ACC into/towards/about
remissionem peccatorum ...
remission-SG.ACC sin-PL.GEN
'I confess one baptism for the remission of sins ...' (Nicene Creed)
- (2) ... qui [locutus est] per prophetas ...
who speak-APRT-SG.NOM BE-IND.PRES.ACT.3SG through prophet-PL.ACC
'...who has spoken through the prophets ...' (Nicene Creed)

7.6 Irregular verbs

7.6.1 The verb *sum*

7.6.1.1 Overview

The verb *sum* has lots of uses in Latin grammar (§ 8.2.3.1), and its inflection is (unfortunately but expectedly) highly irregular. It's also defective: it has no passive forms, either finite or nonfinite. The principal parts (§ 7.2.1) are *sum*, *esse*, *fuī*, with the supine form being absent – usually replaced by the future active participle *futūrus*.

From the principal parts, we find the perfect stem is *fu-*, and the supine stem – if we insist on defining it – is the same, although the perfect passive participle is absent and so is the supine, and therefore the supine stem only appears in the future active participle.

The present stem is not well-defined: the second principal form *esse* doesn't have the regular infinitive ending *-re*, though we can roughly recognize something like *es-* or *e-*; the first principal form *sum* gives *su-* or *s-*. The two stems appear in the finite paradigm in an unpredictable manner, also with irregular though still recognizable endings. Besides *s-* and *es-*, there is also *fo-* seen in one variant of the future active

infinitive (§ 7.6.1.2), which also appears in variants in the subjunctive active imperfect part of the finite paradigm.

7.6.1.2 The nonfinite paradigm

The only nominal form is the future active participle *futūrus*. The three active infinitives forms are all attested. The present active infinitive is *esse*. The perfect active infinitive is *fuisse*, regularly formed by the perfect stem *fu-*.

The future active infinitive can be regularly formed by adding *esse* to the future active participle, and therefore is *futūrum esse*. There is also a free variant *fore*.

7.6.1.3 The perfect system

The perfect forms – finite forms and the perfect active infinitive – of *sum* can be formed regularly (§ 7.3) according to the perfect stem *fu-*.

7.6.1.4 The imperative system

The present imperative system, which is known for reflecting the present stem, is formed regularly using *es-*: the singular second person present imperative is *es* and the plural second person present imperative is *este*.

7.6.1.5 The present system

The imperfect forms of *sum* are highly irregular, though patterns can still be found. In the indicative part (Table 7.6):

- The PRESENT forms show no pattern except the personal endings. Note that here *-m* instead of *-ō* is used for the first person singular form.
- The IMPERFECT forms are formed by adding the standard personal endings (*-m*, *-s*, *-t*, *-mus*, *-tis*, *-nt*) to *erā*, the vowel *ā* of which undergoes shortening according to rules in § 7.3.1.
- The FUTURE forms are formed by the same personal endings seen in the first and the second conjugations, although the tense marker isn't the same: the stem-tense marker complex is *er-* instead of the stem plus *-b-*.

Table 7.6: The indicative paradigm of *sum*

PRESENT	IMPERFECT	FUTURE
<i>sum</i>	<i>eram</i>	<i>erō</i>
<i>es</i>	<i>erās</i>	<i>eris</i>
<i>est</i>	<i>erat</i>	<i>erit</i>
<i>sumus</i>	<i>erāmus</i>	<i>erimus</i>
<i>estis</i>	<i>erātis</i>	<i>eritis</i>
<i>sunt</i>	<i>erant</i>	<i>erunt</i>

In the subjunctive paradigm (Table 7.7), we find that in the PRESENT system, the stem-tense marker complex is fused into *sī-*, and in the IMPERFECT system, the stem-tense marker complex is fused into *essē-* or *forē-*, both of which are then attached to the standard *-m*, *-s*, etc. personal endings, and the vowel shortening rule in § 7.3.1 works.

