

Inversion, V-to-C, and verb-second:

An investigation into the syntax and word order
of Old French and Late Latin

Espen Klævik-Pettersen

Department of Literature, Area Studies and European Languages
University of Oslo

This thesis is submitted for the degree of
Doctor of Philosophy

July 2018

Declaration

This dissertation is the result of my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and in the Acknowledgements.

The total word count, including footnotes, headers and references, is 126,776 words.

Acknowledgements

This thesis was conducted as part of the research project ‘Traces of History’, which was funded by grant 231431 from the Research Council of Norway. I hereby express my gratitude for the financial support that made the thesis possible.

Several people have contributed quite directly to this thesis with comments and suggestions at conferences or workshops or by taking the time to discuss various aspects of it with me one-on-one. I would like to thank my research group at the ‘Traces of History’ project for the feedback during our workshop at Ullershov in June 2016, and also for the very enjoyable days we spent together. Thanks to Christine for organizing.

My warmest thanks go to my supervisors, Christine Meklenborg Salvesen and George Walkden. Christine, thanks for the feedback on my drafts and for the discussions, for all the relevant literature recommendations, and for your Stoic calmness in the face of my procrastination. In general, thank you for making it so easy for me to work independently. George, I owe you thanks for your perceptive comments that helped improve many a section of this thesis. Your knowledge of the literature on modern and historical Germanic as well as on V2 in general has been an invaluable asset to me. I would also like to thank you for receiving me in Konstanz and for enabling me to present parts of my work there.

I was extremely fortunate to have Adam Ledgeway come from Cambridge to perform the mid-term evaluation of my thesis in May 2017, and even more fortunate that he had recently published his own paper on the *Itinerarium Egeriae*. I cannot emphasize enough how important his feedback was at a crucial moment of my research process. I owe him thanks for his detailed comments, critiques, and suggestions, many of which have been adopted in my final approach to the text.

During my time in Oslo, my social life at the campus has been intimately connected with the SynSem group. Thanks to all of you, and in particular those of you who decided that first-impressions can be deceptive and that I could only improve with time: Kjell-Johan (whose attempts at imitating my Northern Norwegian dialect improve from lunch to lunch), Atle (for sharing with me the burden of coming up with inappropriate jokes at the appropriate moment), Dag, Matthew, Alexandra, and all the rest of you. Thanks in particular to Alexander for lending me his pen in un-German fashion although we had only quite recently met.

And then there was someone else, let me see now... Yes: Elena. As you see, I am not as forgetful as you constantly claim. Thanks for all the fun moments shared during my two years in Oslo, both during the organization of the WOLP workshop and in general. It fills me with pride to see that you can now open the door on the 8th floor on your own. Also, your French pronunciation is constantly improving and must now be characterised as well above average. And might I add that you have also become a very good friend.

I must thank Dag Haug for taking the time to discuss problems related to the annotation process as well as some general issues pertaining to the chapter on Latin. Lieven Danckaert also deserves special mention for sharing material and thoughts very openly, for literature suggestions, and not least for inviting me to Lille for the opening meeting of the PALAFRA corpus. At that meeting, I met Andreas Dufter, who graciously invited me to come and present parts of my work for his Romanist colleagues at the LMU Munich. Thanks for that, and for taking so much time off to show me around the beautiful city of Munich.

From late October 2017 until April 2018 I had the opportunity to spend a semester in Konstanz. I had the most wonderful time there, and I wish to express my gratitude to Anke Hagedorn at the Welcome Center for being extremely helpful with all formalities related to

my enrolment and other practical issues. Georg Kaiser and Michael Zimmermann deserve thanks for coming to listen to my talk and for the interesting discussions and literature exchanges. Also, I would like to express my warmest thanks to my 'Biergarten' friends for including me into the group: Carmen, Katja, Marieke, Erlinde, Stefano, Alex, Andrea, Sergio: I absolutely loved our evenings at Zum Guten Hirten (and elsewhere) and I already miss the Dünnele.

Rounding off, I must take the opportunity to mention two persons who have been very important to me in many ways, both academically and on a personal level. One of them is my first Latin teacher, Synnøve de Bouvries. Your extensive knowledge of the Latin language and literature is only surpassed by your devotion towards your students. Along with very many others, I am indebted to you for the invaluable work, often performed under difficult conditions, that you have laid down over these long years in bringing the Classical culture to the North. Consider this thesis among the many fruits of those efforts.

The second person is my former French teacher at the University of Tromsø, Mona Markussen, who was also the supervisor of my master's thesis. Mona tragically passed away before time four years ago. She was the most inspiring and helpful teacher I could ever have wished for, and she also became a close friend who never ceased encouraging me to return to the academic world during my years of wandering. I am forever grateful for the friendship we shared.

Finally, I must thank my dear Karen for being so supportive over these last three years, and for all the insightful comments and ideas, several of which have contributed directly to this thesis. I am looking forward to returning the favour when you write your own thesis one day.

Some remarks on glossing

This thesis focuses mainly on word order. In consequence, the glosses will be kept maximally simple for ease of reading and are generally more translations than standard glosses, including only such morphosyntactic tags as deemed necessary. The tags that are provided loosely follow the conventions of the Leipzig Glossing Rules, but for the most part, they are heavily simplified. Case morphology is for instance generally only employed when needed to avoid confusion. Furthermore, apart from the simplification of the morphology, the glosses themselves are adapted somewhat to the actual meaning of the word in the given context, rather than a faithful word-to-word translation. One and the same word may therefore receive different glosses in different sentences.

For one of the languages under investigation in this thesis, namely Latin, the word order is so free and the morphological system so synthetic compared to English that it was found to be most practical to gloss the morphology properly, since a ‘translational’ approach to the glosses turned out to be rather more confusing than enlightening. But here as well, morphological tags are generally only used to disambiguate; in a noun phrase with adjectival modification, only the noun is glossed for case if the adjective is adjacent and no confusion is possible. A phrase like *sanctus episcopus* – ‘(the) holy bishop’ – is therefore glossed as: holy bishop-*NOM*, rather than : holy-*NOM* bishop-*NOM*. Gender tags are omitted, and number is only (occasionally in Latin) indicated on the verb, but not on the noun, where the singular/plural distinction is rather reflected directly in the gloss.

No attempt is made to follow a consistent system apart from the guiding principle of making the glosses easy to read.

