

Notes about Latin grammar

Jinyuan Wu

June 28, 2022

1 Theoretical orientation and descriptive terms

1.1 Tree diagrams

The theoretical orientation can be found in [my notes about CGEL](#), especially § 1. There is one barrier to adopt a surface-oriented binary-branching tree diagram analysis, since Latin has a largely arbitrary constituent order. Many works have been attributed to this topic ([Danckaert 2011](#); [Devine and Stephens 2006](#), among others), but whether they are enough to account for the variations observed in Latin is questioned ([Spevak, 2007](#)), and all of them (unsurprisingly) involve lots of movements, which are not acceptable for a surface-oriented analysis – indeed, translation between contemporary Minimalist theory and the more surface-oriented typologically oriented theory (the so-called Basic Linguistic Theory (BLT)) is not trivial and needs special attention ([Broekhuis, 2010](#)). Note that, however, that argumentation pertaining to c-command relations, e.g. binding, has been studied in the generative literature ([Mateu and Oniga, 2017](#)), and hence for the surface-oriented descriptive target in this note, positing a CGEL-like surface-oriented framework plus a (possibly pragmatic) scrambling operation is still a good idea: the former visualizes relevant grammatical relations and categories, while the latter respects the free-order property.

1.2 Distinction between form and function

Traditional grammar often confuses syntactic form and function.¹ A clause able to fill an argument slot is considered as nominal, which, despite having overlapping function with noun phrase (NP)s, may have nothing in common with NPs in form, if it is finite, and have limited similarity with NPs if it is nonfinite. The answer to “what it (a specific construction) is” in traditional grammar is therefore often function in form’s disguise or the inverse. What is a content clause? It is a “nominal clause” (read “a clause able to fill an argument position”, where “nominal” means “argument”). This is an example of terms about form actually denoting a possible function. What is a participle in English? It is a “nonfinite verb form that happens to coincide with the gerund but somehow is different” (read: “a predicator of a nonfinite relative clause with *-ing* ending”). This is again an example of terms about form actually denoting the specific function of the construction in a given environment (we may say “an *-ing* verb like an adjective is a participle”, and here “like an adjective” seems to be about form but actually specifies the syntactic environment i.e. function). Calling a content clause a complement clause is the inverse: it uses the function (being a complement) as the name of the form of the construction.

Practically, this approach works well, especially in a comparative grammar context, because of course function constraints form and for a large number of constructions, a one-to-one mapping can be established, and how one form fills two function slots may be formulated as “two forms coincide”, while how one function can be realized by two forms may be formulated as “one form has two subtypes”.

In English, for example, the term *complement clause* is not a proper term about form, because so-called *complement clause* can fill temporal and locational adjunct positions, but for others, complement clauses and adjunct clauses have significant morphosyntactic distinction, and hence the term *complement clause*, though a term for function, also serves well as a term for form. Now a comparative grammarian has good reason to use the term *complement clause* to denote content clauses in

¹Note that here “function” *does not* mean “pragmatic function” as in functionalism.

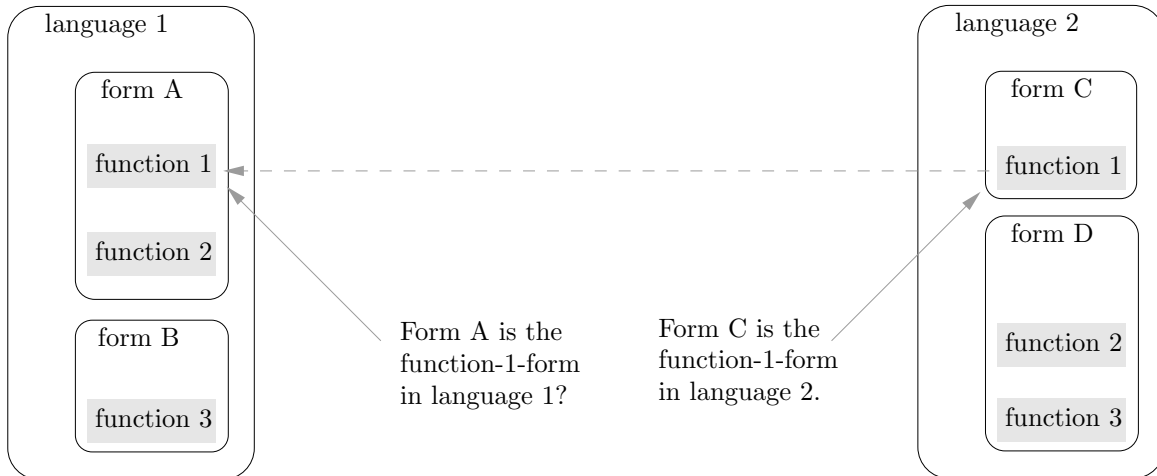


Figure 1: Why confusion form and function in practice does not create much inconvenience: only form C in language 2 is able to fill function slot 1 means in language 2, we can call form C as the function-1-form, and hence we can call form A – which realizes function 1 – as the function-1-form as well. The fact that form A can also fill a function 2 slot, then, is formulated as “function-2-form and function-1-form coincide”. Swapping “form” and “function”, the diagram still holds.

a language like English where complement clauses serve as adjuncts, because this term works in a language with morphosyntactic distinction between complement clause and adjunct clauses anyway, and hence by the definition by prototype approach, it seems reasonable to also call clauses that satisfy as arguments *complement clause* in *any* language. The fact that content clauses can be both complements and adjuncts, then, may be formulated as “a complement clause and the corresponding adjunct clause coincide”. The scheme is shown in Fig. 1.

The position of this note is to respect the traditional grammar terms while point out their true meanings.

1.3 Terminology

1.3.1 Lexical and phrasal categories

noun phrase, adjective phrase (AdjP), adverb phrase (AdvP), preposition phrase (PP)

The term *nominal* sometimes means “noun-like in function”, i.e. able to fill argument slots, and sometimes means “noun-like in form”, i.e. declines. The two meanings are all frequent in existing works. Examples of the former include “infinitives are nominal”, while examples of the latter include “the adjective is a nominal category” (De Kuyper, 1991). The term *adjectival* has similar dual meanings. Something is *adjectival* because it is able to fill a predicative complement slot, or because it has declension.