Table 7.7: The subjunctive paradigm of *sum*

PRESENT	IMPERFECT
<i>sim</i>	<i>essem, forem</i>
<i>sīs</i>	<i>essēs, forēs</i>
<i>sit</i>	<i>esset, foret</i>
<i>sīmus</i>	<i>essēmus, forēmus</i>
<i>sītis</i>	<i>essētis, forētis</i>
<i>sint</i>	<i>essent, forent</i>

7.6.2 The verb *faciō*

The verb *faciō* looks pretty regular regarding the endings, except for one thing: its *stem* alternates according to the voice.

7.7 Auxiliary verb constructions

An auxiliary verb, despite morphologically being a verb, is embedded into the TAM system of a language. The following aspects can be used to identify an auxiliary verb from a lexical one:

- *Standalone or not.* A lexical verb may appear on its own; an auxiliary verb never does this.
- *Scope and hierarchy compared with other TAM marking.* The relative order of complement-taking lexical verbs can be arbitrary; this is not true for auxiliary verbs.
- *Interplay with argument structure: voice.*
- *Interplay with argument structure: complement types.* A lexical verb usually imposes semantic confinements on the complement
- *Interplay with clause type.* Some auxiliary verbs never appear in certain types of non-finite constructions, because the TAM categories labeled by them are absent in these non-finite clauses.

Latin is usually perceived as a language with few analytic properties. Traditionally only *sum* and its forms are considered auxiliary verbs when they are used with the perfect passive participle to form perfect TAM constructions in passive voice. Some however argue that there are more auxiliary verbs.

Chapter 8

Verb frames

This chapter is mainly about verbs that don't take complement clauses as arguments. The phenomena discussed in this chapter mostly apply to complement clause constructions as well, but complement clause constructions have their own peculiarities (§ 12.1).

8.1 The argument-adjunct distinction

A clear complement-adjunct distinction – telling peripheral arguments from core arguments or oblique arguments – is hard to establish in Latin. Below I assess several parameters usually used to draw the distinction listed in Huddleston and Pullum (2002, § 4.1.2):

- *Content of clausal dependent.* Latin peripheral arguments do not necessarily have prepositions.
- *Easiness of topicalization, etc.* Latin is highly free-ordered and therefore all clause dependents can leave their base positions.
- *Obligatoriness.* Latin is also highly *pro*-drop, and even uncontroversial core arguments can be omitted.
- *Licensing by verb.* Oblique arguments are frequent in Latin, as is the case in English (consider *run away from* or *get into*).
- *Anaphora referring to core verb phrase.* Latin doesn't have systematic way to replace the core predicate (i.e. without adjuncts) by an anaphora.

Selection, licensing, and obligatoriness tests can still be used in more subtle ways (e.g. reference in discourse to identify omitted argument)

these criteria are however hard to use for a classical language.

Box 8.1: Argument and adjunct: directional construction

If a directional phrase obligatorily has the subject/object as its subject, then it has to be a complement and not an adjunct.

Thus, despite I'm fully aware that clausal dependents concerning place, instrument, mean, etc. may be licensed by both the argument structure of the verb and by clausal adjunct positions and may have clear structural differences in

other languages (as in English), currently no distinction between the two cases is made. Following the traditional notion, the subject, several kinds of objects, the copular complement (Huddleston and Pullum (2002) calls it *predicative complement*) are identified as complements.

8.2 Argument positions and verb frames

Note that a clausal dependent position is an all-in-one label concerning several grammatical relations: the label *subject*, for example, covers both the pivot role in the nucleus clause (i.e. “SpecTP” in generative grammar) and the most external, agentive role in the argument structure (i.e. “SpecvP” or “SpecVoiceP” in generative grammar).

Box 8.2: Psych-verbs

Structural and semantic properties of psych-verbs; UTAH.

8.2.1 Prototypical intransitive verbs

8.2.1.1 The MOTION type

8.2.2 Prototypical transitive verbs

With complement-taking verbs temporarily excluded, a prototypical transitive verb is more or less close to the AFFECT type, with an A argument which is the causer of the event

8.2.3 Copular verbs

8.2.3.1 The verb *sum*

It’s also possible to use *sum* with an indirect object, and the meaning because ‘something be to [someone]_{indirect object}’. In this case we get the possessive dative construction (Allen and Greenough, 1903, § 373).

