Introduction

Latin constituent order is variable[1] in the sense that the position of a constituent in its sentence or phrase is not determined by the grammatical structure of that sentence (for example, whether it is subject in its sentence), nor is it indicative for an addressee (a hearer or reader) of its grammatical function. In Latin constituent order is rather determined by and indicative of (i) the role of a constituent within the discourse to which its sentence belongs (for example, whether the constituent is present in some way in the preceding discourse), (ii) the speaker or writer's estimation of what the addressee knows and expects, and (iii) how important the speaker and hearer consider a constituent within the overall communication. For these three, partially overlapping, aspects I will use the term "pragmatic value". In other words, in Latin a constituent can change its place without changing its syntactic function;[2] however, this does not mean that the information conveyed is the same.

Much has been written about Latin constituent order. In this introductory chapter, I will briefly sketch the main approaches and the problems involved. I will first deal with different approaches to Latin constituent order: the traditional (Section 1), typological (Section 2), generative (Section 3), and pragmatic (Section 4) approaches. As pragmatic approaches are the most important ones for my description, I will pay special attention to Firbas's communicative perspective as it was applied to Latin by Panhuis (Section 4.1) and to a general presentation of the perspectives of Functional Grammar (4.2). In Section 5, I briefly mention other aspects of Latin constituent order. The aim of my study and information about the methodology and corpora used are presented in Section 6.

1. For the term "free" constituent order, see below, Section 1 and Chapter 2, Section 1.

2. It is worth pointing out that already Weil (1844: 25) rightly observed that there is a – pragmatic – difference between the following sentences: *Romulus Romam condidit.* 'Romulus founded Rome.' *Hanc urbem condidit Romulus.* 'This town was founded by Romulus.' *Condidit Romam Romulus.* 'Romulus founded Rome.' Weil wrote: "These three sentences inform us about different things because their elements, remaining the same, are distributed in a different way." See Chapter 2, Section 1.5 on question tests; indeed, these sentences convey information answering three different implicit questions: what did R. found?; who founded this town?; and what happened?

1. The traditional approach

The traditional approach to Latin constituent order arose from a comparison of Latin with other languages: modern European (Romance, Germanic and Slavic languages) and ancient ones (Sanskrit and Ancient Greek). From this point of view, constituent order in Latin was labelled "free", in contrast with the "fixed" order in French and German, for example. This distinction was introduced by Weil (1844: 54) and elaborated by Meillet (1903). The traditional approach focuses on several properties that are observed in other languages, especially the position of the finite verb and the placement of enclitics.

In traditionally oriented works, the verb is considered the most important element of a sentence, and therefore, its position has drawn much attention. As Classical Latin authors, in particular Caesar and Cicero, frequently put the verb at the end of a sentence,[3] this position is described as "normal"; the other empirically observed positions, initial and internal, are described as "deviations" (cf. Schneider 1912). Kühner and Stegmann's Latin grammar (1914, II: 590 ff.) shares the same point of view and talks about "traditional" and "occasional" order. The same distinction between final, initial and internal verb position is found in Marouzeau's monograph (1938). He considers patterns with a final verb as "unmarked"; on the other hand, "marked" initial and internal patterns have to be explained with the help of a set of factors responsible for these "deviant" placements of the verb.

2. The typological approach

Ideas about "normal" or "unmarked" order, which have a long tradition in descriptions of constituent order,[4] led to a typological classification of languages by Greenberg (1966).[5] According to him, languages are supposed to exhibit one basic, unmarked order of the S(ubject), the O(bject), and the V(erb); consequently, other orderings are viewed as deviations. Although it is generally stated that from a typological point of view, Latin exhibits a basic SOV order, the question is more complicated. An SOV language is also claimed to use postpositions rather than prepositions, and to place adjectives as well as genitive modifiers before their head nouns, which does not hold true for Latin. Discussion of Greenberg's criteria showed that Latin is perhaps not only

3. This statement is based not only on empirical observation but also on a passage from Quintilian (*Inst.* 9.4.26) who advises his readers to put the verb in the sentence-final position, because the 'force' of the sentence is in the verb. For a recent analysis of this passage, see de Jonge (2008: 321 ff.).

4. They mainly reflect a distinction between the subjective and the objective order, made first by psychologically oriented studies on word order and also used by Mathesius, a founder of the Prague School.

5. For typological approach to Latin, see Bauer (2009: 241–271).

an SOV language, but also SVO.[6] This divided scholars into two groups: some of them hold that Archaic and Classical Latin has the SOV order as the basic one, which underwent a typological shift to SVO, the dominant order in the Romance languages – at least in some of them, especially in French. Others argue that the SVO order is already present in Archaic Latin and that SVO coexisted along with SOV (Adams 1976a) throughout the language's history.

