Note on Latin Grammar

Jinyuan Wu

October 2, 2022

Chapter 1

Overview

1.1 Historical notes

This note is about Classical Latin and Ecclesiastical Latin. That's to say Old Latin, vulgar Latin (with prototypes of Romance articles), etc. are not discussed.

1.2 Phonology and the writing system

1.3 Parts of speech

1.3.1 Classification

Latin word classes can be defined easily via morphology, and these classes prove to have morphosyntactic significance. Traditionally speaking, word classes with none or poor morphology are called **particles**, and non-particle words 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 particles include **prepositions**, **adverbs**, **interjections**, and **conjunctions**. The adverb class and the preposition class have a large overlap: often a preposition has an intransitive counterpart, which is similar to a prototypical adverb. Conjunctions may be seen as "prepositions for clauses". The functions and etymologies of particles are highly diverse.

Latin nouns, verbs, and adjectives are all open categories. They are able to head constituents, and so are correlatives (though correlatives can be listed in the grammar). The preposition class is closed and is a part of the grammar, just like conjunctions. However, conjunctions are purely functional, while certain prepositions may be argued to head attributive expressions: 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. This is also the case of adverbs: some adverbs seem to be periphrastic markers of TAME categories and therefore may be considered as a part of the grammar, while others seem to carry "real" meanings. Fig. 1.1 is a visualization of the classification of Latin word classes.

> **Box 1.1: Lexical and function classes**
>
> In § 5.1.2 in this note, by words with "real category labels", I mean words that have "real" meanings and serve as lexical heads of constituents (i.e. being surrounded by function words and dependents). Certain adverbs and prepositions have "real category labels", and they appear at the left side in Fig. 1.1. Prepositions can be enumerated and therefore are considered as a part of the grammar, so they are always at the lower side in Fig. 1.1. Other adverbs and prepositions are light in their semantic and are purely functional, so they appear in the southeast corner of Fig. 1.1.
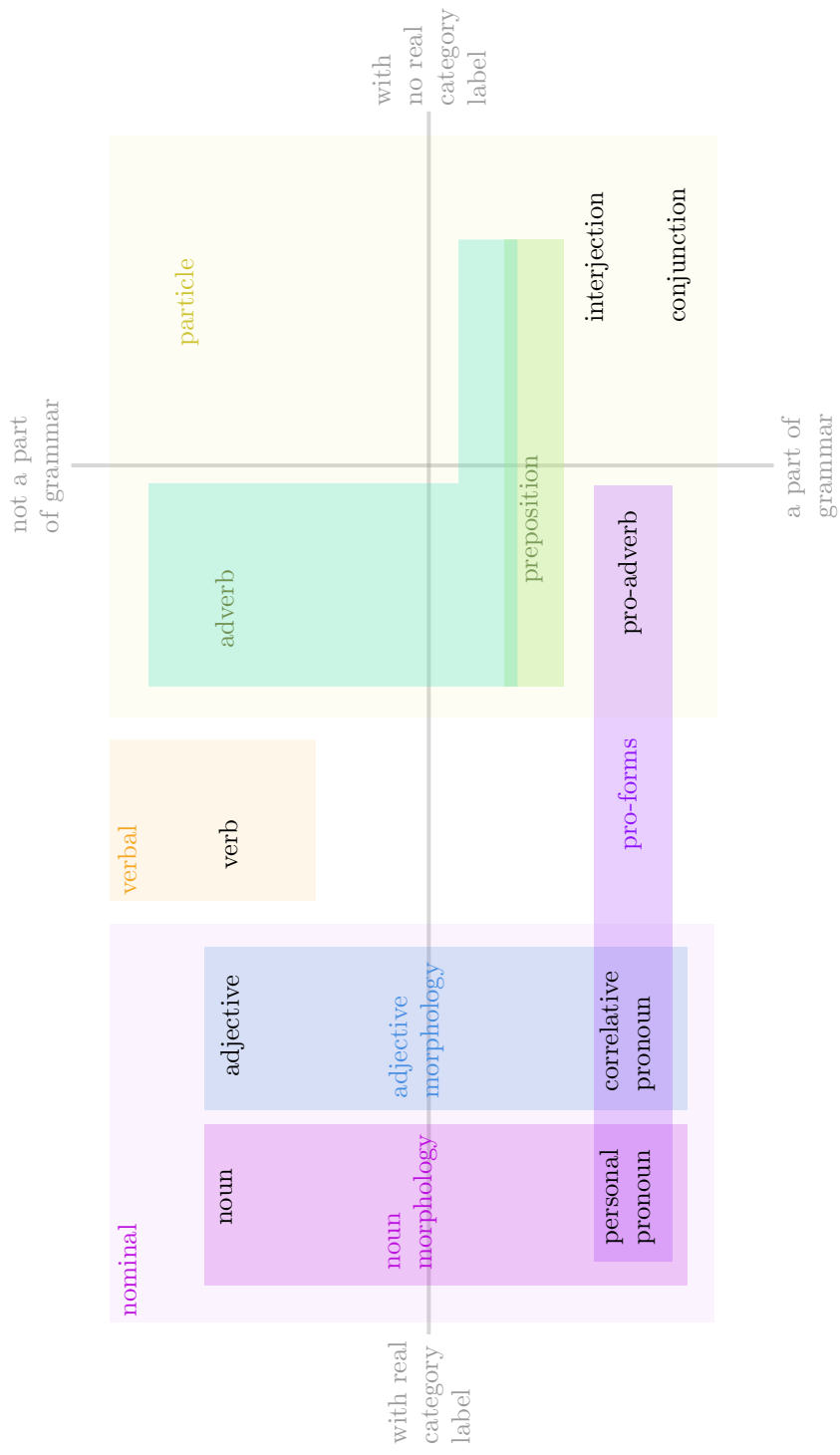
Figure 1.1: Latin word classes