TODO:

- (1) Nono, utrum uti debeat metaphoricis vel
tenth whether at.any.rate owe-SUBJ.PRES-3SG.ACT metaphoric-PL.DAT/ABL or
symbolicis locutionibus.
symbolic-PL.DAT/ABL speech-PL.DAT/ABL
‘Whether (it) at any rate uses metaphoric or symbolic speech.’ (*Summa*, I q. 1 pr.)

8.3 Preverb

Although roughly speaking, the preverbs are adverbs or preposition incorporated into the verb word, non-trivial morphosyntactic interaction between the preverb and the verb root exists. Prefixation only takes place when a PATH category is licensed by

the verb (which in Latin requires the event to be telic, and therefore the IMPERFECT tense and the like may be a problem for prefixation); Some case distributions of prefixed verbs, which seem to be more or less affected by the verb, are never attested if the preverb is not attached to the verb root; the meaning of the verb may also vary depending on whether the preposition is attached to the verb (Mare, 2018). Interestingly, the GROUND NP can be dative and this case is semantically meaningful because the GROUND indeed is the goal of the action, which seems to indicate that the requirement that the NP governed by prepositions has to be accusative or ablative is at least partially due to realizational reasons: when a preposition like *circum* appears in the locational construction, it dictates the case assignment, but when it is moved out of it, the case assignment can be whatever semantically makes sense.

Chapter 9

Clause structure

9.1 Internal arguments

9.1.1 The direct object

Here is a list of grammatical properties of the direct object:

- *Coding of semantic role*: In a prototypical transitive argument structure, the direct object is the P argument, i.e. the most patientive argument.
- *Case marking*: Direct objects are always accusative when it makes sense to talk about case – but not all accusative arguments are direct objects (§ 5.4.1.2).
- *Passivization*: If an argument is coded as the direct object, then it can regularly be promoted to the subject position in a passive clause (§ 9.2). Secondary objects are less frequently promoted in passivization (§ 5.4.1.2).

9.1.2 The indirect object and the secondary object

Latin also has two complement positions named as object: the indirect object and the secondary object. The indirect object is distinguished by the following grammatical properties:

- *Coding of semantic role*: in a AGT-type argument structure, the indirect object is usually the G argument. Intransitive clauses sometimes also have indirect objects, and an indirect object, in this case, is also a G argument.
- *Case marking*: indirect objects are always dative (§ 5.4.1.4).
- *Passivization*: indirect objects are always retained in passive clauses. They are never promoted to subjects in passivization.

The secondary object is distinguished by the following grammatical properties:

- *Coding of semantic role*: in an AGT-type argument structure that is always about information flowing, the T argument (i.e. the thing asked about or taught about) is the secondary object. The G argument (i.e. the person who is asked or taught) is the direct object. Sometimes the G argument is ablative, and in this case, there is only one accusative argument: the secondary object. Another place

where secondary objects appear is clauses headed by a verb with a compounded accusative preposition.

- *Case marking*: secondary objects are always accusative.
- *Passivization*: secondary objects can be passivized, but much more rarely than direct objects.
-

The distributions of the secondary object and the indirect object are mutually exclusive. This means for ditransitive verbs of type GIVING, Latin shows a clear and strong tendency to identify the T argument with the monotransitive O, while for ditransitive verbs about teaching, the inverse is true.

Box 9.1: Comparison with English

It can be found that the Latin indirect object has more similarity with the English *to*-PP, which is also called the indirect object in some grammars, but not in Huddleston and Pullum (2002). The Latin indirect object differs from the English (accusative) indirect object in passivization. Since in Latin, verbs with AGT-type argument structure do not have alternation of complementation pattern – in English we have *give sth. to sb.* and *give sb. sth.*, while in Latin there is only the former one, but *to sb.* is replaced by a dative, (always with no preposition) – the G argument is identified with the E argument, and the T argument is identified with the P argument. In other words, in Latin, there is only the *John gave [goods]_T to [charity]_G* pattern: the double-object *John gave charity goods* pattern is absent.