The typological approach to Latin constituent order suffers from several difficulties that were discussed in detail by Pinkster (1991; cf. 1995: 218). The first is that Latin constituent order largely depends on pragmatic rules and has no syntactic implications. Consequently, it is not very helpful to describe it in syntactic terms such as subject, object, verb because pragmatic values cannot be translated into the formulas such as SOV, SVO, or OVS. On the other hand, without claiming a one-to-one correspondence, we can say that a syntactic order is the outcome of certain pragmatic conditions (cf. Pinkster 1992: 522). The second difficulty is that descriptions using the syntactic functions of subject, object, and verb are centred on clauses containing transitive verbs; other types of verbs (monovalent and trivalent) are usually neglected. The third problem concerns the considerable variety of patterns occurring not only in different Latin authors, but also in the works of one author (Pinkster 1992: 521 f. and 1991: 72).[7] For all these reasons, I will not discuss the question of the "basic order". Indeed, how is one to decide what order is the basic one when we have to deal with a large range of pragmatic arrangements? In sum, Greenberg's typology poses more problems than it solves and therefore, it is no longer adopted as a sustainable framework (Sörés 2004).

3. The generative approach

The approach to word order within the framework of Generative Grammar resembles in a sense the typological one: it is based on the hypothesis that there is one basic, unmarked word order and a set of transformational rules producing marked patterns. Latin word order is supposed to exhibit one neutral, basic SOV order; other orderings represent deviations. In order to explain these "deviations", Generative Grammar works with the concept of movement rules, scrambling, as well as pragmatic features (Topic, Focus, Contrast, Emphasis, Heaviness).[8] Recent contributions[9] within this framework to Latin constituent order are Devine and Stephens (2006), Polo (2004), and Salvi (2004). As I have discussed several problems posed by these contributions

6. See Panhuis (1984a) for details. He himself considers Latin as an ambivalent language (p. 154).

7. See also Panhuis (1981) on non-final placement of the verb in descriptive passages in Caesar, and Spevak (2005a).

8. The concepts of Topic and Focus are different in Functional Grammar, see below, Section 4.2.

9. The transformational approach was first applied to Latin by Ostafin (1986).

elsewhere (Spevak 2006c and 2007c), I limit myself here to saying that Generative Grammar is a well-constructed theory in which nowadays pragmatics and discourse play a certain role. However, due to its continuing insistence on a basic (underlying) word order it does not really help a Latinist to better understand Latin constituent order, especially because of the great number of "deviations" that we must admit in that case. It also poses problems with deciding what counts as empirical evidences, for there is much variation in constituent ordering not only between Latin authors but also between works of individual authors.

4. Pragmatic approaches

4.1 Firbas's communicative perspective (Panhuis 1982)

Panhuis's study (1982) was the first attempt to provide a systematic description of Latin constituent order from a pragmatic point of view. As a general framework, he uses the theory of Functional Sentence Perspective, elaborated by linguists of the Prague School (mainly Mathesius, Daneš, and Firbas).[10] In the theory of Functional Sentence Perspective, the distribution of elements in a sentence is viewed as a result of both the communicative aims of the speaker or writer and the function of that sentence in its context, textual as well as situational. The analysis works with the distinction between a "theme" and a "rheme".[11] The theme is defined as the less informative element whereas the rheme is the most informative one. Since communication represents a dynamic process where information is conveyed from a speaker or writer to an addressee, it is supposed to proceed in a forward direction. Elements of a sentence contribute in different degrees to the progression of the message: they have different degrees of "communicative dynamism". What governs the distribution of elements in a sentence (in a normal, non-contrastive or non-emphatic context) is the principle of increasing communicative dynamism: the element bearing the lowest degree of communicative dynamism comes first, the element with the highest degree – or we can say the most informative one – comes last. Firbas's theory was elaborated on the basis of constituent order phenomena in Czech that he compared with English and German. Whereas the constituent order is predominantly governed by the Functional Sentence Perspective

10. Firbas (1992) concerns a synthesis of his previous works.

11. It is worth noticing that the distinction between "theme" and "rheme" was already suggested by O. Behaghel (1932: 4): "the element that is less important (or already known to the listener) is placed before the more important element" (Behaghel's Second Law). Firbas's (1992) concept of the "theme proper" partially corresponds to Dik's (1997, I-II) concept of Topic. However, Theme in Functional Grammar covers another pragmatic function (see below, Chapter 2, Section 6). Firbas's "rheme proper" corresponds to Focus in Functional Grammar.

(i. e. by increasing communicative dynamism) in Czech, English and German are less sensitive to this principle.

For Latin, Panhuis (1982: 54 f., cf. 1981: 300) claims three main arrangements: (i) a normal, non-emotive arrangement {theme > rheme > verb}, and two "emotive" arrangements (ii) {rheme > theme > verb}, and (iii) {verb (rheme) > theme}. In other words, a Latin sentence is supposed to obey the principle of communicative dynamism just as in Czech. The only difference is that the Latin verb takes the final position regardless of the degree of its communicative dynamism in arrangements (i) and (ii). In order to show the normal, non-emotive arrangement, I borrow an example from Panhuis (1982: 44 and 55) that illustrates the principle of increasing dynamism. In a passage from Plautus, a young man Phaedromus complains to his slave about his situation because a pimp (*leno*) is tormenting him. When the slave asks why (*Quid est?*), Phaedromus replies:[12]