1.3.2 Morphology

Latin has a clear inflection-derivation distinction. I follow the standard practice and call verbal inflection **conjugation** and nominal inflection **declension**.

1.3.3 Grammatical categories

- *Tense*: the term has two meanings: one includes aspectual information, the other does not (§ 2.3.4).
- *Mood*: the term means

2 Morphosyntactic overview and typological information

In the following sections, I discuss grammatical systems contained in all levels of Latin grammar, their realizations and typological information.

2.1 Note about section organization

I intentionally do not employ the section division an experienced grammar writer will use to describe a language with rich morphology, for example Jacques (2021, chap. 2), to present Latin in a more “content-free” way, dealing with only abstract (and typologically comparative) aspects of the grammar.

It can be seen that describing a language purely in terms of “what happens in what construction” as in following sections is usually cumbersome. For example, since argument alignment happens all within the clause, it has to be placed under the subsection about clause structure and hence it is a subsubsection, and then there is no indexing space in the title for further division of the topic: the section about alignment is already a subsubsection, and there is no sub-sub-sub-section for agreement or the subject position. § 2.3.1 therefore becomes too long and structureless, but it has to be so to fit the top-down approach used here.

Should the organization of Jacques (2021, chap. 2) be used and hence constructions be described in a bottom-up way, with grammatical categories being introduced in discussions about smallest constructions that show them, since § 2.3.1 is obviously too long, and deserves at least a subsection, it would be renamed into “core and oblique arguments” and promoted to a subsection, with topics like agreement and subject-predicate relation occupy subsubsections. Contents in § 2.3.1 and the rest of § 2.3, therefore, would be split into two subsections, which are § 2.4 and § 2.5 in Jacques (2021). Sections from § 2.3.2 to § 2.3.6 are too short. They can be all inserted into § 2.3.7. § 2.3.7, in turn, can be promoted to a subsection and divided into sections about methodological issues to distinguish inflection forms, conjugation in finite clauses, conjugation in infinite clauses, etc.

2.2 Parts of speech

2.2.1 Morphological classification of words

Traditionally, Latin lexical categories are distinguished by morphology.

What do not have rich inflection are **particles**, which, according to the prototypical function, can be divided into adverbs, prepositions, conjunctions, and interjections. Adverbs can also be distinguished morphologically, because they appear in the comparative construction and have comparative and superlative forms (§ 2.8), so they still have certain inflectional forms, though much less than non-particles.

What **conjugate** (§ 2.3.7) are verbs. Early grammarians thought what **decline** (§ 2.4.10) are all nouns, but then people realized more detailed classification can be made. Purely by morphology, nouns and adjectives can be tear apart. In the category of nouns, the category of **personal pronouns** is distinguished, for a personal pronoun has explicitly marked person feature. Personal pronouns do not code gender morphologically. Contrary to them, **correlatives** are always third-person but have gender morphologically coded, and hence they form a subcategory under the morphological adjective category. However, correlatives are referential and by no means purely adjectival, so correlatives and personal pronouns form a **pronoun** category (§ 2.5), which breaks the boundary of morphological classification.

2.2.2 Phrasal categories

Nouns are heads of NPs.² Adjectives are heads of AdjPs. We also have AdvPs. It is often said that argument positions (§ 2.3.1) are filled by *nouns*, but in CGEL’s terms, actually they are filled

²Or in generative terms, “project into”, though in modern generative syntax, the term *head* has a large shift in its meaning, and the contemporary formulation of “a noun projects into a NP” is “a noun is the deepest lexical stem of a DP, which is the largest nominal projection”.

by NPs: a bare noun may be seen as a single-branched NP.³ The CGEL notation avoids calling NPs nouns, and hence we can use the two terms in different contexts: when we talk about *nouns*, morphological issues may be talked about (§ 2.4.10), while when we talk about *NPs*, NP dependents are talked about (§ 2.4). Of course, grammatical relations involved in nominal morphology and syntax are intermingled and that is why § 2.4.10 is placed in § 2.4. There is no well-recognized distinction between morphology and syntax: what is syntax is what older generations of Latinists decided as syntax, i.e. things do not cooperate that closely with phonology. From a descriptive perspective, the main reason to contrast nouns and NPs is to separate grammatical relations to different layers to make them well-organized, in the same way CGEL contrast nominals and NPs.

2.2.3 Note about the function-oriented conception of nouns and adjectives

Traditional grammar also consider what fill argument slots as nouns. In this way, content clauses (term in CGEL) or complement clauses (function-oriented term actually denoting form) or substantive clauses (term in Allen and Greenough (1903)) are nouns. They are often neither nouns nor NPs in modern terms. The only sensible interpretation of “substantive clauses being noun” is that substantive clauses fill argument slots. Infinitives, for example, have nothing similar to nominal morphology (compare § 2.3.5 and § 2.4.10). They can fill the subject position, though. Compared to infinitives, gerunds are more like nouns in that they did not take objects in Classical Latin, but they did so in older and later forms of Latin.

Similarly, relative clauses may be regarded as adjectives. They are not. Nor are they AdjPs. Relative clauses differ from AdjPs in form. Relative clauses differ from adjectives in form and function – adjectives appear in the comparative construction (§ 2.8), but relative clauses do not.

2.2.4 Relation between the four particle categories

Categories are often correlated, both historically and synchronously. Here I briefly review the relation between adverbs, prepositions, conjunctions, and interjections.