Therefore, for typical ditransitive verbs, i.e. verbs like *give*, Latin shows a clear and strong tendency to identify the T argument with the monotransitive O, which is more typical than English^a, but for verbs with meaning of TEACH or ASK, there is also a clear and strong tendency to identify the G argument with the monotransitive O. The term *secondary object* is coined to cover this grammatical position.

^aIn English, in the *give sb. sth.* construction, it is the person i.e. the G argument that is passivized, while the T argument i.e. *sth.* cannot, though the latter is identified with monotransitive O according to other criteria.

9.1.3 Copular complements

Latin also has copular complements. A copular complement, just like its counterpart in English, basically can be viewed as a displaced attributive or appositive (and hence is prototypically filled by an NP or an AdjP) but is a little more peripheral (manner, state, factitive, etc.) in its meaning than an attributive or appositive.

Latin has nominative predicate and accusative predicate: as hinted by their names, the nominative predicate gives a property of the subject and agrees with it, and the accusative predicate gives a property of the direct object and agrees with it. In passivization of the direct object, the accusative predicate becomes the nominative predicate.

Other types of copular complements without agreeing with the subject exist. TODO: ablative of quality, price, etc. The syntactic status of copular complements here are closer to PPs: we may say they receive *inherent cases*, while the nominative and accusative copular complements receive *structural cases* (§ 5.4.1).

9.1.4 (Change of) location

TODO: considering moving this section to the case section

Various semantic roles can be summarized as SOURCE, and the source clausal dependents – adjunct or complement – have the following properties. Note that we are dealing with a *group* of clausal dependents.

- *Coding of semantic role*:
- *Case marking*: a source argument is in the ablative case. It may come together with the prepositions *ex* or *ab*.
- *Passivization*: not available.

9.1.5 Others

There are other clausal dependents with semantic roles and case/preposition markings different from any other type mentioned above; they are skipped here for brevity. A full list of these clausal dependent slots can be found by checking the usage of each preposition.

9.2 Voice

9.3 Preverbs and other verbal derivations

9.4 Verb frames

9.4.1 Overview

In this section I consider the correspondence between semantics of verbs and their syntactic properties. Attested combinations of the argument slots include

On the semantic side, I follow the semantic classification by Dixon (2005, Part B), Dixon (2010, § 18.5), and Dixon (2009, § 3.3). Verbs are classified semantically into three subgroups:

1. Primary-A, which contains verbs that don't take arguments with meanings similar to those of complement clauses,
2. Primary-B, which are semantically **complement-taking**¹ and **lexical**, which have arguments that are semantically equivalent to complement clauses (but not necessarily syntactically coded as complement clauses) and have meanings more complicated than what's expected for grammatical items, and
3. Secondary, members of which have the same *meaning* of certain grammatical constructions in the verbal system, but not the same syntactic properties (for example, they may just take complement clauses instead of being auxiliary verbs).

¹Here the term means "semantically equivalent to a complement clause construction".

Note that the valency class of the verb is strongly related to but is not determined by the semantics of the verb.

I will generally follow the classification in Pinkster (2015, chap. 4); some verb classes enumerated by Pinkster are discussed in more details in chap. 12, since a full account of their behaviors is closely related to the structure of complement clauses.

Chapter 10

Constituent order and the information structure

Patterns in Latin constituent order are often overlooked in traditional grammar. Still, fine-grained constituency is demonstrated by the relation between *non* and the auxiliary (§ 10.3), radical change of VO/OV frequency when structural ambiguity is controlled (§ 10.6), and usual constituency tests (Danckaert 2017, § 1.6; TODO: ref to my own analysis in conjunction). The *non*-before-auxiliary condition implies a auxiliary hierarchy just like the one in English, although it's not as developed as the latter since Latin is inflectionally rich. The constituency tests hint on at least the subject-VP binary branching. The fact that superficial VO/OV orders may have structural ambiguity means it's likely that some of the constituent orders are comparable to English poetry in imitation of Latin (Allen and Greenough, 1903, § 600), Japanese scrambling and topicalization (TODO: ref).