(1) Alias me poscit pro illa triginta minas,
sometimes me-ACC he asks for her thirty-ACC minae-ACC
alias talentum magnum
sometimes talent-ACC great-ACC
'Sometimes he asks me thirty minae for her, sometimes a great talent.'(Pl. *Curc.* 62)

This sentence can be divided into two parts: a thematic and a rhematic one. The understood subject (*leno* 'pimp') belongs to the thematic part, as well as the setting element *alias* 'sometimes' and the personal pronoun *me* referring to the speaker. In the rhematic part, we have the verb *poscit* '(he) asks' (that may be labelled as "transition" in Firbas's terminology), followed by its non-obligatory complement *pro illa* 'for her' that is rhematic but less than *triginta minas* 'thirty minae' that is the most informative constituent of the sentence. This sentence can schematically be represented as follows:

Thematic part			**Rhematic part**		
(*leno*)	*alias*	*me*	*poscit*	*pro illa*	*triginta minas*
theme proper >	theme >	theme >	rheme >	rheme >	rheme proper

12. Examples quoted in this book are accompanied by interlinear word-to-word glosses; their aim is to provide information about constituent ordering, and not a full morphological analysis (a complete morphological analysis of examples is available on Perseus, collection Classics, www.perseus.tufts.edu). Case marking is provided for nouns and noun phrases; prepositional phrases are signalled by prepositions. In some cases, when not interrupted by other elements, constituents are treated as a whole and bear only one case marking, when it is evident what belongs together. Only sentences or clauses that are analysed are glossed; preceding or following context is provided between brackets [ ]. Abreviations used for glossing follow the Leipzig glossing rules (www.eva.mpg.de/lingua/pdf/LGR09_02_23.pdf).

Panhuis's aim to analyse sentences in their context and to determine what is new information deserves appreciation. On the other hand, the principle of communicative dynamism raises several other questions, such as how to decide about the degree of communicative dynamism of individual constituents. This difficulty concerns not only the arrangement of thematic constituents (cf. Panhuis 1982: 55 f.) but also rhematic ones. Which criteria allow us to say that one constituent is more thematic or more rhematic than another one? The absence of native Latin speakers renders this task difficult. The principle of increasing communicative dynamism governs the ordering of constituents in Czech, as Firbas, among others, has convincingly shown, and in normal conditions (in the absence of contrast), the final sentence-position receives the most informative element, the rheme proper. However, does the same principle always work in Latin? In Panhuis's analysis, a number of examples can be regarded as conforming to the principle going from the less informative to the most informative constituent. In other cases, in particular in sentences with a final verb, the most informative constituent does not necessarily stand before the verb.[13] In other words, the final or pre-verbal position is not reserved for the rheme proper constituent, and hence this position does not seem to be a reliable indicator that the constituent occupying it has rheme function.

4.2 The perspective of Functional Grammar

4.2.1 *Pragmatic functions in Functional Grammar*

In order to describe constituent order, Functional Grammar distinguishes between two main pragmatic functions: Topic, the entity the sentence is about, and Focus, the salient element of the sentence (Dik 1997, I: 312 f.). These pragmatic functions affect the formal side of the sentence: special positions are reserved for Topic and Focus constituents and they can be marked by specific means.

4.2.1.1 *Topic*

The pragmatic function of Topic, "what is being talked about", can be assigned to entities (persons, objects, localities, etc.).[14] A discourse may contain several Topics: one can be more central than others, the importance of Topics can change and new Topics can be introduced. According to their status in the discourse, Functional Grammar distinguishes several types of Topics: Discourse Topic, Future (New)[15] Topic, Given Topic, Resumed Topic and Sub-Topic.[16]

13. See Chapter 2, Section 1.3.2. p. 29 ff.

14. Several scholars argue that also verbs can be considered as Topics – H. Dik (1995: 207) and Bolkestein (1998a: 197).

15. I will follow the suggestion made by Bolkestein (2000) and call it "Future Topic". Indeed, in spite of its name, Future Topic has focal character. It functions as Topic not in the presentative sentence but afterwards (Toth 1994: 178, and Wehr 1984: 7).

16. In my study, I will not use all these types, but only some of them.

Discourse Topic (Dik 1997, I: 313–326) is the entity about which a discourse (story, paragraph, chapter, etc.) gives information. In order to function as Discourse Topic, the entity has to be introduced and presented. This procedure, necessary for ensuring a coherent development of the discourse, is labelled the "introduction of Future Topic" (Dik 1997, I: 315). Once introduced, the new entity is treated as a Given Topic and can be maintained in the text for a shorter or a longer time. Nevertheless, it is necessary to keep the entity present in the mind with the help of topical or anaphoric chains. For example:

(2) Yesterday, I got a phone call from the tax inspector. He/The man/The joker wanted me to come to his office, and he/Ø gave me the impression that I was in for some trouble.[17]

The tax inspector is a new entity that is introduced into the discourse (Future Topic). It is maintained as Given Topic with the help of a personal pronoun (*he*), a noun that specifies the category it belongs to (*the man*), a qualifying noun (*the joker*), or zero anaphora (Ø). These resumptive means differ as to their force: a noun is a stronger form of anaphora than the anaphoric pronoun, zero anaphora is the weakest means (cf. Givón 1983: 17). After the intervention of another Topic, the speaker can return to the previous one, by referring to it as *the tax inspector*.