Contents

1	Inversion, verb-second and Old Romance	9
1.1	Background: a brief introduction to verb second	9
1.2	More background: Old Romance and Old French	10
1.2.1	The external hypothesis: V2 as the result of language contact	12
1.2.2	The internal hypothesis: V2 as the result of organic development	13
1.3	The structure and research questions of the thesis	15
2	Theory	17
2.1	Introduction	17
2.2	V2: the basic facts	18
2.2.1	The information structure (IS) of the prefield in V2 languages	20
2.2.2	Why V2 is not just a linear constraint	21
2.2.3	The main/embedded asymmetry of V2	23
2.2.4	Linear V2 and the 'field model' of the clause	24
2.3	The 'traditional' analysis of V2	25
2.3.1	Deriving the prefield: The EPP-feature and the ban on CP-adjunction	31
2.3.2	The controversy of subject-initial clauses	34
2.3.3	Outside Continental Germanic: some problems	34
2.3.4	More problems: 'symmetric' V2 languages	40
2.3.5	Embedded V2 in non-complement clauses	43
2.4	Cartographic approaches to the Left Periphery	46
2.4.1	Cartography and V2	47
2.4.2	Verb second = V-to-C?	49
2.4.3	'Relaxed' and 'strict' V2 languages	50
2.4.4	Bottleneck approaches to V2	52
2.4.5	Feature scattering and the bundled-CP approach	58
2.5	Linear non-V2 orders in Germanic V2 languages	60
2.5.1	'V3 adverbs'	60
2.5.2	V3 with 'maybe'-adverbs	62
2.5.3	Miscellaneous other linear $V \geq 3$ orders	64
2.5.4	V1 orders and empty prefields	66
2.6	The approach adopted in this thesis	69
2.6.1	Why it is hard to define a V2 language	69
2.6.2	The definition of V2 adopted in this thesis	72
2.6.3	The general framework	74
2.6.4	Cartographic hierarchies and phrase structure	76

2.7	The methodology of the annotation	80
2.7.1	Annotated categories	80
2.7.2	The <i>bona fide</i> principle	81
3	Old French: the main clause	83
3.1	Introduction	83
3.1.1	Old French and verb-second; some preliminary remarks	83
3.1.2	The texts	86
3.2	The corpus data: Linear V2	88
3.2.1	The linear distribution of the finite verb	88
3.3	The prefield in Old French	89
3.3.1	The prefield in quantitative terms	92
3.3.2	The information structure of the prefield	95
3.4	Inversion	98
3.4.1	Foulet's generalization and the position of the null-subject	100
3.5	The acquisition of Old French phrase structure	101
3.5.1	Subject-initial strings: SVX and SpVX	102
3.5.2	Non-subject-initial V2 strings:	105
3.6	Verb-initial clauses (VX, VSX)	111
3.6.1	<i>Et-V clauses</i>	112
3.6.2	<i>Ne-V clauses</i>	117
3.7	Linear V3 strings; CSVX, CCVX, CCVSpX, SCVX	120
3.7.1	Lexically triggered V3	121
3.7.2	Left-dislocation (LD) structures	131
3.7.3	The initial subordinate clause	134
3.7.4	The formal analysis of initial subordinate clauses	142
3.7.5	Remaining V3 patterns	144
3.7.6	Foulet's generalisation and null subjects again	150
3.8	Summary	153
3.8.1	A V-to-I model for Old French?	154
4	Old French: embedded clauses	156
4.1	Linear order and the prefield	157
4.1.1	The prefield	159
4.2	Embedded inversion	161
4.2.1	Inversion in complement clauses	162
4.2.2	Stylistic Fronting	164
4.2.3	SF or V2?	166
4.2.4	Inversion in adverbial clauses	171
4.3	Embedded linear V3	173
4.3.1	Embedded linear V3 after the NPIs <i>onques/(ja)mes</i>	173
4.3.2	Embedded linear V3 in verb-second contexts	175
4.3.3	The Force-V2 hypothesis	175
4.4	A formal analysis of Old French clausal syntax	180
4.4.1	Main clause strings	181
4.4.2	Stylistic Fronting in main clauses	185
4.4.3	Embedded strings	186
4.4.4	Embedded V2 and embedded V3	189

5	Late Latin	192
5.1	Introduction	192
5.2	The transition from Latin to Romance and the nature of the evidence	193
5.2.1	Classical Latin and the ‘submerged’ spoken language	194
5.2.2	Which sources can we trust?	196
5.2.3	Itinerarium Egeriae	197
5.2.4	A worst case scenario: Christian Latin as a ‘Sondersprache’	198
5.2.5	Free word order and its consequences	199
5.3	Itinerarium Egeriae: the main clause	204
5.3.1	Linear distribution of the verb and the prefield	204
5.4	Inversion	210
5.5	The position of the subject	213
5.5.1	Adverbs	221
5.6	Itinerarium Egeriae: embedded clauses	225
5.6.1	Linear distribution of the verb	225
5.7	Head-finality in the IP	227
5.8	Inversion and the position of the subject	234
5.8.1	The position of the subject	237
5.8.2	A scrambling/operator area above IP?	246
5.9	Summary	250
6	Bridges in time	253
6.1	Introduction	253
6.2	Old French and Late Latin	253
6.2.1	Late Latin	255
6.3	From Latin to Romance . . . and Old French	256
6.3.1	The particular status of Old French	260
6.3.2	Extension and reanalysis	261

Chapter 1

Inversion, verb-second and Old Romance

1.1 Background: a brief introduction to verb second

Verb second, or simply V2, could be described pretheoretically as the requirement that the finite verb be the second constituent of declarative clauses. The verb-second phenomenon is above all associated with the Germanic branch of the Indo-European languages, within which it is almost completely generalized; with the exception of English, all the modern Germanic standard languages are V2 languages, and the same has been claimed of some German dialects spoken in Northern Italy (Grewendorf and Poletto 2011; Cognola 2013). Cross-linguistically however, it is a rare thing. Outside the Germanic family, only a handful of rather diverse and geographically scattered languages such as Breton, Estonian, Sorbian, Kashmiri, Karitiãna and some dialects of Himachali are purported to be V2 (Holmberg 2015).

As an introduction to the V2 phenomenon, consider a simple example from Modern Norwegian (1). Norwegian is an SVO language, meaning that in the unmarked word order of main clauses, the subject precedes the finite verb (boldfaced), which is linearly second (1a). This permits a linear word-for-word translation into English. In (1b) on the other hand, the direct object has been fronted to the initial position of the clause, but crucially, the finite verb still remains in second position, which is the only grammatical order here in Norwegian. Observe that this causes *subject-verb inversion*, as the subject (S) and the verb (V) swap places from SV...X to X...VS. This contrasts sharply with English, where the corresponding word order would be ungrammatical. As the translation of (1b) shows, fronting of the direct object is also possible in English (although often slightly awkward), but this does not alter the relative order of the subject and verb, which remains SV...X:

- (1) a. [*Jeg*] **har** allerede lest den boken.
I have already read that book-the
'I have already read that book.'
- b. [*Den boken*] **har** jeg allerede lest.
That book-the have I already read.
'That book, I've already read.'

While there is a lot to say about V2 on a theoretical level, this simple example serves well to illustrate the two most salient properties of verb second, namely the linear restriction on the verb which has earned the phenomenon its name (V2), as well as the *inversion* that it entails whenever a non-subject constituent is placed in initial position of the clause. This thesis will to a considerable extent revolve around such phenomena of linear order, inversion, and how to properly account for them within a formal theory of verb-second.