Available evidence supports the tradition in existing secondary literatures that Latin is thus better described as a discourse-configurational language, with multiple topicalization and focusing structures (Oniga 2014, p. 189; Danckaert 2017, p. 77; Devine and Stephens 2006, among others). Initial positions in clauses clearly bear information structure functions. Constituents that are able to move to the positions include almost everything: arguments, adverbials, the negator *non*, and also the verb (TODO: aux) (Allen and Greenough, 1903, § 598). TODO: is it possible for the main verb to move to the initial point only? Note that fronting of the verb may be used to focus the verb root or the *tense* (Allen and Greenough, 1903, p. 397); this means preposing of the verb is comparable to stressing the verb in English.

Apart from the information structure, prosody is also an important factor TODO

Pragmatically unmarked sentences have the arguments as pragmatic topics and the rest of the VP are the focus: 'we already know Baebius, his army, and Pinarius (topics); the piece of new information is that a transferring process happens involving the three'. This kind of sentences, known as "broad scope focus" sentences in Devine and Stephens (2006, p. 15) because the scope of focusing is broad and not restricted to a single argument, demonstrates a constituent order which may be referred as the **neutral order** in Latin (Devine and Stephens, 2006, p. 79), although its frequency – without controlling the information structure – isn't significant higher than other constituent orders. This order clearly isn't a faithful representation of the argument structure, since the direct object – the argument that is supposed to be the closest one to the main verb – appears far from the main verb. The reason is likely to be that the

“neutral order” also marks the aforementioned “unmarked” information structure, in which the arguments are by default topicalized.

Deviation from the “neutral” order may be divided into the following levels:

1. Word order alternation that does not change the broad scope focus of the clause: this is the realignment of the (rather weak) information structure statuses of the arguments.
2. Scrambling, where a constituent is shifted leftward and when information-neutral manner adverbs are present (whether the adverb undergoes some movement can be tested by its scope (Devine and Stephens, 2006, p. 100)), to the left of the manner adverb. The scrambled constituent generally receives a presuppositional or weakly focal interpretation (Devine and Stephens, 2006, pp. 102,108-109).
3. Finally, we have strong (and more prototypical) topic and focus constructions (Devine and Stephens, 2006, chap. 3).

The usual typological classification of Latin as a SOV language, despite being misleading in suggesting a rigid base order, still makes sense in pointing out that indeed constituents that are “higher” are moved leftwards.

In periphrastic conjugation, the constituent order is subject + object + verb + *sum*. This may also show a mismatch between the dependency structure and the linear order, since if we consider the auxiliary *sum* to be the analytic counterpart of the inflectional suffix and has a higher position compared to the main verb (i.e. the participle), then since the subject and topicalized constituents usually appear on the left side, it also should appear on the left side compared with the participle under it. But the case may just be that the participle is focused and is moved leftward by default, just like the topicalized direct object, so the auxiliary then appears at the end of the clause in the neutral order.

10.1 Positioning of arguments

10.2 Positioning of the verb (without auxiliary)

Box 10.1: Position of the verb

The position of the verb involves a theoretical question: is its appearance away from the unmarked clause-final position due to phrasal movement (the verb root being moved to a new position, carrying all suffixal realizations of TP functional heads together with it), or is it due to being attracted by some sort of functional head (in this case the verb root is just like a head in head movement; similar mechanisms appear in, say, *on the top of the mountain lies a small village*)? Since this distinction is hard to test, I refrain from picking up one explanation.

10.3 Positioning of auxiliary and negation

One piece of evidence suggesting the grammatical status of *sum* is somehow different from a lexical verb is that its position has non-trivial interaction with the position of

non. The negator *non* usually appears before the verb (Danckaert 2017, § 1.5, TODO: or aux?), and apparent violations seem to be constituent negation as opposed to sentential negation (Danckaert, 2017, p. 43).

10.4 On so-called postposing constructions

Whether postposing exists as an information structure marking device is still not completely clear. It's said that postposing is never used for emphasis (Allen and Greenough, 1903, p. 395), and apparent counterexamples are all "afterthoughts" TODO; but

The postponed subject is likely to be an afterthought, coindexed with a zero pronominal in the rest of the sentence before it Devine and Stephens (2006, p. 87), comparable to English *it kills three people, the wandering puma*.