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.

1.4   Morphology

Latin has rich morphology, which enables a rather free – but still not completely arbitrary – constituent order. Latin has a clear inflection-derivation distinction. Despite its richness, Latin derivation is largely historical, with meanings of derived forms having shifted and no longer regularly inferrable. Latin inflection is always suffixal, while derivation is predominantly prefixal. Concatenative morphology (affixation and compounding) is prominent but isn't the only morphological device: the following non-concatenative mechanisms are all attested:

- *Reduplication*: formation of the perfect stem (TODO: ref)

- *Subtraction*: dropping of first-conjugation stem-final vowel (§ **??**).

- *Infixation*: TODO: ref 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.

These mechanisms, however, are largely historical, just like their concatenative counterparts.

> **Box 1.2: Constituency deemphasized in Latin grammar**
>
> The largely free constituent order means description of Latin grammar is mostly dependency-relation based or BLT-based, because surface-based constituents other than NPs and clauses are hard to define. Still, generative (constituency-based, though the introduction of movements and the structure of Cinque hierarchy gives it certain flavor of dependency grammars) approaches exist for Latin constituent order. There is evidence suggesting Latin is configurational, i.e. has phrase structures (Danckaert, 2017). This is probably not surprising because even the most non-configurational languages show certain degree of configurationality (Niedzielski, 2017; Morris, 2018; Legate, 2002, among others). Then, *how* non-configurational Latin is is a question needing addressing. Is it closer to a typical non-configurational language, say Warlpiri, or is it closer to Japanese where we have more localized scrambling? I will address this question in TODO: ref , though unfortunately, we still does not have a very clear answer.

1.5   Noun phrases and nominal morphology

1.6   Verbal morphology and clause structure

Most clausal grammatical categories are marked on the verbal morphology. Sometimes a grammatical category is there but is not reflected in the morphology. For example, in English we have infinitive clauses, but strictly speaking, there is no such thing as "infinitive verb": the head verb of an infinitive clause has exactly the same form of a non-third person singular present tense verb. This is not the case in Latin. For example, the head verb of a infinitive clause in Latin indeed has a separate position in the paradigm. Thus, grammatical categories of the clause are listed in this section.