It is also possible to deduce a Topic by association with another word. For example, on the basis of shared knowledge that music is played at parties, *the music* is not a new element but a Sub-Topic, which explains the presence of the definite article *the*:

(3) John gave a party last week, but *the music* was awful.

4.2.1.2 *Focus*

Focus is the salient, or most informative, element of a sentence (Dik 1997, I: 326 ff.). The pragmatic function of Focus can be assigned to entities as well as to actions. Focus adds new information or replaces information that we already have. It need not convey completely new information but may be in contrast with another element, either explicit or implicit. This is exemplified in (4a) with Focus represented by new information and in (4b) with Focus that contrasts with another, implicit element. The sentence in (4b) is a reaction to (4a). Saliency due to explicit contrast is illustrated in (4c), presenting two judgments concerning John and Bill; *nice* and *boring* function as explicit contrastive Focus constituents.

(4) a. I've just bought *a Peugeot*.[18]
b. Did you buy *a Peugeot*???
c. John and Bill came to see me. John was nice, but Bill was rather boring.

17. Here and below, I quote Dik's examples (1997, I: 314 and 318).

18. In this section, I quote Dik's examples (1997, I: 326 and 328).

Languages have various means of signalling the saliency of an element. These include prosodic prominence, one or several special positions reserved for this type of elements, focusing particles, and special syntactic constructions, such as cleft constructions (cf. also Miller 2006: 128).

4.2.2 *General principles and hierarchies*

Functional Grammar, mainly based on works by Dik (1997, I-II),[19] views constituent order as an interplay of a diverse range of syntactic, semantic, and pragmatic factors. The factors involved can be subcategorised and described as linearization rules or linearization hierarchies. The remarks that follow are grounded in observations of Siewierska (1988: 29), who discusses in detail three main groups of hierarchies: formal, dominance, and familiarity hierarchies.

4.2.2.1 *Formal hierarchies*

The formal hierarchies are formulated as the Language Independent Preferred Order of Constituents (LIPOC) schema (Dik 1997, I: 411). They include the rule of increasing complexity[20], a principle that is responsible for the rightward placement[21] of syntactically heavy constituents such as complex noun phrases, for example *M. Scaurus* in (5), which is expanded by a subordinate clause and an apposition.

(5) In quis fuit *M. Scaurus*, de quo supra
among them was M. Scaurus-NOM of whom above
memoravimus, consularis et tum senatus princeps.
we have spoken ex-consul and then of Senate leader
'Among them was Marcus Scaurus, of whom I have already spoken, an ex-consul and at the time the leader of the Senate.' (Sal. *Jug.* 25.4)

The formal hierarchy also includes the rule of domain integrity[22] according to which "what belongs together should be kept together" (Dik 1997, I: 402). Constituents belonging to a domain such as noun phrase, infinitive or participial clause, etc. are not expected to leave it and go to another domain. An example of

19. The work referred to as Dik (1997, I-II) is the second, enlarged version of his previous grammar published in 1978.

20. When talking about the principle of increasing complexity, it is worth recalling Behaghel's "Law of Increasing Terms" (*Gesetz der wachsenden Glieder*, Behaghel 1909 and 1932: 6); according to him, "given two phrases, when possible, the shorter precedes the longer".

21. It is important to stress that, in contrast with the Generative Grammar framework (see Section 3), linearization principles following Dik (1997, I: 392) are template principles and not "movement" principles.

22. Cf. "First Behaghel's Law": "Elements that belong close together intellectually will also be placed close together" (Behaghel 1932: 4).

contiguous syntactic units is given in (6). As we will see (p. 22), Latin may rather easily violate this rule.

(6) His rebus impulsus / equitatum omnem / prima nocte /
by these things urged cavalry-ACC all-ACC at nightfall
ad castra hostium / mittit / ad flumen Bagradam.
to camp of enemies he sends to river Bagrada
'Under the influence of these factors, he sends off all his cavalry at nightfall to the enemy camp at the River Bagradas.' (Caes. *Civ.* 2.38.3)

4.2.2.2 *Dominance hierarchies*

"Dominance" hierarchies concern priority given to two categories: person and animate entities, and semantic roles, as is represented in the following schema (from Siewierska 1988: 30):

THE PERSONAL HIERARCHY
1st person > 2nd person > 3rd person human > higher animals > other organisms > inorganic matter > abstract entities

THE SEMANTIC ROLE HIERARCHY
agent > patient > recipient > benefactive > instrumental > spatial > temporal

In accordance with the personal hierarchy, the first person has priority over the second person; nouns denoting human beings are preferably placed before nouns denoting inanimate or abstract entities. The effect of the semantic role hierarchy on the linearization of constituents is that the agent has priority over the patient, the patient over the recipient, and so on. At the same time, the semantic role is largely connected with the eligibility of the agent as syntactic subject, and of the patient as object.