Adverbs, prepositions, and conjunctions fill so-called peripheral arguments (§ 2.3.1). This hints at how adverbs were historically created: derivation from adjectives (“in . . .’s manner”), fossilization NPs, PPs, etc. Actually, in Forker (2020), adverbs are indeed placed in the nominal category part. Just like the case in English, lots of Latin adverbs themselves are prepositions. Indeed, adverbs are intransitive particles, while prepositions are transitive particles that take NPs as complements, and conjunctions are transitive particles that take substantive clauses as complements. In a synchronous perspective, since a verb can be both transitive and intransitive, so can a preposition. In a diachronic perspective, many Latin prepositions are historically adverbs, and because of semantic reasons (e.g. to describe an action towards a target, an adverb describing “toward some target” and a NP about what the target is routinely occurred together), which was recognized as an independent dependency relation (i.e. not clearly reflected in the surface form, maybe involving movement of the NP out of the underlying PP, maybe by other mechanisms) by younger generations, which further developed into a direct dependency relation and the adverb became a typical preposition.

Negators (§ 2.3.3) are traditionally also recognized as adverbs, but they do not fill peripheral argument positions, and hence

2.2.5 Nouns and adjectives

2.3 Clause structure

2.3.1 Argument structure and alignment

Latin is a typical nominative-accusative language, both syntactically and morphologically. A **subject** can be identified for all clauses, though it is frequently omitted. Grammatical behaviors re-

³Unfortunately, even highly experienced field linguists may occasionally use the term *noun* to denote NPs. Friesen (2017) is a well-organized grammar, striking a balance between the top-down approach and the bottom-up approach, but in § 5.4, we can still see words like “nouns as modifiers”, but a closer look (discussion in § 5.4.1) reveals the section is actually talking about NPs as modifiers.

stricted to the subject are summarized in the following list:

- *Coding of semantic role*: in an active clause, the subject is always the most agentive argument. In a passive clause, the subject always corresponds to
- *Case marking* (§ 2.4.3): subjects are always nominative for finite clauses. Nonfinite clauses may be argued to be subjectless in the surface form (a reasonable claim, since they have deficient TP layers, and hence it is possible that no canonical subject position exists), but in accusative-infinitive constructions, the accusative may be seen as the non-canonical subject of the infinitive.
- *Agreement* (§ 2.3.7): the number and person features on the subject leave marking on the predicator. Latin does not have verbal agreement with arguments other than the agreement with the subject.
- *Category*: a subject is a NP (§ 2.4) or a complement clause (§ 2.6), usually an infinitive but never a gerund (§ 2.3.5).

A minimal clause – without any adjuncts, negation, etc. – can therefore be analyzed as a subject plus a predicate, with the predicate headed by the predicator (which can only be filled by a verb in Latin) and its internal arguments (the subject is the external argument).

In the predicate, sometimes a **direct object** position can be identified. If there is a direct object, then the predicator is called a **transitive verb**. If there is none, it is **intransitive**. 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* (§ 2.4.3): direct objects are always accusative – but not all accusative arguments are direct objects.
- *Passivization*: in a passive clause, the subject corresponds to
-

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.

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 is distinguished by the following grammatical properties:

- *Coding of semantic role*: in an AGT-type argument structure (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 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 English⁴, 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.

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 a 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.

Beside the subject, various types of objects, and predicative complements, there are more peripheral clause dependents, like purpose, direction, location, etc. They may be analyzed as adjuncts or adjunct-like complements in CGEL (§ 4.7 in [my notes about CGEL](#)). Besides clauses and NPs (with or without prepositions), their categories also include AdvPs. The fact that NPs in certain cases, with or without prepositions, AdvPs, and conjunctions (which introduce substantive clauses) have largely shared function means they are often etymologically correlated.

This note follows the traditional practice of Latin grammar research, and classify clausal complement and adjunct types according to their case marking (§ 5).

Since Latin is highly free-ordered, highly *pro*-drop, and peripheral arguments do not necessarily have prepositions (§ 2.4.12), criteria like category, position, and argumenthood in CGEL 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. Possible complementation patterns and their correlation with the semantics are introduced in § 8 in this note.

2.3.2 Voice

Latin only has active (the canonical one) and passive voices. Passivization can be done morphologically in all circumstances except the case of the passive perfective, which is realized resembling the English passive, i.e. via a copula and the perfect passive participle.

2.3.3 Polarity

Unlike languages like Japanese, polarity is not marked morphologically in Latin. Latin realizes negation in a largely regular syntactic way: the negation operator *nōn* can be placed into the clause.

The exact position of *nōn* is kind of nuanced. Latin is largely free-ordered, and a good approximation is to assume the negator can appear anywhere it wants. Cross-linguistic investigation of negation, however, suggests a fine-grained hierarchy in which several functional projections can be

⁴In 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.

Table 1: Latin tense and aspect

	past	present	future
imperfect	imperfect	present	future
simple	perfect		
perfect	pluperfect	perfect	future perfect

negated, but not others, and this hierarchy obviously imposes some restrictions on the linear constituent order and also relates the constituent order with semantics. It is *possible* that a language *completely* does not show any restriction on the position of the negator, and nor does its position has any semantic implication, since by scrambling and PF dislocation, it is easy to shuffle the constituent order, but it would be *surprising* should this happen, because completely throwing away surface constituent order as a handy way to realize negation scope is somehow uneconomical. Latin allegedly does show some constituent order-related constraints on negation (Tierney, 2018). The details are to be discussed later.

2.3.4 Tense, Aspect, Mood (TAM)

In traditional grammars, Latin has six tenses: present, perfect, future, imperfect, pluperfect, future-perfect. These so-called tenses are recognized by morphology, and are better considered as labels of verb conjugation. Modern analysis separates aspect from tense, and aspect itself is a catch-all term for several systems (BLT § 3.15). Since the tense system is quite clear (past/present/future), the question is then what aspect system can be recognized in Latin. In the first glance, a imperfect/perfect aspect system seems to be answer, but this is misleading: it can be demonstrated that the so-called (present) perfect tense is able to describe an event simply happening in the past without any specification on its aspectual properties. The correct analysis, therefore, is Table 1, where the imperfect and simple aspects for the present and future tense are identified, and simple past is identified with present perfect for their obvious semantic resemblance. In the following discussion about nonfinite verb forms, the term *perfect* and the term *past* are often used interchangeably.