1.2 More background: Old Romance and Old French

The Modern Romance family is generally non-V2, with the exception of a few Rhaeto-Romance dialects that have been claimed to exhibit a V2 grammar (Poletto 2002; Anderson 2004; Kaiser and Hack 2013). This does not only mean that they have no requirement to place the verb in second position. Rather, while most Romance languages make ample use of inversion, they pattern like English in not generally allowing the word order constellation in (1b).

Interestingly, the modern situation belies a completely different scenario in the past, as written sources from the medieval period provide testimony of a widespread pattern of inversion structures parallel to (1b) in the Romance-speaking area, as the following example from Old French illustrates (notice that *ne* is a clitic and does not count for the purpose of deciding the linear order of the verb):

- (2) [*Autre chose*] *ne pot li rois trouver.*
 Other thing.OBL NEG could the king find.
 ‘The king could not find anything else.’

(*La mort Artu* (79.24), taken from Vanelli et al. 1985:166. Glosses and translation added.)

Starting with Benincà (1983), a considerable number of researchers have come to analyse nearly all the major Romance languages as obeying a V2-constraint at some point in their historical development, e.g. Old Spanish (Fontana 1993; Pinto 2011; Wolfe 2015c, but pace Sitaridou 2011) Old Italian (Vanelli et al. 1985; Benincà 2004; Poletto 2006, 2014), Old Portuguese (Ribeiro 1995; Salvi 2004, but pace Fiéis 2003; Eide 2006; Rinke 2009), Old French (Adams 1987b, 1989; Roberts 1993; Vance 1997), but pace (Kaiser 2002; Rinke and Meisel 2009), as well as non-standard Romance varieties such as Old Neapolitan (Ledgeway 2008), Old Occitan, Old Sardinian, Old Venetian or Old Sicilian (Wolfe 2015b). Some comparative studies considering different branches of Old Romance or the family as a whole have reached similar conclusions (Salvi 2000, 2004; Benincà 2004, 2006; Wolfe 2015b), but for a different view, see Sitaridou (2012). While the evidence does not command complete consensus, there is in other words a strong tradition for regarding medieval Romània as a generalized verb-second area. The sum of all of these individual studies therefore constitutes a claim, which I will refer to as the *Pan-Romance V2 hypothesis*.

As already alluded to, not all linguists agree on the proper analysis of the Old Romance textual evidence. According to some researchers, the Old Romance languages were not truly V2 languages, because the latter not seldom allowed the verb to surface in different positions of the clause, a fact that indicates that there did not really exist any requirement as such on the verb to appear in second position. This is illustrated with a V3 clause from Old Florentine (3) and a V4 clause from Old Sicilian (4):

- (3) *[Ad ogni matto] [i savi] paiono matti...*
 To each mad.person the wise.persons seem.3PL mad
 ‘To every crazy person, the wise seem crazy...’
 (Old Florentine, *Novellino* (40). Adapted from Benincà 2004:276.)
- (4) *[tamen poy di la morti loro], [li ossa loro] [pir virtuti divina] operannu miraculi.*
 Then after of the death their the bones their by virtue divine perform.3PL
 miracles
 ‘Then after their death, their bones perform miracles through divine virtue’
 (Old Sicilian, *Sanctu Gregoriu* (262). Taken from Wolfe 2015b:26)

V3 and V4 sequences of this kind would be ungrammatical in all the modern Germanic V2 languages. This shows that the Old Romance languages share some commonalities with the modern Germanic V2 languages, namely the general availability of inversion structures like (2), while at the same time lacking the linear restriction on the verb, as illustrated in (3)–(4). With respect to the latter property, there was also significant variation within the Old Romance family, as some dialects made quite frequent use of linear V3, V4 and even V>4 orders, while others, notably Old French and some Northern Italian dialects (Benincà 1983) – possibly alongside Old Spanish (Wolfe 2015b) – generally only permitted a rather restricted set of V3 orders. This variation has led to the suggestion that the Old Romance languages can be split into a group of *relaxed V2* languages, imposing no restriction on the linear order of the verb, and a group of ‘strict’ V2 languages (Benincà 1983, 2004; Poletto 2006; Wolfe 2015b). Apart from raising the question of how this distinction should be accounted for in a formal theory of syntax, this proposal also has clear typological consequences, since it would entail that the notion ‘V2 language’ is an umbrella term for different sub-groups of languages which can differ in non-trivial ways from each other.

Although Old French is often taken to be the Old Romance language whose syntax most closely resembles that of the modern Germanic languages in this respect, the proper characterisation and analysis of the language has generated a considerable debate in the literature. Starting with Kaiser (2002), a number of linguists have called into question the V2 status of Old French, claiming deviations from linear V2 reveal that it was not a V2 language, at least not in the same sense as the modern Germanic languages. (Ferraresi and Goldbach 2002; Rinke and Meisel 2009; Rinke and Elsig 2010; Sitaridou 2012; Zimmermann 2014). Other researchers have considered the linear position of the verb to be of little relevance, maintaining that it is the general availability of a particular kind of inversion structure that is central to verb-second (Benincà 2004, 2013; Wolfe 2015b). In other words, alongside the actual empirical controversy on how to interpret the facts, the debate around Old Romance inversion has turned into a (largely implicit) tug-of-war over the proper definition of the verb-second phenomenon. There is something which is undisputable, namely that the Old Romance languages displayed widespread, Germanic-like inversion of a kind which is no longer grammatical in the modern Romance languages, and then there is something which is disputable and disputed, namely whether these inversions systems should be characterised as V2 systems or not. This debate contains not one apple of discord, but two: first the empirical question of how to actually analyse the Old Romance inversion systems syntactically, and secondly the definitional question of what constitutes a V2 language.

While there is not complete consensus on the status of Old French and Old Romance inversion from a synchronic perspective, there is a general lack of knowledge of its origin and

diachronic evolution, as the research is rather heavily lopsided towards the subsequent loss of the phenomenon (Adams 1987a, 1989; Kroch 1989; Roberts 1993; Côté 1995; Platzack 1995; Vance 1997; Andrade 2018) to the neglect of its very origin within the Romance family. Part of the reason for this surely lies in the texts available for analysis, as the growing amount of prose texts from the 13th century onwards facilitates diachronic study. Nonetheless, two different hypotheses have been raised in the research literature regarding the origin of the Old Romance inversion systems. I will now take the time to briefly present these two alternative views, which I will refer to respectively as the ‘external’ and the ‘internal’ hypotheses.