4.2.2.3 *Familiarity hierarchies*

The hierarchies labelled as "familiarity" (Siewierska 1988: 61) represent a group of hierarchies concerning the speaker's personal involvement, and encompasses mainly givenness, definiteness, referentiality and iconicity.

THE FAMILIARITY HIERARCHY
more familiar > less familiar
given > new
definite > non definite
referential > non-referential

According to these hierarchical principles, the entity that is more familiar to the speaker has priority; given information tends to be placed before new information, a definite entity before an indefinite one, a referential entity before a non-referential one. The familiarity hierarchies occupy the highest place on the scale of hierarchies. This is why the Topic > Focus arrangement (given > new hierarchy) does not necessarily coincide

with the agent > patient preference in the semantic role hierarchy (Siewierska 1988: 65). The preference for Topic being placed first is observed in Latin (Pinkster 1995: 123) as well as in Dutch and in Polish, whereas in French and English initial non-subject constituents are unusual. The following example is taken from Caes. *Civ.* 2.23.5.

(7) Hunc secutus Marcius Rufus quaestor navibus XII.
him-ACC followed M. Rufus-NOM quaestor with 12 ships
Hem volgde de quaestor Marcius Rufus met twaalf schepen.
Za nim podążył kwestor Marcius Rufus z 12 statkami.
Le questeur Marcius Rufus l'avait poursuivi avec douze vaisseaux.
'The quaestor Marcius Rufus followed him with twelve ships.'

Linearization of constituents can also reflect our perception of extra-linguistic facts. Following the iconicity principle, especially concerning the temporal order (Siewierska 1988: 79), actions or events are presented in their actual temporal succession, as in (8). Information about Pompey's trip is organised according to the iconic principle: his point of departure (*Luceria*) is given first, then comes the action (*proficiscitur*), followed by the direction (*Canusium, Brundisium*), in their order of occurrence.

(8) Pompeius... Luceria proficiscitur Canusium atque
Pompey-NOM from Luceria sets out to Canusium and
inde Brundisium.
thence to Brundisium
'Pompey... goes from Luceria to Canusium and thence to Brundisium.'
(Caes. *Civ.* 1.24.1)

5. Other aspects of Latin constituent order

5.1 Colon subdivision

Describing the structure of Latin sentences, Fraenkel (1964[2] and 1965) used an ancient rhetorical concept and divided a simple or a complex sentence into syntactically defined units called *cola*. A *colon* may be a participial and infinitive clause, an ablative absolute, an adverbial prepositional phrase, or indeed any subject, direct object, and indirect object composed of more than one element. Habinek (1986) elaborated Fraenkel's theory in more detail.[23] Apart from the above-quoted Example (6), that illustrates domain integrity as well as *colon* subdivision, I give another one in (9). Slashes mark a *colon* boundary.

23. Habinek (1986: 11) proposed a distinction between rhetorical *cola* and rhythmical *cola*. Especially what he calls rhythmical *cola*, defined as "short, detachable, grammatical constituents of the Latin sentence", are of interest for our subject. The *colon* division also inspired Adams (1994b: 109) when explaining the placement of Latin pronouns.

(9) Cuius adventu nuntiato / L. Plancus / qui legionibus
his arrival being announced L. Plancus-NOM who legions-DAT
praeerat / necessaria re coactus / locum capit superiorem /...
commanded by necessity compelled place-ACC takes higher-ACC
'On the report of his approach Lucius Plancus, who was in command of the legions, under the stress of necessity occupies the higher ground...' (Caes. *Civ.* 1.40.5)

In a number of cases, Latin complex sentences can be seen as combinations of several successive *cola*. On the other hand, not every sentence allows a clear *colon* subdivision, as pointed out by Janse (1997: 111; cf. Salvi 2004: 145). Furthermore, it would make no sense to talk about *cola* in cases such as (10) which contain only single constituents.

(10) Curio Marcium Vticam navibus praemittit.
Curio-NOM Marcius-ACC to Utica with ships sends forward
'Curio sends Marcius ahead to Utica with the ships.' (Caes. *Civ.* 2.24.1)

5.2 Stylistic ornamentation

Latin constituent order in literary prose (as well as poetry) can reflect various stylistic figures that serve as ornamentation (see, for example, Kühner and Stegmann 1914, II: 616–622). These figures will not be described in this book. I only mention two of them: firstly, chiasmus, which is an arrangement following the pattern *a b – b a*; for example in (11), *a* are the nouns *mala* and *voluptates*, *b* the adjectives *praesentia* and *praeteritae*. In other words, the second noun phrase has a different order of elements. Secondly, parallelism, which copies the same ordering of constituents, for example the verb and the subject in (12).