1.6.1   The finite paradigm

1.6.1.1   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. Whether passivization happens is recorded by the category of **voice**. A verb (and hence the clause headed by it) is therefore either in **active voice**, or in **passive voice**.

1.6.1.2 TAME categories

Latin has fused tense and aspect: the composition of three tense values and three aspect values gives nine options, but in Latin, there are only six morphologically distinguished options, as is shown in Table 1.1. When people talk about **tense** in Latin (and in many other Indo-European languages), they are often taking about things like the six options, instead of the past/present/future system.

Table 1.1: Latin tense and aspect

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

Similar fusion between categories is shown in the category of **mood**. 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.

1.6.1.3 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).

1.6.1.4 Compatability of categories

There are compatibility problems of these 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 TAME 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 is shown in Fig. 1.2. Here mood and tense are realized in one morpheme, and voice, person and number are realized in one morpheme. The exact realization is divided into four conjugation classes, and the details are too complex to show here. 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.
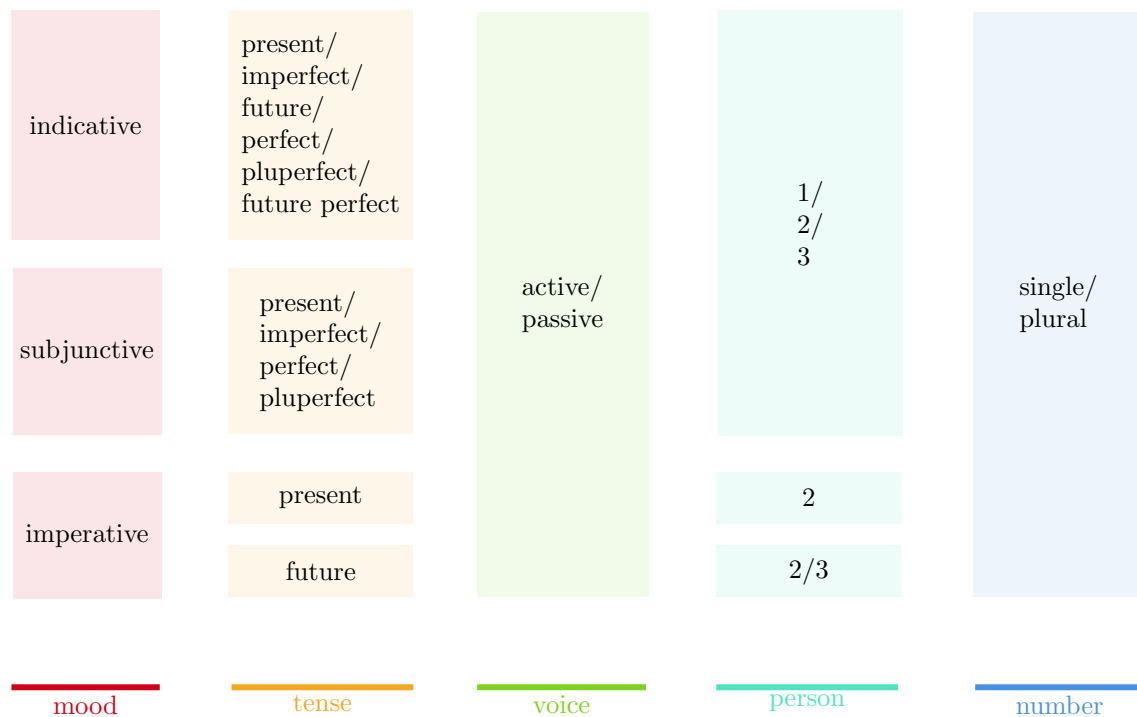
mood	tense	voice	person	number
indicative	present/ imperfect/ future/ perfect/ pluperfect/ future perfect	active/ passive	1/ 2/ 3	single/ plural
subjunctive	present/ imperfect/ perfect/ pluperfect			
imperative	present		2	
	future		2/3	