1.2.1 The external hypothesis: V2 as the result of language contact

Prior to the important paper of Benincà (1983), it seems to have been implicitly assumed that Old French was unique in displaying verb-second effects. Old French inversion has repeatedly been explained in 20th century philological and linguistic literature as the result of Germanic influence. This superstrate-theory has a long and eminent pedigree in French, witness for example the words of Antoine Meillet:

... these usages ... reflect the fact that, for several centuries, men who were accustomed to speak both Latin and Germanic constructed their Latin sentences like their Germanic sentences. (Meillet 1931:37 – translation added).¹

This claim has subsequently resurfaced several times with varying degrees of explicitness (von Wartburg 1958:128, Thomason and Kaufman 1988:53,² Posner 1996:53, Vincent 2000:62, Mutz 2009:61) and has found its most recent champion in Mathieu, who suggests that Germanic influence on French was twofold, ‘first through the invasion of Gaul by the Franks, and second, by the Normans in the North-West’ (Mathieu 2009:345). Harris also cites Germanic influence as ‘a most likely factor favouring the topic-initial phase in Middle French’, but also prudently adds an important caveat: ‘Plausible though this is, it is by no means proven, since we know that there are natural iconic reasons for topics to occur initially ...’ (Harris 1984:198)

The fact that language contact can lead to syntactic change is well established in the contact literature in general (Thomason and Kaufman 1988; Harris and Campbell 1995; Lucas 2015) and in the particular domain of V2 it is conspicuous that non-Germanic V2 languages in Europe have evolved in areal contiguity and intimate contact with Germanic-speaking peoples, such as the Rhaeto-Romance dialects of Southern Tyrol (Poletto 2002) or Switzerland (Anderson 2004), the Slavonic language Sorbian, or the Finno-Ugric language Estonian. It is equally conspicuous that bilingual communities in Western Flanders show signs of deviation from linear V2 in the form of *adjunct-subject-verb*-sequences, that these seem to increase statistically with increased vicinity to the French border, and that in French Flemish, that is Flemish spoken in Northern France, V3 order in this context constitutes the rule rather than the exception (Haegeman and Greco 2016). Furthermore, in the urban

1

... ces usages ... traduisent le fait que, durant plusieurs siècles, des hommes habitués à pratiquer à la fois le latin et le germanique ont construit leurs phrases latines comme leurs phrases germaniques. (Meillet 1931:37)

²Thomason and Kaufman also claim the V2 rule was not inherited from Latin and that it is ‘unknown elsewhere in Romance’ (Thomason and Kaufman 1988:128).

vernaculars of multiethnic communities in the Scandinavian countries and Germany, V3 orders similar to those adduced by Haegeman and Greco are frequent (see Walkden 2017 and references therein). Moreover, the literature on heritage languages robustly attests that heritage speakers of V2 languages occasionally produce non-V2 orders that would be ungrammatical in the target language (Schmid 2002; Larsson and Johannessen 2015; Arnbjörnsdóttir et al. to appear; Westergaard and Lohndal to appear). The accumulated evidence therefore strongly suggests that language contact can be instrumental both in bringing about V2 grammars as well as in destabilising them.

1.2.1.1 Pan-Romance Germanic influence: a historically unrealistic scenario

The preceding observations notwithstanding, in the particular case of the *Pan-Romance V2 hypothesis* the facts as reported in the research literature do not lend themselves so readily to an explanation in terms of language contact. Although the superstrate-theory may be historically realistic for French and perhaps to some extent for the Northern Italian dialects, it is considerably less so for Spanish and Portuguese. The only lasting Germanic presence on the Iberian peninsula was that of the Visigoths, an East Germanic tribe that established their kingdom on the decline of Roman power in the 5th century and were expelled by the Umayyad conquest in the early 8th century. On traditional accounts, the Visigoths were romanized and (possibly monolingual) Latin-speakers even before crossing the Pyrenees (Harris 2000b:2, Green 2000:119, Wright 2002:30), soon converted to Catholicism, and their vernacular was never established as a language of administration. Penny concludes that ‘the influence exercised by Visigothic upon the Latin of Spain was [...] small. Apart from a number of lexical loans [...] such influence is limited to a few morphological features...’ (Penny 2002:14) This view is shared by Marías (1990:75): ‘The inhabitants of Hispania quite soon began to speak Latin; but after the fifth century they did not speak a Germanic language; rather, the Visigoths also spoke Latin.’ (See also Green 2000:118-119) In a similar vein, Rinke (2009:312) claims there was no ‘substantial Germanic influence’ on Old Portuguese,³ and Head and Seménova (2013) argue that the Germanic impact on Portuguese was restricted to the lexicon. It should be added that much of the Germanic vocabulary in Ibero-Romance is part of the common stock already found in Vulgar Latin, thus suggesting lexical diffusion through Latin rather than direct loans (Meyer-Hermann 1988, Green 2000:119, Parkinson 2000:164-165, Wright 2002:30). The importance of the Germanic influence on Neapolitan, Sicilian and Sardinian seems equally dubious from a historical perspective.

Against the backdrop of these facts, it seems highly unlikely that V2 should have passed from Germanic to all of these Old Romance varieties. This means that the contact-theory, although relevant for certain varieties, does not square well with the generalised claim embodied in the Pan-Romance V2 hypothesis. This state of affairs invites us to consider the null-hypothesis, namely that Romance V2 was the result of organic development from Latin.

1.2.2 The internal hypothesis: V2 as the result of organic development

The view that the inversion systems of Old Romance evolved out of structures already present in Latin is no less time-honoured than the theory of the Germanic superstrate.

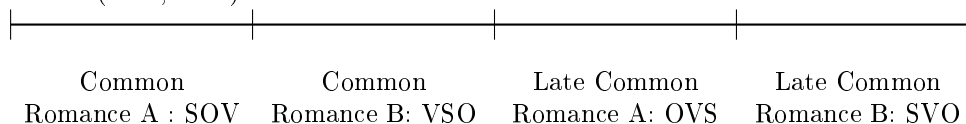
³However, as noted above, Rinke does not analyze Old Portuguese as a V2 language. She also adds that ‘Vulgar-Latin was not a V2 language.’ (Rinke 2009:312)

Rudolf Thurneysen, the philologist normally accredited with the discovery of verb-second in Old French (Thurneysen 1892), seized upon the discovery of Jacob Wackernagel’s famous law, which had been published only some months before (Wackernagel 1892), to suggest that Old French V2 had generalized from light verbs such as the copula and some auxiliaries, which often occupy second position in Latin in accordance with Wackernagel’s Law. This essentially phonological explanation of the origin of verb-second has also been reiterated occasionally (Anderson 1993; Dewey 2006).

In more recent years the internal hypothesis has been reinvoled by a handful of linguists who, to the extent that they are explicit about it, argue for a syntactic rather than phonological explanation. Without mentioning verb-second explicitly, Herman claimed that the characteristic feature of Late Latin texts was the sequence SVO or OVS, adding that ‘both of these orders seem to have gained ground statistically since Classical times, and in some texts they form the majority’ (Herman 2000:86). Unfortunately, Herman did not specify which texts he had in mind.