(11) Non enim video, quo modo sedare possint mala
NEG indeed I see how allay can evils-ACC
praesentia praeteritae voluptates.
present-ACC past-NOM pleasures-NOM
'I do not see how past pleasures can allay present evils.' (Cic. *Tusc.* 5.74)

(12) Luget senatus, maeret equester ordo.
mourns Senate-NOM grieves equestrian order-NOM
'The Senate mourns; the equestrian order is inconsolable.' (Cic. *Mil.* 20)

6. Aim and methodology of this study

This book aims to provide a systematic description of constituent order in Classical Latin prose from a pragmatic point of view, in order to elucidate the pragmatic value of Latin sentences; in other words, what information is conveyed by a sentence or a clause. For doing so, I adopt the approach of Functional Grammar, combining it with several

aspects worked out by the Prague School, such as contextual dependency. However, my aim is not to apply a theory to Latin constituent order but to try to understand more about it. My study integrates and discusses recent works on pragmatic aspects of Latin constituent order (Bolkestein, de Jong, Panhuis, Pinkster, among others); whereas these works are concentrate on particular questions or expressions – after all, Panhuis's monograph is a case study just as well –, my aim is to describe Latin constituent order in a systematic way. Furthermore, the reader will find suggestions to solve several questions that are traditionally discussed in Latin grammars, such as the placement of the verb in sentence-initial position, the placement of pronouns, coordinators, and ellipsis. This book also aims to treat problems of more general linguistic interest. Indeed, the Latin language does not figure in the Eurotyp project (Siewierska 1998a) and other typologically oriented studies. Thus, attention is paid to various peculiarities of Latin in the domains of the noun phrase, focusing particles, and cleft constructions.

The present study is based on several corpora. The main one is a corpus of Classical Latin prose including the works of Cicero, Caesar and Sallust. It contains different text types: philosophical treatises, speeches, correspondence, and historical narrative.[24] There is a comparable amount of material – about 21,000 words – in each of the text types. This corpus has served for the chapters on interrogative and declarative sentences. For the examination of Latin noun phrases (Chapter 6), another corpus was set up.[25] The chapter on imperative sentences mainly rests upon an examination of data provided by LASLA.[26] Also, various case studies were made with the help of the CD-ROM *Bibliotheca Teubneriana Latina* (*BTL*); for example, with respect to the positioning of the verb *mitto* 'to send', or verbo-nominal constructions in the chapter on declarative sentences. However, the present study concentrates on the verb and its obligatory arguments; it does not examine in a systematic way optional complements and subordinate clauses. In the near future, I will elaborate these points in detail and devote some articles to them.

The book is organised in the following way: Chapter 1 provides a survey of constraints and liberties in Latin constituent order. Chapter 2 presents my approach to Latin constituent order and principles of analysis, as well as the pragmatic functions Focus, Topic, Theme, Tail, and related phenomena. Chapters 3–5 discuss three main sentence types: declarative, interrogative, and imperative sentences. Chapter 6 is devoted to the order within the noun phrase. The most important results are summarised at the end of each section or chapter.

24. This corpus 1 includes: Cicero's treatise *Tusculan Disputations* (books 1 and 3); speeches *On His House* and *Philippics* (1 and 4); correspondence (*Letters to Atticus*, 13.50 – 16); Caesar's *The Civil War* (1 – 3.30), and Sallust's *The Jugurthine War.*

25. The corpus 2 contains Cicero's correspondence (*Letters to Atticus*, 1–4, 30,954 words), Caesars's *The Gallic War* (books 1–5, 28,696 words), and Sallust's *The Conspiracy of Catiline* and *The Jugurthine War* (32,360 words).

26. LASLA (*Le Laboratoire d'Analyse Statistique des Langues Anciennes*) is available for registered users on the Internet: http://www.cipl.ulg.ac.be/lsl.htm.

CHAPTER 1

Placement constraints and liberties in Latin constituent order

In the variable constituent order of Latin, there are a certain number of placement constraints as well as certain liberties that are not common in modern European languages such as Romance and Germanic, nor in the more flexible Slavic languages. The aim of this chapter is to catalogue at least the most important characteristics of Latin constituent order. One may note that Latin poetry is much more flexible than prose and shows even more liberties in the placement of constituents.

1. First sentence- and clause-position

Because of their semantic-referential and functional properties, coordinators, connective particles, subordinators and interrogative words are usually found in initial position in Latin. The regular position of coordinators such as additive *et, atque, ac* 'and', *neque, nec* 'and not', alternative *aut, vel, seu* 'or', and adversative *sed* 'but' is at the beginning of a clause or, if they coordinate constituents, before the coordinated constituent.

In Classical Latin prose, the initial position is obligatory for the connective particles *nam* 'for' and *itaque* 'so'. Whereas other connective particles may appear in another position, for example *igitur* 'therefore' (see Example 1) and *ergo* 'accordingly, then', *nam* (2) and *itaque* never leave the initial position, and thus they are placed before subordinators and interrogative words.