10.5 Notes on some typologically rare constituent orders

10.6 Historical evolution

Without sentences in which the OV/VO alternation can be alternatively analyzed as topicalization, VO frequency no longer shows significant change as time went by, indicating a well-defined extended verb phrase (Danckaert, 2017, § 1.5, p. 29).

10.7 Information packaging constructions

10.8 Existential clause

In the existential construction, the *sum* verb always appears first (Allen and Greenough, 1903, p. 396).

10.9 Cleft construction

nequitia est quae te non sinit esse senem

10.10 Question formation

Chapter 11

Adverbial clauses

Chapter 12

Complement clause constructions

12.1 Overview

12.1.1 Infinitives

12.1.1.1 *Accusativus cum infinitivo*, or the autonomous infinitive

Despite the superficial resemblance to the English object raising constructions, Latin complement infinitives with accusative subjects are licensed even after complete nominalization of the complement-taking verb. Therefore the accusative subject of the complement clause can't be seen as an object of the complement-taking verb, since the nominalized verb no longer takes object in Latin. Therefore, the *accusativus cum infinitivo* construction is comparable to English *for sb. to do sth.*, where the subject of the complement clause is autonomous; the accusative case here is the case assigned to the subject of a non-finite clause where the nominative case is not available. This is not completely unexpected, since even absence of an explicit complementizer is observed in the *ut* clause as well (Oniga, 2014, pp. 290-292).

Chapter 13

Relative constructions

13.1 Agreement properties

The case of a relative pronoun is determined by its syntactic position in the relative clause, and *not* the case of the antecedent, though the number and gender categories are determined by agreement with the antecedent.

Chapter 14

Coordination

Latin coordination in the nominal system and the verbal system shows strong correspondence, with most conjunction words being shared by the two systems.

Chapter 15

Texts

Below are some examples of Latin texts, in an order from the easiest to the hardest, with remarks on their vocabulary and grammar.

15.1 Liturgy texts

15.1.1 Short formulae in the Roman Mass

Examples in this section are short formulae found in the Roman Mass in the order of their appearance. In (1, 2), *nomine* and *patris* are third declension nouns, while *spiritus* is a fourth declension noun.

- (1) In Nomine Patris, et Filii, et Spiritus
in name-SG.ABL Father-SG.GEN and Son-SG.GEN and SPIRIT(M)-SG.GEN
Sancti.
holy-M.SG.GEN
'In the name of the Father, and of the Son, and of the Holy Spirit.'
- (2) – Dominus vobiscum. – Et cum spiritu tuo.
Lord(M)-SG.NOM 2PL.ABL and with spirit(M)-SG.ABL your-M.SG.ABL
'– The Lord be with you. – And with your spirit.'
- (3)

15.1.2 Nicene Creed

- (4) Credo in unum Deum, Patrem
believe-IND.PRES.1SG in one-M.SG.ACC God(M)-SG.ACC father(M)-SG.ACC
omnipotentem,
omnipotent-M.SG.ACC
'I believe in one God, (the) omnipotent Father,'
- (5) factorem caeli et terrae, visibilium omnium et
maker-
'maker of'

15.2 Vulgate bible

15.2.1 Excerpts in John 1

- (6) in principio erat Verbum et Verbum erat apud Deum et Deus erat Verbum
in be.IMPF
'In the beginning' (John 1:1)
- (7) omnia per ipsum facta sunt et
all-N.PL.NOM through DEM-ACC make.PPRT-N.PL.NOM be.IND.PRES.3PL and
sine ipso factum est nihil
without DEM.ABL make.PPRT-N.SG.NOM be.IND.PRES.3SG nothing.NOM
quod factum est
REL.N.3SG make.PPRT-N.SG.NOM be.IND.PRES.3SG
'All have been made through exactly this (i.e. the Word of Lord), and without exactly this, nothing that has been made has been made.' (John 1:3)

As an example, below I show how (7) can be parsed. First we can see a *et* dividing the sentence into two branches.