The imperfect/simple/perfect aspect system itself can be further divided. The details have to be postponed to sections devoted to this topic.

Latin has the realis/irrealis contrast in mood (or modality in BLT terms), which are much more frequently called as the indicative/subjunctive distinction. Subjunctive clauses are never in the future tense regardless of the aspect (i.e. never in future or future-perfect “tense” in traditional grammar’s sense), which is not typologically rare.

2.3.5 Finiteness

Nonfinite clauses are often said to be “nominalized” ones. This term is often misleading because it confuses form with function: argument positions are prototypically filled by NPs, and to this extent nonfinite clauses and NPs have largely overlapping *function*, but they do not necessarily (though possibly) have quite similar form: no accusative NP, for example, is able to appear as a complement of a noun, but nonfinite clauses may take accusative objects and may be modified by AdvPs. Nonfinite clauses may also have functions that NPs do not have. In languages like English, nonfinite clauses can be adjunct clauses by adding a word like *when*, while no such mechanism exists for NPs.

In Latin, besides the usual canonical clauses, Latin has several nonfinite clauses – much more than English. There are simple active, perfect passive and future active participles, a gerund, a gerundive which may be also called as the future passive participle, and present active, present passive, perfect active, perfect passive, future active, future passive infinitives. Strictly speaking, finiteness is a category of the clause, but since Latin is rich in morphology, finiteness may also be understood as a verb morphological category, and inversely, the inflection form of the predicator is also a classification criterion of the whole nonfinite clause, and we have notions like *infinitive clause*.

Note that this is not always the case: the perfect passive, future active and future passive infinitives are periphrastic.

Latin gerunds and participles have case feature marked by declension, so compared to English, saying Latin gerunds and participles are “nouns” are “adjectives” captures more information about their form. In Classical Latin (though not other varieties of Latin), gerunds and gerundives typically do not take objects, and hence their verbal properties vanish almost completely (except the fact that they are modified by AdvPs), since there are no longer “gerund clauses” or “gerundive clauses”. They are therefore better described as nouns and adjectives derived systematically from verbs in Classical Latin. This means some may disagree with the nonfinite verb status of gerunds and participles and place them into the chapters about *noun and adjective* morphology, and indeed, they are absent in some conjugation tables. This note still takes them as nonfinite verbs, because in historical stages other than Classical Latin (and even under limited circumstances in Classical Latin), gerunds and gerundives do take objects. This fact, plus the fact that they are modified by AdvPs, makes them much more verb-like.

Historically, the traditional grammar concept that nonfinite verbs are somehow nominalized originates from the fact that Latin gerunds and participles are really nominal and adjectival, which unfortunately distorts description of other languages where nonfinite verbs are not that nominal or adjectival.

Infinitives are not declinable, and do take objects. Some authors consider them also as nouns, since they fill argument slots (actually, in argument positions, infinitives are in complementary distribution with gerunds: a gerund is never nominative, and infinitives fill subject positions), but this time the aforementioned criticism towards the “nonfinite = nominalized” notion hits the target. Expectedly, infinitives never disappear from conjugation tables.

As can be inferred from their names, nonfinite clauses have (aspect-less) tense and voice categories, but with limited values.

2.3.6 Clause types and illocutionary forces

The term *clause type* is reserved for the syntactic marking of illocutionary force in CGEL. Latin has three syntactically coded clause types: declarative, interrogative and imperative. The interrogative clause does not have much syntactic markedness, and hence is usually not considered as a separate clause type in Latin grammars. The imperative is restricted in TAM features: the realis/irrealis distinction vanishes, and the aspect feature is unavailable.

2.3.7 Verb inflection

Traditional grammars place the labels *voice*, *tense*, *mood*, etc. on the verb, and hence the big, big conjugation table. From a modern perspective, this is not quite correct: labels like *voice*, *tense*, *mood* etc. are introduced by functional heads in the derivational process of the clause and should be considered, in an abstract and cross-linguistic way, as labels attached to the clause, not the predicator (usually filled by a verb). Specifically, from a cross-linguistic way, these labels are not to be considered as labels of verb conjugation: the number of conjugation forms and how to name them should be decided solely with syntactic distributional tests. If in clausal environment 1 and clausal environment 2, the verb has exactly the same appearance, then we say there is only one verbal conjugated form in these two environments, not two. It is thus in principle not appropriate to label verb conjugated forms according to clausal grammatical categories. However, from a more practical (and describe-a-language-in-its-own-terms) perspective, since in Latin (and languages with rich verbal morphology, like Japhug), these categories are marked on the verb and nowhere else, and there are never two forms identical when changing these clausal categories, naming verb conjugated forms with respect to the clausal environment is well justified. Actually, this is also how CGEL names non-finite clauses and verb forms in them (CGEL § 2.14).

It also follows from the modern linguistic view taken here that in Latin, the passive voice, perfect aspect forms of a lexical verb, strictly speaking, do not exist, since what really conjugates is the copula *sum* and not the lexical verb (the lexical verb is always the perfect participle, which agrees with the subject with number, gender and case), the latter being always in the perfect passive

participle form. Nonetheless, showing these so-called conjugated forms on the conjugation table is more user-friendly, and this is still the standard practice taken when studying and teaching Latin. These so-called syntactic “conjugated” forms are *periphrastic* conjugation, i.e. what is usually conveyed by pure morphology is marked by an idiom construction. The conjugation table including both morphological and periphrastic conjugation, therefore, is more a table about clause templates rather than the purely morphological paradigm.

So now I review categories marked by verb conjugation. They include person and number of the subject (§ 2.3.1), voice (§ 2.3.2), TAM features (§ 2.3.4), finiteness (§ 2.3.5), and the clause type (§ 2.3.6). In § 2.3.4 it has already been seen that tense and aspect are strongly intermingled in Latin, and hence in most Latin grammars, the two concepts are merged into one “tense” category when making the verb conjugation table. In § 2.3.6, it is introduced that Latin imperatives do not have contrast of mood, and the interrogative clause type does not have morphosyntactic marking besides the interrogative pronoun or adverb. Therefore, the clause type category and the mood category introduced in § 2.3.4 can be zipped into one extended mood category, values of which are indicative, subjunctive, and imperative.