The internal hypothesis is also associated with the work of Dardel, who in a series of publications championed a comparative-reconstructive account of Proto-Romance word order (Dardel 1983, 1989, 1996). Categorizing word order constellations on the basis of their pragmatic, surface properties (SVO and OVS are considered two different word orders), Dardel considered the distribution of these different patterns in the oldest stages of Romance as traces of older historical stages. Based on an evaluation of the available evidence, Dardel reconstructed different historical phases, each characterised by a different unmarked word order. In Dardel (1983), these stages are *SOV*, *VSO*, *SVO*, and in Dardel (1989), an unmarked stage *OVS* is inserted in the middle, giving *SOV*, *VSO*, *OVS*, *SVO*:

Figure 1.1: Four stages in the development of Proto-Romance word order according to Dardel (1983; 1989)



A more syntactic approach is taken by Salvi, who in several publications has argued for an internal Romance genesis of V2 (Salvi 2000, 2004, 2012). Salvi points out that Classical Latin allowed fronting of the under verb certain pragmatic conditions. Apart from verb-initial orders associated with a particular illocutionary force such as questions, imperatives, hortatives and the like, fronting of the verb could also be employed in declaratives in the case ofthetic clauses where the focus scopes over the entire event (see also Devine and Stephens 2006:144-150). According to Salvi, this kind of verb-fronting was originally in complementary distribution with fronting of constituents to the left periphery of the clause under topicalisation or narrow focus. However, in Late Latin this complementary distribution was broken as verb-initial orders gradually became unmarked, yielding first *focus-verb* and subsequently also *topic-verb* sequences (Salvi 2004:101-117).⁴

One text in particular has attracted much attention in the literature on Late Latin word order, namely the late 4th century *Itinerarium Egeriae*. According to Clackson and Horrocks, there is evidence for a verb-initial pattern in the text, ‘with an optional focus slot

⁴Another claim occasionally made is that verb-fronting, particularly in later Latin, is associated with ‘lively style’ or narration (Bauer 2009:277-279).

before it' (Clackson and Horrocks 2007:291-292; Wolfe (2015) cites some suggestive passages from the text and argues that these provide indications of an incipient V2 syntax. Until recently, however, these claims were considerably underdeveloped, lacking sufficient backup from quantitative and qualitative analysis. This situation has been remedied recently by Ledgeway, who has presented a series of arguments, based on a complete annotation of the text, that the *Itinerarum Egeriae* is indeed an early specimen of verb-second syntax and a forerunner of the later Romance systems (Ledgeway 2017).⁵

1.3 The structure and research questions of the thesis

The preceding sections have served to give some relevant background to issues that will be of central concern in this thesis and to set the stage for the following chapters. Against this backdrop, I will now explain the structure of the thesis and lay out some of the most central research questions.

The thesis is organised as follows. In chapter 2, the verb-second phenomenon is introduced and explored in depth on the basis of empirical evidence from the modern Germanic languages. The chapter focuses on the historical development of verb-second theory within the mainstream generative paradigm, and in particular those more recent trends that will be most relevant to the empirical investigation and theoretical discussion in later chapters. A distinction will be made between a V2 construction and a V2 language, and concrete definitions are provided for both. The theoretical framework is clarified and some assumptions regarding the acquisition of phrase structure are presented. The methodology of the corpus annotation is briefly discussed. Some central research questions in this chapter include the following:

- (5) How should a V2 construction be defined? What role, if any, should be accorded to linear order in this definition? What role should be accorded to inversion? If the label 'V2' can be appended to languages which exhibit rather different word order properties, is the notion of V2 language really well-defined? Or to put it slightly differently: how much variation should we allow for under the label of V2 language?

On the empirical side, this thesis will contribute directly to the debate on Old French inversion on a synchronic level. This is the topic of chapters 3 and 4, which are devoted to an analysis of two Old French prose texts that have not been extensively exploited before in studies on Old French word order, namely *Le roman de Tristan en prose* and the prose version of *La vie de Saint Eustace*. Since both texts date from the first half of the 13th century and show very similar syntax, they are considered jointly as a piece of synchronic evidence for the state of the language of this period. For the ease of the reader, this bulk has been divided into two chapters; chapter 3 is devoted to main clauses, while chapter 4 focusses on the syntax of embedded clauses. The research questions can be stated in preliminary form as follows, pending more elaborate formulation:

- (6) How should the inversion system of Old French be analysed syntactically? What is the syntax of main and embedded clauses, and in what respect do they differ? (To what extent) was Old French a V2 language?

With chapter 5, this thesis aims to contribute to the issue of the origin of the Old Romance inversion systems. As a point of departure, the internal hypothesis of a development

⁵Antonelli (2015) has argued that wh-questions in the Vulgate show V2 order.

within Latin is adopted, and this hypothesis is then tested against the evidence from the late 4th century prose itinerary *Itinerarium Egeriae*. Central research questions include the following:

- (7) Had Late Latin already developed generalised subject-inversion structures of the Old Romance kind? If so, how should these structures be analysed? Had Late Latin developed into a V2 language?

Finally, the diachronic aspect will be taken into consideration in chapter 6, which rounds off by summarizing the findings from the previous chapters and by discussing their relevance to our understanding of Romance diachrony. A concrete, stagewise scenario is offered leading from Late Latin to Old Romance in general, and to Old French in particular. The possible impact of Germanic influence on Old French will also be discussed. Some important questions to be discussed in this final chapter include:

- (8) How did the Old Romance inversion systems evolve diachronically? What might be considered plausible stages in this process? In what sense is Old French special within the Romance family, and how do we account for this special status?

Having stated my research objectives, I would like to express my hope that this work can be of interest outside the circle of theoretically oriented linguists. In particular, I hope that philologists within both the Latin and the Romance tradition and indeed anyone interested in the diachrony of the Romance languages might find it worthwhile to read. In order to make the text as accessible as possible, I have chosen to include in chapter 2 a rather lengthy introduction to the verb-second phenomenon which does not require extensive background knowledge. The empirical data as well as all the relevant theory are introduced, in addition to some thoughts of my own on the proper way to delimit V2 from similar inversion systems. Some background knowledge of syntactic theory is assumed; for an accessible introduction to the Minimalist framework used in the analysis, see Adger (2003).

Chapter 2

Theory

2.1 Introduction

This chapter serves several purposes. First, it introduces the major facts of verb-second as the phenomenon manifests itself in the modern Germanic V2 languages. This gives us a solid empirical basis for comparison when confronting the evidence from the inversion system of 13th century Old French in chapters 3-4. Such comparison with Germanic is often made in work addressing the Old Romance inversion systems. This practice of using modern Germanic as a benchmark is quite understandable, as it would be senseless for a corpus linguist considering a V2 hypothesis for a dead language to deprive herself of the potential insights offered by contemporary languages whose syntax is much better understood. At the same time, this practice can run the risk of becoming somewhat less than methodical if the dead language is expected to behave in fashion identical to the ‘control group’, and in particular if this assumption stays implicitly in the background rather than being explicitly stated.