(1) Hanc tu *igitur* dedicationem appellas...?
this-ACC you therefore dedication-ACC call
'Is this, then, what you call dedication...?' (Cic. *Dom.* 118)

(2) *Nam* quis hoc non intellegit...?
for who this NEG understands
'For who cannot see that...?' (Cic. *Ver.* 1.9)

The first sentence-position is reserved for the connecting relative *qui*, which has an anaphoric function. In the case of its co-occurrence with another candidate for the initial position, for example with a subordinator (*quia*), the connecting relative has priority, as illustrated in (3) where *cuius* picks up the preceding *memoria rerum gestarum* 'the recording of past events'.

(3) (memoria rerum gestarum) Cuius de virtute *quia*
of which about virtue because
multi dixere, praetereundum puto.
many-NOM have spoken to be passed over I consider
'(the recording of past events) Since many have spoken of its merits, I do not think I need say more about it.' (Sal. *Jug.* 4.2)

Subordinators are expected to appear in the first position of a subordinate clause. In Classical Latin, relative pronouns and subordinators introducing a complement clause or a final clause mainly have this introductory function and mark the beginning of the subordinate clause. But temporal (*cum* 'when', *dum* 'as long as', *ubi* 'when'), causal (*quia* 'because', *quoniam* 'since') and conditional (*si* 'if') subordinators give way to another constituent(s) in 20% to 30% percent of cases.[1] They are typically preceded by anaphoric pronouns or noun phrases containing one (cf. Example 3 *cuius de virtute* 'of its merits'), and by Sentence Topics (*Tubero* in 4).[2] Focal or contrastive (*palam* 'in public') constituents are also found in sentence- or clause-initial position, as is shown in (5).

(4) Tubero *cum* in Africam venisset, invenit in provincia
Tubero-NOM when to Africa had come finds in province
cum imperio Attium Varum.
with command Attius Varus-ACC
'Tubero on reaching Africa finds Attius Varus in the province in military command.' (Caes. *Civ.* 1.31.2)

(5) Vbi id a Caesare negatum et, palam *si* conloqui
when this by Caesar was refused and publicly if negotiate
vellent, concessum est...
they want was accorded
'While Caesar refused this request, but allowed them, if they wished, to negotiate in public...' (Caes. *Civ.* 1.84.2)

This placing of subordinators introducing adverbial clauses is not restricted to Classical literary prose.[3] Similar arrangements also appear in technical prose; for example in Vitruvius's *On Architecture*. In (6), the initial position is occupied by a noun phrase containing the connecting relative *qui* (Topic) followed by the contrastive Focus *Auster* 'the South wind' and then comes the subordinator *cum*.

1. 30% in Caesar (*Civ.* 1), 21% in Cicero (*Tusc.* 1). These figures are even higher than Marouzeau's (1949: 123) estimation of 15–20% in Caesar and Cicero (without distinguishing subordinate clauses). In Sallust, other placement than initial is exceptional (see Marouzeau, *ibid.*). Latin grammars usually speak about *traiectio*, displacement of the subordinator (Kühner and Stegmann 1914, II: 614). See also Chapter 2, Section 3.6.3. p. 70.

2. See Panhuis (1982: 80 ff.), Pinkster (1995: 221), and Amacker (1998).

3. Poetry is even more flexible; a focal element standing before a subordinator may also be the

(6) In qua civitate Auster *cum* flat, homines aegrotant;
in this city Auster-NOM when blows people-NOM are ill
cum Corus, tussiunt.
when Corus-NOM cough
'In this city when the South wind blows people fall ill; when the North-west, they cough.' (Vitr. 1.6.1)

Question-words and interrogative particles (*num, nonne*) usually start questions, i. e. stand as the first word; they may be preceded by a coordinator (*et, sed*) or by a sentence-initial connective particle (*nam, itaque*). When a question-word is used in a noun phrase it can be separated from its head noun (7). In about 7–13% of cases, interrogative pronouns and particles occupy a position other than initial and are mainly preceded by a Topic constituent. Every type of Topic is found, contrastive (*tu* in Example 8), Sentence Topic or Resumed Topic (*totum prope caelum* in 9).[4] However, apart from instances of ellipsis, interrogative pronouns and particles do not appear in sentence-final position.[5]

(7) *Qui* sunt *homines* a Q. Metello, fratre tuo, consule in
who are men-ACC by Q. Metellus your brother consul in
senatu palam nominati...?
Senate publicly named
'Who are the men whom the consul Quintus Metellus, your brother, publicly named in the Senate...?' (Cic. *Dom.* 13)

(8) Atque ille tamen ad collegium rettulit, tu *ad quem* rettulisti?
and he yet to College submitted you to whom submitted
'Yet he submitted his dedication to the College; and to whom did you submit yours?' (Cic. *Dom.* 132)

(9) Totum prope caelum, ne pluris persequar, *nonne*
whole-NOM nearly sky-NOM NEG further I pursue is not
humano genere completum est?
by human race filled up
'Is not nearly the whole sky (to give no further examples) filled up with the human race?' (Cic. *Tusc.* 1.28)

4. In poetry, we can find instances where the verb precedes an interrogative word, for example (*separor* is contextually dependent): *Separor a domina* cur *ego saepe mea?* (Ov. *Am.* 2.16.41–42)
'Why am I often separated from the mistress of my heart?'