Historically, since the traditional grammar was invented to study the Latin language, mood (in the narrow sense, i.e. modality) and clause type are unfortunately mixed together as one category, which causes lots of confusion. So is the case for tense and aspect.

So here is a summary of factors involved in the conjugation of a finite verb:

- *Person*: 1, 2, 3.
- *Number*: sg, pl.
- *Tense (in the broad sense)*: present, perfect, future, imperfect, pluperfect, future-perfect.
- *Mood (in the broad sense)*: indicative, subjunctive, imperative. (As is discussed below, sometimes the infinitives are regarded to have the infinitive mood, which is a nonfinite mood.)
- *Voice*: active, passive.

The list above, strictly speaking, is about morphological forms of the verb, but unlike the case for infinite verbs, a finite verb in, say, the second person singular present active indicative form, appears *nowhere* besides the predicator position in a clause with a second person singular subject, active voice, present tense (= present tense + simple aspect), and indicative mood. The relation between the conjugated form of a finite verb and the clause categories headed by that verb is one-to-one, compared to nonfinite verb forms.

There are still periphrastic conjugations beside the conjugation paradigm mentioned above. If *sum* + perfect participle is considered as a periphrastic conjugated form, then so should *sum* + other participles be so. The *sum* + present participle construction is rare, but the other two – *sum* + future participle and *sum* + gerundive – are indeed recognized as periphrastic conjugation (Allen and Greenough, 1903, § 193-196).

Now I turn to nonfinite verb forms. The terminology of them is completely chaos, and beginners are often confused by simple questions like how many nonfinite verb forms there are. Morphological forms include

- Four participles. Participles are adjectives in traditional terms, which means their prototypical roles are attributive modifiers in NPs. Participles include
 - The present participle, or the present active participle.
 - The future participle, or the future active participle.
 - The perfect participle, or the past participle, or the perfect passive participle. Despite the name “past participle”, the perfect participle appears as a part of the periphrastic conjugation in finite, passive, and perfect/pluperfect/future perfect clauses.
 - The gerundive, or the future passive participle.

Sometimes the gerundive is excluded from the participles and discussed together with the gerund.

- The gerund, which is traditionally considered as a noun, which, in modern terms, means it often heads nonfinite content clauses.

The gerundive is sometimes discussed together with the gerund and not participles, because their close relation both in morphology and in syntactic function: note, especially in Classical Latin, that a gerundive modifying a NP is usually used in place of a gerund taking an object.

Quite frequently, a *supine* form is distinguished. This is an example of function in form's disguise: in a nonfinite purpose clause, often the neutral accusative declension of the past participle occurs as the predicator, and since the purpose clause position differs from the usual syntactic context (i.e. attributive) of the past participle so significantly, a new name *supine* is invented to cover the former.

- The infinitives. The infinitives are said to have the infinitive mood by some, and in this framework, the infinitive mood is compatible with the present, perfect, and future tenses, and both the active and passive voices, so there are six infinitives, as in § 2.3.5. Nonetheless, only three of them have distinct morphological marking. Here are they:
 - The present active infinitive.
 - The perfect active infinitive.
 - The present passive infinitive.

The other three are periphrastic one.

Note that Latin infinitives are rarely used as purpose clauses in Classical Latin.

Therefore, if only considering morphological distinct forms, there are 4 (the four participles) + 1 (the gerund) + 3 (the three morphological infinitive forms) = 8 nonfinite verb forms. The gerundive may be classified as a variety of the gerund, and thus the equation will be $3 + 2 + 3 = 8$. If the supine is recognized as a separate form, then there are $8 + 1 = 9$ nonfinite conjugated forms. If the three periphrastic infinitives are taken into account, then there are $9 + 3 = 12$ nonfinite forms. This is how Allen and Greenough (1903) organizes nonfinite verb forms. There is no periphrastic present passive and perfect active participles, but there are strategies to express the same meaning in lieu of them (Allen and Greenough, 1903, § 491-493).

2.3.8 Verbal derivation

Derivations end up with a verb are listed here:

- *Denominal derivation*: adding a affix to a noun stem or adjective stem and getting a verb.
- *Verb from other verb*: inceptive, intensive, etc.
- *Compounding with particle*: compounding a preposition or an adverb, or a so-called **inseparable particle** (i.e. an affix that looks just like a preposition or adverb but nevertheless never appears freely). These devices are so-called *syntactic* compounding, in that the inner structure of the resulting verb has much interaction with the uncontroversially syntactic environment, for example, a verb arising from compounding with an accusative preposition takes a secondary object (§ 2.3.1). Of course, the line between morphology and syntax is highly blurred, and hence the notion of Distributed Morphology and morphosyntax. Here in *syntactic compounding*, *syntax* merely means what Latinists call syntax.
- *Compounding with noun stem*:

2.3.9 Constituent order in the clause

The actual rules about Latin clausal constituent order are still debated. There is evidence suggesting Latin is configurational, i.e. has phrase structures (Danckaert, 2017), but even when putting the everlasting functionalism v.s. formalism quarrel aside, since even the most non-configurational languages show certain degree of configurationality (Niedzielski, 2017; Morris, 2018; Legate, 2002, among others), *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? Unfortunately, it seems this question still does not have a clear answer.

What can be confirmed is that the constituent order has nuanced meanings: there is a SOV tendency, an initial verb means a sudden event, topics are usually placed initially, focuses are usually placed finally or just before the verb, etc. Problems like whether there is a neutral order, whether the constituent order variation can be explained by scrambling, etc. are yet to be settled.