Furthermore, the modern Germanic V2 languages are far from homogeneous with respect to all aspects of the verb-second syntax. The second objective of this chapter is therefore to clarify what is common and what is not, and to separate what is central to a verb-second grammar from what is strictly speaking extrinsic, in order to develop a theoretically explicit definition of verb-second which ultimately will serve as the standard against which to evaluate the data in this thesis. When armed with such a language-neutral definition, the possibility of comparison offered by the modern Germanic languages may serve as a very useful heuristic in later chapters.

In general terms, the present chapter aims to provide an overview of the research literature on verb-second within the mainstream generative paradigm (read: transformational grammar based on the T-model), and to problematize certain aspects of this research. Various empirical phenomena which pose problems to prevailing theory will be pointed out and discussed. In particular, this chapter makes a novel theoretical contribution by showing that some theoretical innovations of recent years are more problematic than what is perhaps commonly assumed, but no serious attempt will be undertaken to solve these problems. The reason for this is that the primary objective of this thesis is not to develop new theory, but to arrive at an adequate account of the syntax of Old French and Late Latin, as well as the diachronic links that unite these languages. In this perspective, the Germanic V2 languages discussed in this chapter play only a secondary role, setting the ground, as it were, for the

investigations in subsequent chapters.

The structure of the chapter Rather than trying to keep data and theory strictly apart, an approach which would probably have led to much repetition, I have opted for a loosely ‘historical’ structure, tracking the development of verb-second theory as it has evolved in the face of ever-growing evidence and theoretical innovations. The literature on verb-second is imposing, and the emphasis here is placed on those recent developments that will play the most direct role in the following chapters.

The structure of the chapter is thus as follows: section 2.2 provides a basic introduction to the verb-second phenomenon through some examples, a first attempt at a definition is offered and then immediately rejected, and a brief illustration is given of an important pre-generative forerunner of later verb-second theory, namely the field model (*Feldermodell*) approach to German word order.

Section 2.3 introduces what will be termed the *traditional analysis* of verb-second and spends some time illustrating its benefits and shortcomings. Section 2.4 introduces recent cartographic approaches to verb-second in general and so-called ‘bottleneck’ approaches in particular. It will be argued that the nature of Germanic V2 eludes a satisfying description in a fully cartographic model. In section 2.5, I will review some of the most frequent deviations from the linear V2 pattern found in the Modern Germanic verb-second languages. Section 2.6 establishes the definition of V2 that will be used in this thesis and clarifies the theoretical framework and assumptions that will be adopted.

Since chapters 3 – 5 will provide data retrieved by manual corpus annotation, section 2.7 briefly explains the methodology employed in the annotation process.

2.2 V2: the basic facts

As an introduction to the V2 phenomenon, let us start out by considering some examples from Modern German.¹ The practice established here will be continued throughout the thesis: the finite verb is boldfaced, and the constituents in the prefield (the area preceding the verb) are enclosed in brackets. When necessary, other means of highlighting such as underlining will be employed; this will be properly indicated at the relevant time.

Example (9) gives five slightly different variants of a German main clause. Notice that the finite verb remains in second position throughout all the permutations and functions as the axis around which the other constituents of the clause group. The area to the left of the verb will be called the ‘prefield’.² There are few restrictions on the category and grammatical function of the first constituent, which can be, among other things, the subject (9a), the object (9b), a locative adverbial PP (9c), a temporal adverb (9d) or even an entire embedded clause functioning as an adverbial (9e):

- (9) a. [*Ich*] **habe** *gestern* *das Buch* *im* *Auto* *vergessen*.
I have yesterday the Book in-the car forgotten.
‘I forgot the book in the car yesterday.’
- b. [*Das Buch*] **habe** *ich* *gestern* *im* *Auto* *vergessen*.
The book have I yesterday in-the car forgotten.

¹Unless otherwise indicated, the examples are my own. In case of doubt native speakers were consulted.

²In this thesis, the term ‘prefield’ is used as a descriptive surface term to designate the area to the left of the finite verb. From this follows that the prefield can in principle range from being empty (in verb-initial clauses) to comprising the entire clause (in verb-final clauses).

- c. [*Im Auto*] **habe** ich gestern das Buch vergessen.
In-the car have I yesterday the book forgotten.
- d. [*Gestern*] **habe** ich das Buch im Auto vergessen.
Yesterday have I the book in-the car forgotten.
- e. [*Weil ich so gestresst war*], **habe** ich das Buch im Auto vergessen.
Because I so stressed was, have I the book in-the car forgotten.
'Because I was so stressed, I forgot the book in the car.'

While all of the major categories and functions illustrated in (9) are permitted in the prefield across the range of Germanic V2 languages, this does not mean that there are no restrictions whatsoever; discourse particles are generally not possible in the prefield (Thráinsson 2007:40),³ and verbal particles are contextually severely limited, although far from impossible (Trotzke and Quaglia 2016). Other restrictions are particular to some Germanic languages. For instance, German does not accept simple sentential negation in the prefield, while this option is possible in certain contexts in some of the modern Scandinavian languages (Holmberg and Platzck 1995), although apparently not in Danish (Mikkelsen 2010:4). While Continental and Scandinavian Germanic can front the entire VP to the prefield, this is generally not possible in Icelandic (Thráinsson 2007:349). The intention here is not to catalogue the actual variation between the Germanic languages in this respect, merely to point out that such variation exists.

As for the quantitative dimension of this variation in the prefield, there does not seem to be sufficient data available for all of the languages to draw firm and detailed conclusions. However, all V2 languages pattern similarly with respect to the most frequent fillers, which are subjects or adverbials, the latter of course being a cover-term for several different types of constituents. In a study on word order in modern German newspaper prose, Fabricius-Hansen and Solfeld (1994) found that 38.6% of main clauses feature an initial adverbial, while 6.6% feature an object (N=984). In her investigation of Swedish prose, Westman (1974) reports 30.8% initial adverbials and 2.3% objects (N=5588). A comparative investigation by Bohnacker and Rosén (2008) finds that the prefield in Swedish is filled by adverbials in 23% and objects in 3% of main clauses (N=535), while the corresponding numbers for German are 42% and 7%, respectively (N=1173). There therefore seems to be some evidence for claiming that Germans front non-subjects, and in particular objects, more frequently than Swedes (and as a conjecture perhaps speakers of Mainland Scandinavian in general). Walkden and Booth searched the Icelandic corpus IcePaHC and found 40.8% 'non-subject, non-object' constituents and 2.9% objects in the prefield (Walkden and Booth to appear). While we must expect style and genre to exert an influence on this distribution in all languages, the relative difference between German and Swedish could also

³Notice that in German, it is possible to front elements like 'doch' – ≈ 'still/yet' in the concessive sense – to the prefield:

- (i) *Und [doch] wird das Licht der Gottlosen erlöschen...*
Und yet will the light of-the godless extinguish...
'But the light of the godless shall die out...'
(*Book of Job* 18:5)

The question is whether 'doch' should be considered a particle in such cases or rather an adverb; it can clearly be prosodically accentuated in the prefield, and the semantics seems somehow richer and more concrete than what is the case with the corresponding middle field particle 'doch'. However, this is not decisive evidence against the status of 'doch' as a particle, since these properties could equally well be afforded by the position in the prefield itself.

be indicative of differences related to information-structure (hereafter also IS) and the use of the prefield in this respect, an issue we will now briefly consider.