Figure 1.2: Categories in the finite paradigm

1.6.2 Non-finite forms

1.6.3 Core, oblique, and peripheral arguments

1.6.3.1 Alignment

Latin is a clear nominative-accusative language. Similar to what is documented in CGEL, Latin core arguments are coded as subject, object(s), and copular complements at the level of alignment. They can be distinguished by the semantic roles, case marking, possible contents, and transformational properties (§ 5.1). In

> **Box 1.6: The term *complement***
>
> In CGEL the term *complement* means the grammatical status of anything that originates from the argument structure of the main verb. Thus the subject, several kinds of objects, the copular complement (CGEL calls it *predicative complement*) are all complements,

and they are also function labels (roughly corresponding to "SpecTP" or "what is in vP and receives the object case from a high light verb"). In traditional Latin grammar *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 CGEL *predicative complement*. Also, the term *complement-taking verb*, despite being confusing, is used to denote a verb that take a complement clause as one of its arguments.

The mapping from the argument structure of a verb to clausal complement types is called alignment.

1.6.3.2   Peripheral arguments

There is no serial verb constructions in Latin (§ 1.6.4), and thus semantic functions like location or instrument are always realized by typical peripheral arguments attached to the core argument structure. These peripheral argument positions sometimes can be filled by adverbs, which also reveals an origin of adverbs.

1.6.4   Clause combining

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

1.7   Constituent order

Chapter 2

Phonology and the writing system

Chapter 3

Nominal morphology

Chapter 4

Verb morphology

4.1 The structure of Latin verbs

The traditional analysis of the structure of a Latin verb can be roughly represented by Fig. 4.1. Derivation in Latin is predominantly preverbal, and hence the conjugation is mostly about the final lexical morpheme in the verb stem. The stem-final vowel is sometimes considered as a part of the stem, and sometimes as a part of the verb ending. It is the residue of the PIE stem suffix, which is after the core stem and before the conjugation ending, and still has morphosyntactic alternation as well as phonological ones in Latin. The uncontroversial components of the verb ending include the **tense and mood marker**, and the person, number and voice marker (here after the **personal ending**, following the terminology in Allen and Greenough (1903, § 165)). But the tense and mood marker is influenced by the personal ending: the same tense and mood may be marked by one marker under one person and one number but by another under another case. The inverse is also true. There is also phonological interaction between the two components of the ending.
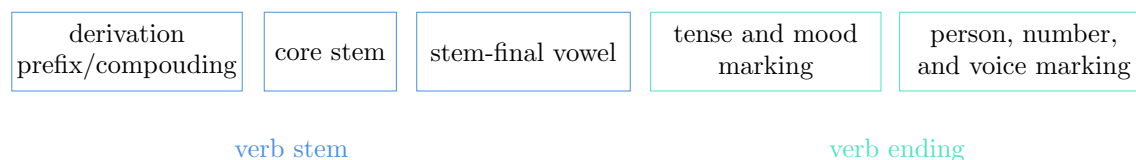
derivation prefix/compouding	core stem	stem-final vowel	tense and mood marking	person, number, and voice marking

<div align="center">verb stem   verb ending</div>

<div align="center">Figure 4.1: The template of Latin verbs</div>

TODO: explain the *-mus* and *-imus* alternation

> **Box 4.1: About the number of verb forms**
>
> Different people use the term *verb forms* – and count them – in different ways. The most generous – and the most syntactically relevant – way is to view the realization of every possible CP-TP-vP projection as a form of the main verb – the verb root at the core of the CP-TP-vP domains. This results in a paradigm in traditional grammar. The problem with this approach is sometimes two cells in the paradigm are always identical, so recognizing them as two morphological forms is weird. (Also, this is not a good idea when dealing with languages like Japanese.) A stingy linguist may then stipulate that conjugation forms are literally about *forms*, and thus there is no such thing as "the subjunctive form" of English verbs, because in subject *clauses*, the main verb always has the same form as the infinitive.
>
> The generous approach fortunately works in Latin because Latin is morphologically rich. The idea of the stingy linguist may lead one to reject the notion of supine in Latin grammar, but since sometimes a verb lacks TODO: argumentation for a separate supine form, for the same reason the infinitive is recognized as a form independent from the "default form" in English in CGEL, the status of supine as a separate form is recognized in this note.