2.4 Noun phrases and prepositions

2.4.1 Unattested categories

Unlike the case in English, a NP in Latin does not have syntactic marking of definiteness, and hence the determiner function (and related parts of speech, like the article) is completely absent. Definiteness can be indirectly marked by demonstratives, which is also how descendants of Latin obtained articles: by grammaticalized demonstratives.

2.4.2 The head

A NP is prototypically headed by a noun, but it is possible to have fusion-head constructions as in English: an attributive modifier, namely a relative clause or an adjective, can function as a NP itself, which may be analyzed as fusion of the head position and an attributive modifier position.

2.4.3 Case

Latin nouns and adjectives have six productive cases: nominative, genitive, dative, accusative, ablative, and vocative.

In traditional grammar (and modern grammars dealing with morphologically rich languages, such as Jacques (2021, chap. 8)), argument and adjunct positions (§ 2.3.1) are often introduced by the case requirements of them. The peripheral argument position about manner, for example, may occur in the chapter about cases, in the section about the ablative case, and in a subsection named “ablative of manner”. This is also the practice followed in this note (§ 5).

Early forms of Latin also had a locative case, which had largely been merged or replaced by dative, genitive and ablative in Classical Latin.

2.4.4 Number

The Latin number category has two values: singular and plural. Number is marked on the noun, and by agreement, all NP dependents with declension.

2.4.5 Gender

Latin has an idiosyncratic gender system, which means gender is *not* explicitly marked on nouns but is reflected on how nouns interact with NP dependents – in the case of Latin, the declension ending of adjectives, pronouns, etc. Latin has three genders: masculine, feminine, neuter.

2.4.6 Possession

Possession is coded in Latin by two means. The first is by possessive case: the possessor NP is possessive, which is placed into the possessed NP as a modifier. The second is by possessive

adjective: some nouns, especially proper nouns (personal pronouns included), have corresponding adjectives, which is placed into the possessed NP as an attributive modifier.

2.4.7 Determiner

2.4.8 Attributive

In Latin, the lines between the possessives, the determiner, the numerals in NPs, and typical attributives are highly blurred. Attributives besides § 2.4.6, § 2.4.7, and § 2.4.9 include

- AdjPs and relative clauses.
- Genitive NPs. Note that Latin does not have simple attributive nouns. In a construction like *apple tree*, *apple* is to be replaced by its genitive version.

2.4.9 Numeral

2.4.10 Nominal inflection

Nominal inflection, or declension, happens on nouns, adjectives, and pronouns. Nouns decline according to number and case. Adjectives decline according to gender, number, and case. They may have additional comparative or superlative forms (§ 2.8). Pronouns decline according to number and case, but they have idiosyncratic person features. Note that gender is *not* morphologically marked on pronouns.

2.4.11 Nominal derivation

2.4.12 Prepositional

Latin prepositions differ with English ones in one major aspect: they can be omitted. This fact leads many (for example BLT) to reject the head status of prepositions, and hence the preposition system is to be analyzed as a “syntactic (and optional) case system”.

Most prepositions either take ablative or accusative complements. A more limited group of prepositions take genitive complements.

2.4.13 Constituent order in the noun phrase (or preposition phrase)

Similar to the case of clausal constituent order (§ 2.3.9), Latin NP constituent order is also a topic without much consensus. The most frequent constituent order is preposition – attributive and determiner – noun – genitive possessor, but there are plenty of variations. The relative order between the noun and modifiers (attributives, determiners, genitive possessors) is highly flexible. The only hard rule seems to be that the preposition precedes its complement, with the rare exception that a monosyllabic preposition may occur between a modifier and the head noun (this phenomenon is called **hyperbaton**).

2.5 *Pro*-forms

Latin *pro*-forms, like all constructions that are called *pro*-forms, fill core and peripheral argument positions, and can be roughly classified into the following categories:

- *Personal pro-forms*, including
 - **Personal pronouns**, which carry the person feature (and usually have definite references). Latin first-person and second-person pronouns do not show gender marking, but third-person pronouns do. In practice this is not a big deal, since Latin is *pro*-drop and gender is usually to be inferred from verb agreement.

Table 2: Schematic correlative table

		quantification of reference →			
		interrogative	relative	demonstrative	quantifier
type of reference ↓	basic				
	pronoun				
	pro-adverb				

- **Possessive pronouns**, which also carry the person feature, but always appear in NPs, and hence the person feature is never marked on a verb. They can be regarded as adjectives corresponding to personal pronouns as proper nouns (§ 2.4.6).

Possessive pronouns and genitive personal pronouns have different distributions. For first-person and second-person ones, a genitive personal pronoun appears with the sense of attributive NP, for example *love of me* (i.e. towards me), while a possessive personal pronoun appears with the sense of possession, e.g. *my love* (i.e. love from me). However, third-person possessive pronouns are purely reflexive (they correspond to *his own* in *He killed [his own] horse* – see below), and third-person genitive pronouns cover all genitive constructions in which the possessor-like argument is not bound by a higher argument, be it *love of* (i.e. towards) *me* or *my love* (i.e. from *me*).

- **Correlatives**, which are always third-person. Correlative pronouns may have definite or indefinite references, and are always inherently third-person. They can be classified in the scheme shown in Table 2.

The axis of quantification of reference corresponds roughly to the quantifier in the logical form of the clause containing a correlative in question.

- A clause containing a **quantifier** has the meaning “I have something to say about all ...” or “I have something to say about some ...”. The quantifier-like thing is \forall and \exists .
- A restrictive relative construction containing a **relative pronoun** means “if there are something that ..., then for these things, I have something to say ...”, and in this case, the quantifier-like thing is Russell’s *description* ι , and a nonrestrictive relative construction means “for the aforementioned things, I have something to say”.
- A clause containing a **demonstrative** has the meaning “here/there is something and I have something to say about it ...”. The quantifier-like thing is still Russell’s ι , but the logical form is $\iota x(x \text{ is here/there and } \dots)$.
- A clause containing an **interrogative pronoun** has the meaning “it seems there are indeed ..., can you tell me more about them?” Naturally, interrogative pronouns and relative pronouns are often identified, but in Latin they do not overlap exactly.