2.2.1 The information structure (IS) of the prefield in V2 languages

The sentences in (9) are not identical in terms of meaning. Since they differ only with respect to what constituent is found in the prefield, it is clear that the prefield offers some pragmatic possibilities that can be exploited in discourse. In subject-initial clauses, the subject by default carries an aboutness topic interpretation, to which the rest of the clause provides the comment. The fronting of a non-subject constituent creates a (sometimes rather slight) discourse effect, the exact IS-value of which is subject to some variation across V2-languages (Holmberg 2015). In the Germanic languages, the fronting⁴ of a non-subject argument most commonly maintains a topic reading, which may or may not be contrastive. Cases like (9d), where a temporal adverbial occupies the prefield, can either be interpreted as a topic, or perhaps more plausibly as an element providing some kind of anchoring or *scene setting* for the rest of the clause, possibly a sub-case of topicalisation. It is worth noting that such clauses are also experienced as relatively unmarked, on a par with subject-initial clauses (see Frey 2004a for a formal analysis of this ‘unmarkedness’):

- (9d) [*Gestern*] *habe ich das Buch im Auto vergessen.*
 Yesterday have I the book in-the car forgotten.
 ‘Yesterday, I forgot the book in the car.’

Foci are also possible in the prefield with contrastive or corrective readings (10a). However, it seems like information focus, understood as non-contrastive, non-presupposed new information (128), is not particularly felicitous in the prefield of Germanic V2 languages (cf. Frey (2006):⁵

- (10) a. *A: Har du sett filmen om dronning Elisabeth? B: Nei. [BOKEN] har jeg lest, men jeg har ikke sett filmen.*
 Have you seen film about queen Elisabeth no book-the have
 I read but I have not seen film-the.
 ‘Have you seen the film about queen Elisabeth? No. I have read the BOOK, but I haven’t seen the film.’
- b. [*I går*] *var en hyggelig kveld. ?? [En film] så vi alle sammen.*
 In yesterday was a nice evening. A film saw we all together.
 ‘Yesterday was a nice evening. We all watched a film together.’
 (Modern Norwegian)

In spite of these general similarities, some subtle nuances between the different Germanic V2 languages with respect to the prefield have been documented in the literature.

⁴Note that XP-fronting to the prefield is in fact often referred to as *topicalisation*, regardless of the pragmatic value of the initial constituent. Used in this sense, it is possible to ‘topicalise’ foci and adverbials. I shall try to avoid this ambiguous term, reserving it for cases of true topicalisation where the initial XP is in fact a topic. In general I will use the term XP-fronting, although I stress that such terminology should not be interpreted as a staunch commitment to a derivational theory.

⁵Throughout this thesis, little or no attempt is undertaken to reproduce the information structure of the V2 languages into the English translations. This is simply because the resulting structures often seem quite marginal in English. The difference between English and the Germanic V2 languages in this respect is far greater than the difference between the latter languages themselves.

Bohnacker and Rosén have shown that native speakers of Swedish do not place rhematic (new) information in the prefield as easily as Germans, and that this native preference is transferred by Swedes when acquiring German as a second language (Bohnacker and Rosén 2008). For instance, the Swedish L2 acquirers showed a tendency to use expletives in the prefield to demote rhematic information to a postverbal position in contexts where German native speakers preferred preverbal rhematic information. In a subsequent study, Bohnacker (2010) finds that the inverse pattern also holds, as German L2 acquirers of Swedish tend to ‘overuse’ the prefield. This difference at the level of information structure must therefore be assumed to be at least partially responsible for the different quantitative distribution of elements in the prefield in the two languages (see section 2.2) above).⁶

Although there is need for more work in this area, the evidence suggests that there are differences between V2 languages with respect to information structure which have both qualitative and quantitative impacts on the use of the prefield, and that these differences may cause some interference effects for Germanic speakers in L2 acquisition.

2.2.2 Why V2 is not just a linear constraint

Since the verb stays in second position no matter what is put in the prefield, V2 thus produces surface inversion whenever a non-subject constituent occupies the prefield.⁷ A particular feature of this inversion structure is that, in the case of periphrastic or compound tenses, the subject surfaces between the finite auxiliary and non-finite main verbs, rather than following both of these, as is the case in modern Romance inversion structures. This is sometimes reflected in the term ‘Germanic inversion’ (Adams 1987a), but rather than using this term, I will follow Poletto (2002) and refer to such constructions as *G-inversion*.⁸

Now, if two constituents appear before the verb as in (11), the result is strongly ungrammatical in all the Germanic standard languages (although they are found in certain non-standard varieties; see Walkden 2017) :

- (11) **[Gestern], [ich] habe das Buch im Auto vergessen.*
 Yesterday I have the book in-the car forgotten.

Since the prefield is restricted in this way, let us as a first approximation define verb-second as a simple linear word order constraint (albeit one that recognizes constituency):

- (12) **Verb-second (preliminary definition):** The finite verb must appear as the second constituent of all declarative main clauses.

⁶Bohnacker’s (2010) study also shows that this discrepancy between the L2 and L1 use of the prefield improves with time, as L2 speakers increase their proficiency.

⁷Much like the term *prefield*, the term *inversion* will also be employed in a descriptive sense to refer to any surface word order where the finite verb appears to the left of the subject, regardless of the actual structural position of either the verb or the subject. Used this way, the term even includes the unmarked word order of VSO languages.

⁸The reason for this is mainly to avoid the two following implications:

- (i) That such inversion structures are the result of Germanic influence.
- (ii) That such inversion strings necessarily have the same syntactic structure as in modern Germanic.

In other words, *G-inversion* is yet another surface term, referring to the particular inversion structure where the subject surfaces between the finite and non-finite forms of the verb. It is undisputable that the Old Romance languages featured widespread G-inversion in this sense of the word, but clearly disputable whether (i) or (ii) holds. Providing answers to these questions is of central concern to this investigation.