The analysis of conjugation forms of the verb, theoretically speaking, is more about vocabulary insertion and readjustment rules, instead of the syntax proper. This is an instance of the *separation principle*: morphophonological features can be irrelevant to morphosyntactic features (Embick, 2000).

Box 4.2: Distributed Morphology of deponent verbs

Embick (2000) analyzes deponent verbs as roots carrying a passive feature themselves. When functional heads higher than the roots are realized, the passive feature – which may come from the root or from the passive light verb – guides the realization of the person and number categories.

4.2 Formation of stems

4.2.1 The three verb stems

Box 4.3: About the concept of stem

Prototypically, the verb conjugation in a language is described by a series of morphological devices that take *the* verb stem as the input, and give conjugated verb forms as the final product. This is indeed the case for Latin nouns (§ **??**) and for English regular verbs: the infinitive form is taken in, and third-person singular *-s*, past tense *-ed*, past participle *-ed*, and the gerund-participle *-ing* are attached according to the syntactic environment. Sometimes the process is a little more irregular but not that irregular: *several* stems can be identified, each of which is fed into different morphosyntactic machines. In other words, we have irregular stem alternation. Again, for English irregular verbs, there are three stems: the infinitive stem (e.g. *go*), the preterite stem (e.g. *went*) and the past participle stem (e.g. *gone*). The step to feed stems into morphosyntactic machine is irregular, but everything else is regular: irregular, in this case, does appear, but it appear *regularly*: it only appears in certain parts.

This phenomenon – that a verb has more than one stem, i.e. irregular stem alternation – is frequent cross-linguistically (Jacques 2021 § 12.2, Forker 2020 § 11.2, among others). Usually, certain correlation between the stem varieties can still be recognized, and verbs can be grouped accordingly, which, if the linguist truly will, can be (though tediously) summarized as more fine-grained conjugation classes. This is also the case for Latin. The irregularity of stem alternation is so prevalent that if the conjugation paradigm of a verb can be described with a few stems, the verb is deemed as regular, despite the fact that such verbs are obviously irregular by the standard of English.

The notion of **stems** isn't really essential in the description of morphosyntactic: it can well be modeled by environment-dependent vocabulary insertion rules and/or post-syntactic operations. When certain correlations can still be built between so-called suppletive forms, what happens may be analyzed as in Embick and Halle (2005), where certain stems receive morphophonological readjustment (according to the aforementioned hyper fine-grained conjugation subclasses). Thus, it's not true suppletion: it's just a corner case of non-concatenative morphology. When these readjustment rules are fossilized, suppletion – like the English *good/better* – may just be the result of conditional insertion, as is outlined in Siddiqi (2009).

A general tendency about suppletion is truly suppletive verbs are usually light verbs (in the surface-oriented sense), with meanings like 'do', 'come', etc. This may come from the fact that conditional realization of the root is somehow "heavy". When we do away with this, e.e. under Embick and Halle (2005)'s assumption, unrestricted suppletion can only be the result of vocabulary insertion rules of functional heads; certain degree of suppletion can still be realized by readjustment rules, which however are restricted in their computational capacity. Thus, real lexical verbs are highly unlikely to have truly irregular suppletion; if a verb is truly suppletive, then it's highly likely to be the spellout of vP functional heads.

Latin shows prevalent and not completely predictable stem alternation. All forms mentioned in § 1.6 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, 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. 4.2.