1. For the first branch, we know *omni-* is a quantifier meaning *all*, and morphologically it's a twin-termination third declension adjective; then from Table 5.1 and the fact that we are dealing with a third declension word, the ending *-a* means neutral and PL.NOM/ACC/VOC. The vocative case is of course impossible here.
2. The word *per* is a preposition taking an accusative object. *Ipsum* is a basic identity demonstrative, with the meaning of "exactly this". Since it follows *per*, the ending *-um* here seems to be the accusative case marker, instead of a neutral nominative case marker.
3. The sequence *facta sunt* contains the indicative perfect 3pl copula *sunt*, and in *facta*, we see the supine stem *fact-* of the verb *faciō* 'make'. The second fact means *facta* should be the perfect passive participle in a certain inflection form. Then *facta sunt*, collectively, is the indicative passive perfect 3pl form of *faciō*. (Here we are fortunate: it's possible that *facta* and *sunt* get scattered to different places.) The *-a* ending can again be looked up for in Table 5.1: the possibilities are PL.NOM/ACC/VOC – note that the first declension singular possibilities are excluded by the fact that *sunt* is in plural form. We expect *facta* to be nominative because it has to agree with the subject, which is always nominative and it turns to be possible.
4. Now we should link things together. The open ends are: the case of *omnia*, and the (3pl) subject of *facta sunt*. Then quite obviously, we find *omnia* should be in the subject position, and therefore everything works well.
5. We can also check gender agreements to make sure our reading is correct.

The second half is done in similar manners. The structure of the text looks like this:

[[omnia]_{subject} [per ipsum]_{instrument:PP} [facta sunt]_{verbal complex}]_{coord} et [[sine ipso]_{adverbial:PP} [factum est]_{verbal complex} [nihil [quod factum est]_{rel}]_{subject}]_{coord}

15.3 *Aeneid*

15.3.1 Introduction

(8) Arma virumque cano, Troiae
weapon(N)-PL.ACC man(M)-SG.ACC=and sing-IND.PRES.1SG Troy-SC.GEN
qui primus ab oris Italiam,
REL.M.SG.NOM first-M.SG.NOM from shore(F)-PL.ABL Italy(F)-SG.ACC
fato profugus, Laviniaque venit
fate(N)-SG.ABL exiled-M.SG.NOM Lavinia-TODO=and go.to-IND.PRES.3SG
litora, multum ille et terris iactatus et alto vi superum saevae
shore(N)-PL.ACC
memorem Iunonis ob iram; multa quoque et bello passus, dum conderet
urbem, inferretque deos Latio, genus unde Latinum, Albanique patres, atque
altae moenia Romae.

‘I sing weapons and a man, who was the first from the shores of Troy to Italy,
was by fate exiled, and ’

In (8), it should be noted that *arma* is in plural only. The *qui* clause is an example of the rule that the relative pronoun doesn’t agree in case with the antecedent (§ 13.1). The copula is omitted in the *qui* clause.

15.4 *Summa Theologiae*

15.4.1 Introduction

Thomas Aquinas, a theologian and philosopher, was recognized by the Catholic Church as one of the Doctors of the Church, usually known as *Doctor Angelicus* (doctor-SG.NOM angle-ADJ-SG.NOM, ‘Angelic Doctor’). *Summa Theologiae* (summary-SG.NOM theology-SG.GEN, ‘Summary of Theology’) or *Summa Theologica* (summary-SG.NOM theology-ADJ-SG.NOM.F), usually abbreviated as *Summa*, is probably his most known work. As its name implies, the book is a summary of Catholic theology, containing necessary information for beginning theological students and for arguing against frequently seen heresies. Apart from its great value within and out of Christianity, the book is a good example of what Medieval Latin looks like.

One salient feature of *Summa* is it’s written in a rather strict and even dull format. The work is divided into three Parts, each containing many Questions. Each Question (i.e. a topic, like “the nature and extent of this sacred doctrine”) is divided into several Articles (i.e. a specific question, like whether the sacred doctrine is a science). Each Article starts with a sentence like (9) indicating its index within the Question it belongs to, and then several Objections are raised to support an unorthodox idea; then the accepted doctrine contrary to the above heretical claims is given (10), followed by a comment by Thomas on this Article (11).

(9) Ad primum sic proceditur.