As can be seen from the above discussion, some kinds of quantification of references (e.g. demonstratives) may be fusions of others.

The list of possible values of quantification of reference in Latin correlatives is too long to be shown here.

The type of reference axis is about what kind of object is referred to.

- A **basic** correlative denotes anything.

- A **determiner** correlative denotes anything and can be a determiner in a NP. Note that the term *determiner* is about function, not form, especially not about lexical category: Latin lacks the article category, but demonstrative determiners work well (compare with English: there is no *a girl*, but there is *this girl*).

It is natural to identify the basic type and the determiner type, which is the case in Latin, or, in more traditional terms, “a demonstrative can be both an adjective (read: determiner with adjectival morphology) and a pronoun”.

A correlative pronoun besides the aforementioned two can be regarded as fusion of function between the determiner correlative and a noun denoting the type. Take English as an example: the interrogative (quantification of reference) human pronoun (type of reference) *who* is the fusion between *what/which* (interrogative determiner) and the noun *person* indicating the type of the referred object. A correlative pro-adverb can be regarded as a correlative pronoun denoting place and manner, and the fusion of function between a correlative pronoun denoting place and manner and a preposition/case: *where* may be understood as *what place*, and it can also be understood as *at what place*.

The list of types of references of Latin correlatives is also too long to be shown here.

Latin pro-adverbs are strictly adverbial, which means the fused case/preposition, unlike English, is always there, and they do not decline (they do not need to, after all).

- There are also **reflexive pronouns**, which are named *anaphora* in generative grammar. When appearing in a clause, they have to be bound by some arguments with higher positions (typically the subject). Latin reflexives are purely third-person: the first-person and second-person personal pronouns can appear as reflexives as well as ordinary pronouns.

The third-person possessive pronouns may be considered as reflexive possessives, since they only appear in reflexive positions.

Latin does not have reciprocal pronouns (“each other”). The reciprocal meaning (i.e. reflexive of a referential expression that is both agentive and patientive) is conveyed by phrases like *inter se*.

2.6 Subordination

Subordination can be roughly divided into clause embedding and clause linking. The former means inserting a clause into a NP modifier (**relative clause**) or clausal argument position (**complement clause**, or in terms of (Allen and Greenough, 1903), **substantive clause**), while the latter means inserting a clause into a clausal adjunct position, which, for example, may be about time, or location, or condition. Of course the two functions do not have a quite clear distinction in general, since the border between clausal arguments and adjuncts is blurred.

In Latin, there is no counterpart of English *that*-relativization: all relative clauses are created by relative pronouns or adverbs (§ 2.5). Both complement clauses and adjunct clauses are realized by adding a **conjunction** before the subordinate clause, though this term is problematic in that unnecessary confusion between syntax and semantics may occur. Conjunctions, in this way, are similar to prepositions.

2.7 Coordination

Coordination is realized almost in the same way of complement clause and adjunct clause: by adding a conjunction.

2.8 The comparative construction

The comparative construction in Latin has the same scheme as in English. The comparison construction applies to AdjPs, acting as attributives and predicative complements, as well as AdvPs, acting as adjuncts. We call *A* in *more A than S* the **parameter** of comparison, while *S* is the

standard of comparison (BLT § 3.23). In Latin, the head of the parameter of comparison – be it an adjective or an adverb – is in its comparative form, while the standard is ablative. Latin also has superlative construction, in which the head of the AdjP or AdvP is in the superlative form.

3 Existing grammars, reference and pedagogical

3.1 Learning Latin

3.2 Reference grammars

Most traditional Latin grammars organize grammatical systems according to their morphological realizations. The reasons are summarized in § 2.1.

Ironically, older grammars (but not too old; at least after the emergence of comparative grammar) tend to have *more* fine-grained segmentation of morphemes than newer materials, the latter mainly using large inflectional paradigms without deeper analysis. Cross-linguistically, not every language can be well captured by Item-and-Arrangement morphology, and even Item-and-Process morphology may be not so handy. It is often assumed by syntacticians that morphology is merely syntax within the word (and various theories in this approach, most prominently Distributed Morphology, have been proposed). These theories are often claimed to be morpheme-based. However, with all these syntactic and post-syntactic operations (which infamously makes syntax completely chaos and even tempts many to assert that syntax varies unboundedly), the so-called “morphemes” in these theories are by no means reflected in surface-oriented analysis, and hence certain Word-and-Paradigm devices would better be used in some stages (Anderson, 2017).

4 Noun morphology

4.1 Noun declension

Latin has five noun declension classes. Every introductory book or reference grammar covers all of them and it makes no sense to repeat them here.

5 Cases, and where to find them

6 Structure of noun phrases

7 Verb morphology

7.1 Verb conjugation

7.1.1 Overview of regular conjugation

Factors involved in conjugation have been introduced in § 2.3.7. This section is about how conjugation is actually realized.

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 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 (tediously) summarized as more fine-grained conjugation classes.

This is also the case for Latin verbs. 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. All forms mentioned in § 2.3.7 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 *irī*.

This process is summarized in Fig. 1.

In practice, the three stems are not what stored in the dictionary, for two reasons. First, fluent users of a language often tend to *not* anatomize the language in detail, since “everything is so natural”, and recording actually attested word forms is hence easier to do compared to the morpheme-based approach. Second, Latin has four conjugation types, and hence the three stems themselves are not 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 scheme the verb paradigm can be found in (Allen and Greenough, 1903, § 166). The following sections in the book discusses how to do local phonological adjustments to get the correct word form. How each conjugation form is formulated can be found in (Allen and Greenough, 1903, § 180).