4.2.2 Rules of stem alternation

4.3 The four regular conjugations

4.4 Deponent verbs

4.5 Periphrastic conjugations

4.6 Irregular verbs
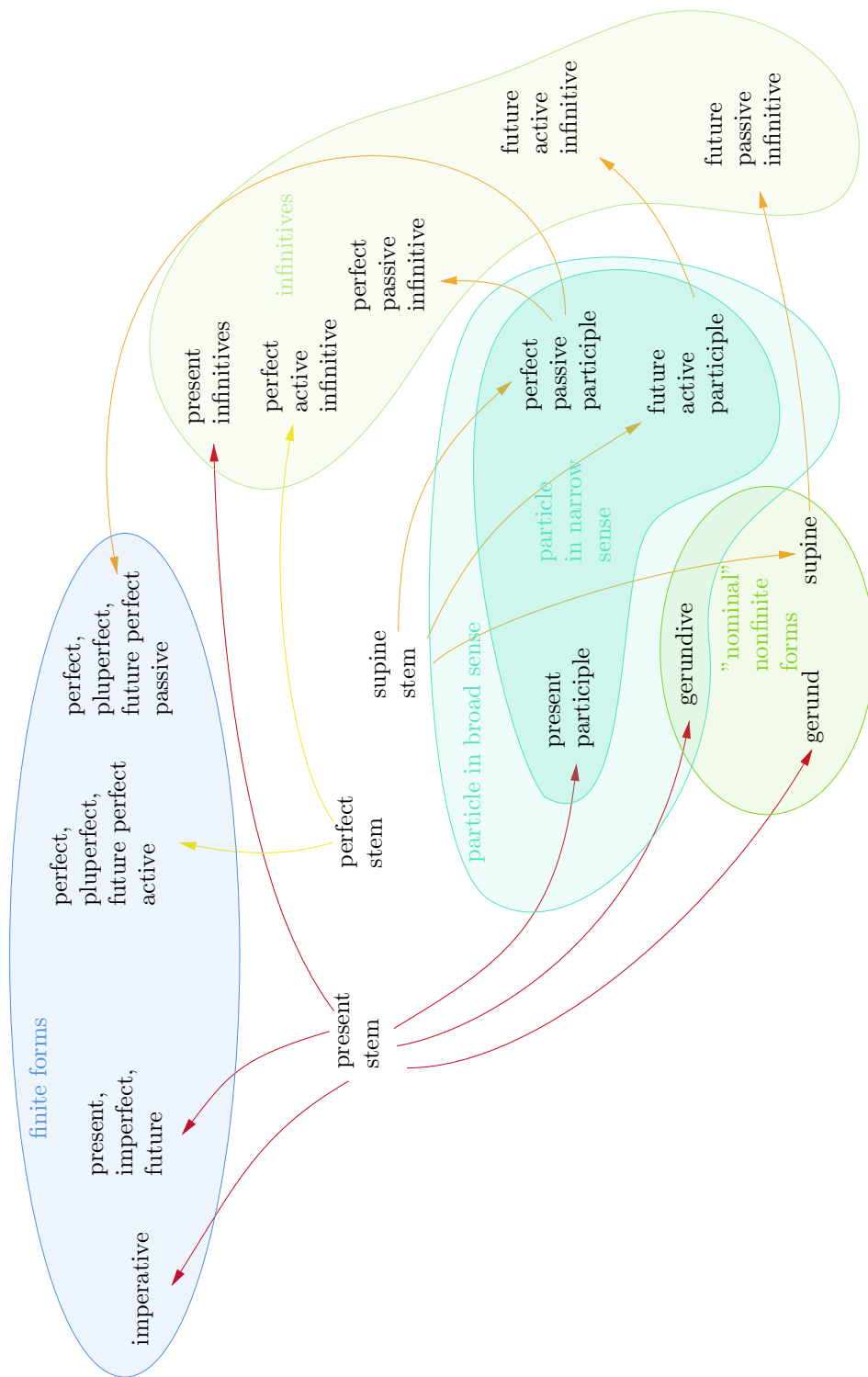
4.7 Verb derivation

Figure 4.2: How to get all conjugation forms from the three stems

Chapter 5

Arguments and verb valency classes

5.1 Overview of marking of core arguments

5.1.1 The subject

5.1.2 The direct object

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

- *Coding of semantic role*: in a AO-type argument structure, the direct object is the O argument, i.e. the most patientive argument.