- (10) Sed contra est quod dicitur II
 but contrarily be.IND.PRES.ACT.3SG what.R.SG.N say-IND.PRES-3SG.PASS two
 ad Tim. III ...
 to(TODO) NAME three
 ‘But on the contrary there is what is said in 2 Timothy 3 ...’ (*Summa*, I q. 1
 s.c.)
- (11) Respondeo dicendum quod ...
 respond-IND.PRES-1SG.ACT
 ‘I respond to what is said that ...’

References

- Joseph Henry Allen and James Bradstreet Greenough. *Allen and Greenough's New Latin grammar for schools and colleges: Founded on comparative grammar*. Ginn, 1903.
- Irena Botwinik-Rotem and Arhonto Terzi. Greek and hebrew locative prepositional phrases: A unified case-driven account. *Lingua*, 118(3):399–424, 2008.
- Lieven Danckaert. *The development of Latin clause structure: A study of the extended verb phrase*, volume 24. Oxford University Press, 2017.
- Andrew M Devine and Laurence D Stephens. *Latin word order: Structured meaning and information*. Oxford University Press, 2006.
- Robert MW Dixon. *A semantic approach to English grammar*. OUP Oxford, 2005.
- Robert MW Dixon. *Basic linguistic theory volume 1: Methodology*, volume 1. OUP Oxford, 2009.
- Robert MW Dixon. *Basic linguistic theory volume 2: Grammatical topics*, volume 2. Oxford University Press on Demand, 2010.
- David Embick and Morris Halle. Latin inflections. *Paper delivered at Going Romance*, 2003.
- David Embick and Morris Halle. On the status of stems in morphological theory. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, 270:37, 2005.
- Dianne Friesen. *A grammar of Moloko*. Language Science Press, 2017.
- Jacopo Garzonio, Silvia Rossi, et al. Functional and lexical prepositions across germanic and romance. In *Continuity and Variation in Germanic and Romance*, pages 450–471. Oxford University Press, 2021.
- Giuliana Giusti and Rossella Iovino. A split-dp hypothesis for latin and italo-romance. *Complex Visibles Out There. Olomouc Modern Language Monographs*, 4:127–143, 2014.
- Nadine Grimm. *A grammar of Gyeli*. Language Science Press, 2021.
- Rodney Huddleston and Geoffrey K. Pullum. *The Cambridge Grammar of the English Language*. Cambridge University Press, 2002. doi: 10.1017/9781316423530.
- Guillaume Jacques. *A grammar of Japhug*, volume 1. Language Science Press, 2021.

- María Mare. Issues on word formation. the case of latin circum. *The Linguistic Review*, 35(1):121–159, 2018. This paper assumes that the main predicate in a clause whose main verb has *circum-* preverb is the preverb instead of the verb root, which is assumed to be adjoined to the light verb; this is not necessary for the conclusions of the paper; also, saying that the verb root is first adjoined to the light verb and then the complex is merged with the locative construction is not very different from saying that a “locationality” functional head selects the verb root as its complement and the locative construction as its specifier.
- Isabel Oltra-Massuet and Karlos Arregi. Stress-by-structure in spanish. *Linguistic Inquiry*, 36(1):43–84, 2005.
- Maria Isabel Oltra Massuet. *On the notion of theme vowel: A new approach to Catalan verbal morphology*. PhD thesis, Massachusetts Institute of Technology, 1999.
- Renato Oniga. *Latin: a linguistic introduction*. OUP Oxford, 2014.
- Harm Pinkster. *The Oxford Latin Syntax: Volume 1: The Simple Clause*. Oxford University Press, 08 2015. ISBN 9780199283613. doi: 10.1093/acprof:oso/9780199283613.003.0004.
- Harm Pinkster. *The Oxford Latin syntax: Volume II: The complex sentence and discourse*. Oxford University Press, 2021.
- Peter L Smith. *Greek and Latin roots: part I-Latin*. University of Victoria, 2016.
- Peter Svenonius. Spatial P in English. In *Mapping Spatial PPs: The Cartography of Syntactic Structures, Volume 6*. Oxford University Press, 08 2010. ISBN 9780195393675. doi: 10.1093/acprof:oso/9780195393675.003.0004.