7.1.2 Verb endings

8 Verb complementation types

9 Important grammatical words

In principle, a dictionary only needs to cover all lexical words: grammatical words, like conjunctions, prepositions, etc. can be enumerated in the grammar. In practice, two factors prevent this from happening. One is the fact that grammar writers and dictionary editors are often different people and no one has attempted to describe a language (a huge project I tell you!) from grammatical rules to lexical entries in a consistent way. A dictionary editor, therefore, does not dare to skip grammatical words, fearing that the reader may have not learned about them somewhere else. The other is the line between lexical words and grammatical words is often not that clear. Is the copula a grammatical word? Of course it is: its forms can be enumerated finitely, and itself has no semantic meaning. But its behavior is just too close to a lexical verb. Prepositions may also be viewed as grammatical words, since they are markers of a “syntactic case system” (§ 2.4.12). But there are just too many of them. A grammar listing all Latin prepositions will be like a dictionary.

In this section, I review

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.
- Stephen R Anderson. Words and paradigms: Peter h. matthews and the development of morphological theory. *Transactions of the Philological Society*, 115(1):1–13, 2017.
- Hans Broekhuis. A typology of clause structure. *Linguistic Variation Yearbook*, 10:1–31, 12 2010. doi: 10.1075/livy.10.01bro.
- Lieven Danckaert. *On the left periphery of Latin embedded clauses*. PhD thesis, Ghent University, 2011.
- Lieven Danckaert. *The development of Latin clause structure: A study of the extended verb phrase*, volume 24. Oxford University Press, 2017.
- Josef GJ De Kuyper. Gerund, gerundive and supines as used by cj caesar: A statistical approach i. *Journal of Humanities*, 5(1):29–40, 1991.
- Andrew M Devine and Laurence D Stephens. *Latin word order: Structured meaning and information*. Oxford University Press, 2006.
- Diana Forker. *A grammar of Sanzhi Dargwa (Volume 2)*. Language Science Press, 2020.
- Dianne Friesen. *A grammar of Moloko*. Language Science Press, 2017.

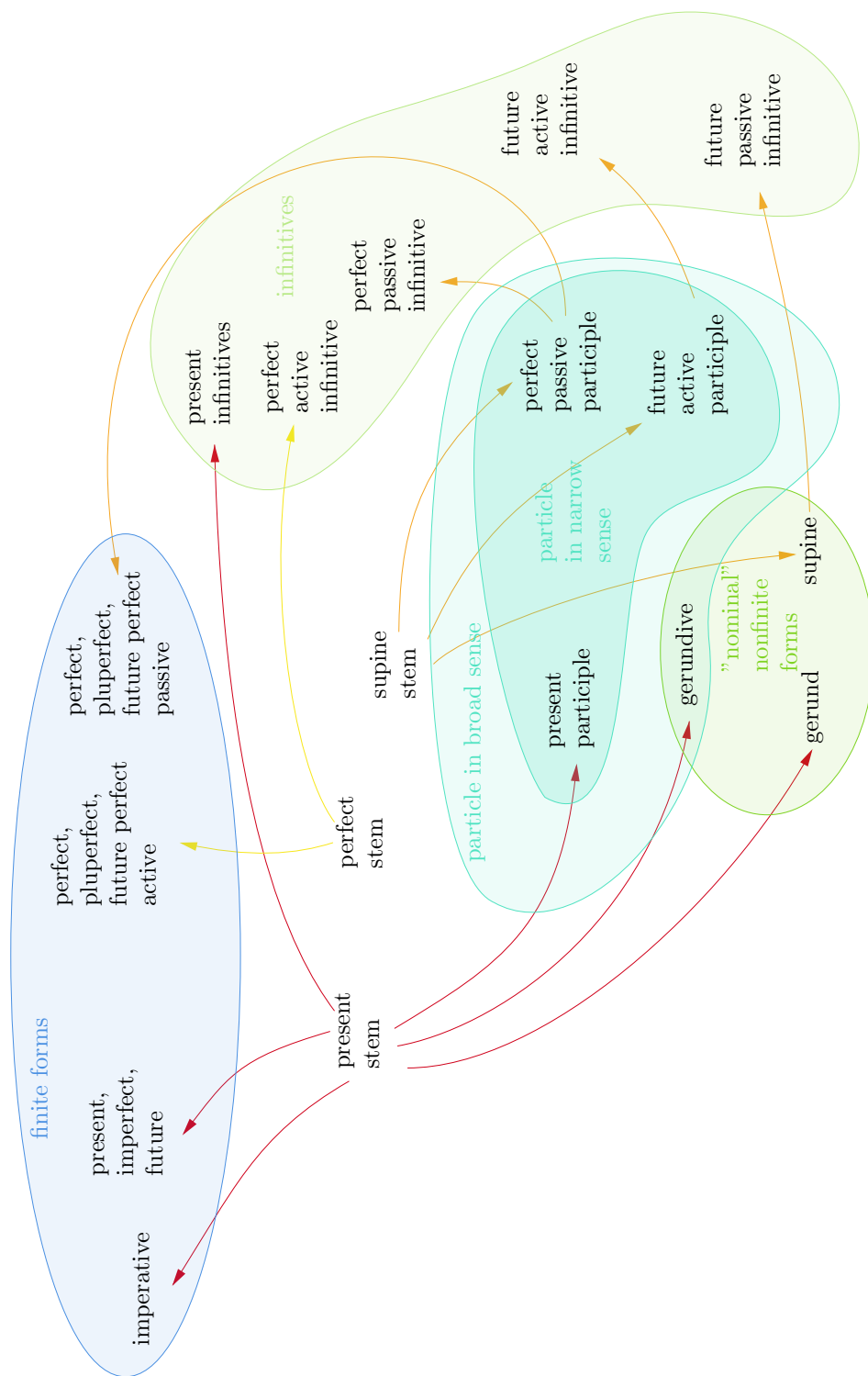


Figure 2: How to get all conjugation forms from the three stems

- 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.
- Jaume Mateu and Renato Oniga. Latin syntax in fifty years of generative grammar. *Catalan journal of linguistics*, 16:5–17, 2017.
- 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.
- Olga Spevak. Latin word order. structured meaning and information. *Mnemosyne*, 60(3):497–501, 2007.
- Maria C Tierney. A syntactic analysis of negation in latin: Negated quantifier phrases. 2018.