- *Case marking*: direct objects are always accusative – but not all accusative arguments are direct objects.

- *Passivization*: in a passive clause, the subject corresponds to

-

5.1.3 The indirect 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.

- *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 5.1: 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 CGEL. 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, and in Latin datives never have prepositions – the G argument is identified with the E argument, and the T argument is identified with the O argument. In other words, in Latin, there is only pattern (1) in my notes about CGEL.
>
> The secondary object, therefore, is like (2) in my notes about CGEL. 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.
>
> ---
> [a]In 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.

5.1.4   Predicative complements

Latin also has predicative complements. A predicative 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 is roughly about the subject and agrees with it, and the accusative predicate is roughly about the direct object and agrees with it. In passivization of the direct object, the accusative predicate becomes the nominative predicate.

> **Box 5.2: How complement types are discussed**
>
> The traditional practice of Latin grammar research is to classify clausal complement and adjunct types according to their case marking. This strategy is also found in modern grammars. Some introduce clausal complement types just in chapters about case marking (Jacques, 2021, chap. 8), while other grammars, despiting giving a brief description of the context of case marking, spare some time to discuss complement types in the chapters about valency and clause structure (Forker, 2020, § 3.4, chap. 19, chap. 22). From a Tree-adjoining grammar (TAG) perspective, the two extremes are different in how they treat function labels: in the former, the function label of a construction appears together with its category label on the root node, while in the latter, the function is described separately from the form of what fills that position. The former is more bottom-up, while the latter is more top-down. The choice between the two, however, is usually language-dependent: grammars for analytic languages, of course, have to lean even further to the "complement type as clause slot" extreme and away from the "complement type as case-form context" extreme. Allen and Greenough (1903) uses a hybrid method: the discussion about case marking (§ 39) is separated from the discussion about complement and adjunct types (§ 338), so the top-down approach seems to be adopted, but the latter is still arranged in terms of case. This arises both from the distinct features of Latin and the intended readers: the relation between complement types and cases is regular enough in Latin, and what is most important for Latinists is to understand, at least sketchily, ancient writings, so a parsing-oriented grammar is much handier.

5.2 Peripheral arguments

Beside the subject, various types of objects, and predicative complements, there are more peripheral clause dependents (or "adjuncts" in the terminology of Huddleston and Pullum (2002)), like purpose, direction, location, etc. Besides clauses and NPs (with or without prepositions), their categories also include adverb phrase (AdvP)s, which is an etymological origin of AdvPs.

Latin peripheral arguments do not necessarily have prepositions. This fact, together with the fact that Latin is highly free-ordered, highly *pro*-drop, means criteria used to tell core arguments from peripheral arguments, like the criteria of category, position, and argumenthood in Huddleston and Pullum (2002, § 4.1.2) all fail to work, and since the problem is how to tell adjuncts from adjunct-like complements, the criterion about role also fails. Latin does not have systematic way to replace the core predicate (i.e. without adjuncts) by an anaphora, and that criterion does not work, either. The remaining criteria are about selection, licensing, and obligatoriness.

5.3 Prototypical intransitive verbs

Chapter 6

Examples of texts

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.

Lieven Danckaert. *The development of Latin clause structure: A study of the extended verb phrase*, volume 24. Oxford University Press, 2017.

David Embick. Features, syntax, and categories in the latin perfect. *Linguistic inquiry*, 31(2): 185–230, 2000.

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.

Diana Forker. *A grammar of Sanzhi Dargwa*. Language Science Press, 2020.

Dianne Friesen. *A grammar of Moloko*. Language Science Press, 2017.

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.

Julie Anne Legate. *Warlpiri: theoretical implications*. PhD thesis, Massachusetts Institute of Technology, 2002.

Paul A Morris. Evidence of a configurational structure in meskwaki. *Proceedings of the Linguistic Society of America*, 3(1):64–1, 2018.

Patrick M Niedzielski. Clausal syntax in the sumerian language, 2017.

Daniel Siddiqi. *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*, volume 138. John Benjamins Publishing, 2009.