2. Second sentence-position and enclitics *enim* 'indeed', *autem* 'but' and *vero* 'however'

Any constituent whatsoever may occupy the second sentence-position, but in Classical Latin there are only three words for which this position is obligatory: the connective particles *enim* 'indeed', *autem* 'on the other hand', and *vero* 'however'.[6] In Caesar, Sallust and Cicero, these particles are excluded from the initial position in 100% of cases. Exceptionally, they are placed third (*enim* 6%, *autem* 3% and *vero* 0.5%) under well-defined conditions,[7] when the author aims to keep together a syntactic unit, for example: *difficile est enim* 'for it is hard' (Cic. *Sulla* 31), *quam multi enim* 'how many (orators)' (Cic. *Brut.* 138), *non est autem* 'but there is no' (Cic. *Div.* 2.41). These placement properties allow us to consider them as *enclitics*, i. e. words requiring a full, autonomous word placed before them. Since they resemble words bearing their own accent, they can be regarded as "non-bound" enclitics (cf. Siewierska 2004: 26 and Spevak 2006b).

It is important to define the second position occupied by these enclitics: it is the position after the first full word in a sentence, as is shown in (10). The negations *non* and *neque*, subordinators such as *si, etsi, ut, cum, ubi,* pronouns and the verb *sum* 'to be' in whatever function, including the auxiliary of passive perfect forms and the perfect of deponent verbs, also count as full words (see *est... profectus* 'is gone' in 11).

(10) Adventus *enim* L. Nasidi summa spe et
arrival-NOM indeed of L. Nasidius with utmost hope and
voluntate civitatem conpleverat.
goodwill city-ACC had filled
'For the arrival of L. Nasidius had filled the community with the utmost hope and goodwill.' (Caes. *Civ.* 2.4.4)

(11) Est *enim* profectus in Hispaniam Dexius.
be-AUX but gone to Spain Dexius-NOM
'For Dexius is gone in Spain.' (Cic. *Fam.* 7.23.4

One- and two-syllable prepositions do not behave as full words in Classical Latin. Thus, sentence enclitics do not follow prepositions but the first word of the prepositional phrase.[8]

(12) De morte *enim* ita sentit, ut...
about death indeed so he thinks that
'For his view of death is to think that...' (Cic. *Tusc.* 5.88)

6. See Spevak (2006a, Chapter 7, Section 8) and Spevak (2006b).

7. See Watt (1980), Adams (1994a), Spevak (2006a, Chapter 7, 8.3), and (Spevak 2006b).

8. The example quoted by Marouzeau (1953: 73), *post autem eum* (Cic. *Tim.* 46) 'however after him' is the only example from Classical Latin prose I know.

As they must come after the first full word, the enclitics *enim, autem* et *vero* split up noun phrases and prepositional phrases. In (13), the first full word is *ea* belonging to the prepositional phrase *in ea civitate.* It is noteworthy that in Latin even such closely related noun phrases as lexical units, for example *res publica* 'State, republic' (14), or proper names may be discontinuous. Sentence enclitics then stand after the first word and, in the case of proper names, between the *praenomen* (first name) and the *nomen gentile* (family name), for example after *Quintus* in (15).[9]

(13) In ea *enim* civitate mentio facta est.
in this indeed city mention-NOM was made
'That was the state in which the matter was mentioned.' (Cic. *Fam.* 3.8.3)

(14) Rem *vero* publicam penitus amisimus.
republic-but-republic-ACC entirely we lost
'But we have lost for ever our republic.' (Cic. *Off.* 2.29)

(15) Q. *enim* Ligarius...
Q. indeed Ligarius-NOM
'Indeed Quintus Ligarius...' (Cic. *Lig.* 2)

3. Enclitics *-que* 'and', *-ve* 'or' and *-ne*

Apart from sentence enclitics placed in second sentence-position, there is another set of enclitic elements in Latin: the additive coordinator *-que* 'and', the disjunctive coordinator *-ve* 'or' and the interrogative particle *-ne*. Whereas *enim, autem* et *vero* resemble full words in the sense that they are phonologically independent and bear accent, *-que, -ve* and *-ne* depend phonologically on their host (Allen 1973: 25). Due to this property, they can be described as "bound" clitics (cf. Siewierska 2004: 26).

The position of these enclitics depends on their function. The coordinator *-que* and the disjunctive *-ve* accompany the first constituent of a clause that is connected to another by coordination or disjunction, or the first (word of a) constituent involved in coordination or disjunction. In (16), *-que* hosted by *eam* coordinates a clause whereas *-que* tacked onto *cellam* coordinates a constituent. If the constituent is a complex phrase, *-que* and *-ve* are tacked onto the first word of the phrase, as in (17). Consequently, they split up noun and prepositional phrases, including lexical units such as *patres conscripti* 'senators' in (18) and proper names, for example *Gaius Norbanus* in (19).

9. There are three counter-examples, among them: *De Gai Gracchi autem tribunatu* (Cic. *Amic.* 41) 'But about the tribuneship of C. Gracchus', mentioned by Devine and Stephens (2006: 274).