Notes about Latin grammar

Jinyuan Wu

June 25, 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 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 in 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 an again 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" specifies the syntactic environment). 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 a language like English where complement clauses serve as adjuncts, because this term works in a

---

[1] Note that here "function" *does not* mean "pragmatic function" as in functionalism.
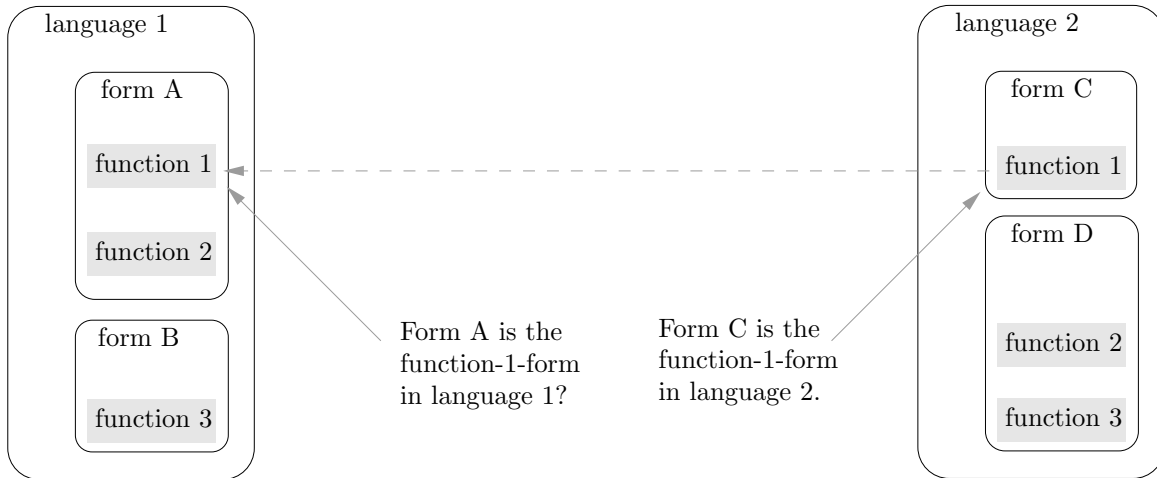
Figure 1: Why confusion form and function in practice does not create much inconvenience: only form C in language 2 is able to fill function slot 1 means in language 2, we can call form C as the function-1-form, and hence we can call form A – which realizes function 1 – as the function-1-form as well. 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.

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 a 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

NP

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 topological 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. Should the organization of Jacques (2021, chap. 2) be used and hence constructions are described in a bottom-up way, with grammatical categories being introduced in discussions about smallest constructions that show them, § 2.3.1 would be placed in a "core and oblique arguments" section,

It can be seen that describing a language purely in terms of "what happens in what construction" as in following sections is usually cumbersome. § 2.3.1 is obviously too long, and deserves at least a subsection, with topics like agreement and subject-predicate relation occupy subsubsections. But since it all happens within the clause, it has to be placed under the subsection about clause structure, 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 and the subject position. Similarly, § 2.3.7 can be divided into sections about methodological issues to distinguish inflection forms, conjugation in finite clauses, conjugation in infinite clauses, etc., but it is already a subsubsection and cannot be further divided.

Jacques (2021), on the other hand, introduces categories in a bottom-up manner, and thus there is no section devoted to the clause structure, so contents in § 2.3.1 and the rest of § 2.3 can be well split into two sections, which are § 2.4 and § 2.5 in his book. On the other hand, sections from § 2.3.2 to § 2.3.6 are too short. They can be all inserted into § 2.3.7.

2.2 Parts of speech

Traditionally, Latin lexical categories are

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 restricted 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.

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**. Otherwise it is **intransitive**. Here is a list of grammatical properties of the direct object:

- *Coding of semantic role*:

Table 1: Latin tense and aspect

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

- *Case marking* (§ 2.4.3): direct objects are always accusative – but not all accusative arguments are direct objects.

- *Passivization*:

Subject, direct object, indirect object, and various

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 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 adverbs. 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 adverbs), 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 adverbs, makes them much more verb-like.

Historically, the traditional grammar concept that nonfinite verbs are somehow nominalized origins 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, 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 conjunction, 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 be 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 them:

  – 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

2.3.9 Constituent order in the clause

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 usually 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 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".

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.

2.4.6 Possession

2.4.7 Nominal inflection

2.4.8 Constituent order in the noun phrase

2.4.9 Prepositional phrases

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".

2.5 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**), 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. It should be noted that

2.6 Coordination

Clause linkers coordinator is much more frequently called **conjunction** in Latin, though the latter term is problematic in that unnecessary confusion between syntax and semantics may occur.

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.

4 Nouns

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 Verbs

7.1 Verb conjugation

8 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" (§ 2.4.9). 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.

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.

Andrew M Devine and Laurence D Stephens. *Latin word order: Structured meaning and information*. Oxford University Press, 2006.

Guillaume Jacques. *A grammar of Japhug*, volume 1. Language Science Press, 2021.

Jaume Mateu and Renato Oniga. Latin syntax in fifty years of generative grammar. *Catalan journal of linguistics*, 16:5–17, 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.