This definition is insufficient in many respects. In particular, the position of the finite verb in the linear surface form of a sentence is neither a necessary nor a sufficient condition of a V2 grammar from a theoretical point of view. One of the the main objectives of the following sections is to illustrate why this simple and intuitive definition of V2 in terms of linear order does not capture the essence of the phenomenon, and to replace it with another notion of V2, defined in terms of syntactic structure, that is theoretically and empirically more adequate. Consider to this end the following contrast between German and the only Germanic non-V2 language, modern English:

- (13) a. [My mother] **works** at the hospital.
 b. [*Meine Mutter*], [*die*] *arbeitet im Krankenhaus.*
 My mother, she works in-the hospital.

Here, the non-V2 language English puts the verb in second position, while the V2 language German allows it to surface in third position. In other words, a distinction must be made from the very outset between *linear V2* and *structural V2*. Many researchers maintain that only structural V2 holds any theoretical importance. I will also rally behind this position, but not without making two important qualifications.

First, it is imperative to bear in mind that linear order has an epistemological priority over structure, as it is *linearly ordered strings* that serve as input to children during acquisition. Without denying that children have access to other cues such as morphology or prosody as well, it follows from this observation that the researcher carries the responsibility of rigorously demonstrating how a structural V2 grammar can be acquired on the basis of linear surface input. My exact stance on these questions will be clarified in section 2.6.4.

Second, I will argue that an appropriate theoretical notion of V2 should be defined in such a way that structural V2 regularly produces linear V2, although it is impossible to define in quantitative terms exactly how tight this bond should be. While it is indeed imaginable that a given sample from a non-V2 SVO language such as Modern English may in fact contain more linear V2 than a sample from a V2 language like German, as was demonstrated in (13), this will rarely be the case, and the longer the sample, the less likely this scenario will be. As will become clear later in this chapter, recent developments in verb-second theory have strained the bond between linear and structural V2 considerably.

But this is getting ahead of our story. Since nothing has been said so far about structural V2, we continue to focus on linear V2. Observe at this point that from the perspective of linear order, the V2 string is by definition *CVX*; in other words, a single constituent followed directly by the verb and then the rest of the clause. Canonically, the V2 string is *CVSX*, which is of course a subtype of *CVX* and entails that the subject is directly postverbal in the surface linear order. However, this adjacency does not hold in full generality, as the subject may in fact be separated from the verb by intervening constituents. In (14), the finite verb and the subject are separated by the adverb *leider* – ‘unfortunately’ – producing a string *CVCSX*. I will follow Vance (1997) and refer to such cases as *non-contiguous inversion*:

- (14) [*Den Film*] **hat leider** *niemand gesehen.* (*German*)
 The film has unfortunately nobody seen.

‘Unfortunately, nobody has seen the film./As for the film, unfortunately nobody has seen it.’

An important lesson to bring from such examples to the corpus data in the following chapters is that while the string *CVSX* is highly characteristic of V2, it is neither a necessary nor sufficient condition for diagnosing structural verb-second.

Non-contiguous inversion is particularly common in the Scandinavian languages, as illustrated by the following, completely unmarked sentence from modern Norwegian:⁹

- (15) [*Brevet*] **har** *dessverre sannsynligvis ikke faren min mottatt*.
 Letter-the has unfortunately probably not father-the mine received.
 ‘My father has unfortunately probably not received the letter.’¹⁰
 (Standard Norwegian)

Here the subject is separated from the finite auxiliary by no less than two adverbs plus negation, although it could be argued that the latter adverb modifies the negative phrase. It should be noted that, across Germanic, the adjacency between verb and subject seems to be stronger with pronominal subjects, as the contrast between (14) and (16) illustrates. These examples show that full DP subjects may often occupy a lower position in the clause than pronouns:

- (16) a. **[Den Film] hat leider sie nicht gesehen*.
 The film has unfortunately she not seen.
 b. [*Den Film*] **hat** *sie leider nicht gesehen*.
 The film has she unfortunately not seen.
 ‘As for the film, unfortunately she has not seen it.’

2.2.3 The main/embedded asymmetry of V2

Before we proceed to theory, let us observe a final ‘major fact’ of Germanic V2, the so-called ‘main clause/embedded clause asymmetry’, again illustrated by German:

- (17) a. *Ich bereue, dass [ich] [das Buch] [im Auto] [vergessen] habe*.
 I regret that I the Book in-the car forgotten have.

⁹Non-contiguous inversion is apparently widespread in Germanic, but the conditions governing its application are complex and vary somewhat from language to language. In German, the equivalent of the Norwegian example in (15) is ill-formed, since non-contiguous inversion featuring low subjects preferably happen with quantified subjects, as in (14). The definiteness of the subject is another impacting factor, the crude generalization presumably being that definite subjects prefer ‘higher’ positions than indefinite ones; see for instance Thrainsson (2007:47-58) for a discussion of these facts in Northern Germanic, with a particular emphasis on Icelandic. Observe also that the example in (150) might be more acceptable if the subject pronoun is strongly stressed, thereby forcing a contrastive reading.

¹⁰Interestingly, the same sentence is perfectly acceptable with the subject in higher position. Thus, in addition to (15), all the following seem fine:

- (i) a. *Brevet har* *dessverre sannsynligvis [faren min] ikke mottatt*.
 b. *Brevet har* *dessverre [faren min] sannsynligvis ikke mottatt*.
 c. *Brevet har* *[faren min] dessverre sannsynligvis ikke mottatt*.

If one adopts the cartographic Principle of Transitivity and assumes that the position of the adverbs are fixed, we are forced to conclude that the subject may climb from the lowest to the highest position in this hierarchy, landing any place on the way, without noticeable semantic effect. We have a strong candidate for optionality in grammar here, then.

‘I regret that I forgot the book in the car.’

- b. *Ich bereue, dass [ich] **habe** das Buch im Auto vergessen.

As the ungrammaticality of (17b) shows, V2 is generally ruled out in embedded clauses in German. While this main-embedded asymmetry is particularly conspicuous in Continental Germanic SOV-languages like German or Dutch, the same discrepancy between main and embedded clauses is observable in the Scandinavian SVO-languages, since the finite verb in embedded clauses follows negation and sentence adverbs. This is illustrated for Modern Norwegian in (18). In the main clause in (18a), the direct object occupies the prefield and there is subject-verb inversion, illustrating that verb-second is operative in Norwegian. Example (18b) shows the correct word order in embedded clauses; notice that the sentential negator *ikke* – ‘not’ – precedes the finite verb. Linear V2 is ungrammatical in this case, whether the embedded clause is subject-initial (18c) or inverted (18d).

- (18) a. [*Boken*] **kjøpte** jeg ikke.
Book-the bought I not.
‘The book, I didn’t buy.’
- b. *Jeg* angrer på at [*jeg*] [*ikke*] **kjøpte** boken.
I regret on that I not bought book-the.
‘I regret that I didn’t buy the book.’
- c. **Jeg* angrer på at [*jeg*] **kjøpte** ikke boken.
I regret on that I bought not book-the.
- d. **Jeg* angrer på at [*boken*] **kjøpte** jeg ikke.
I regret on that book-the bought I not.

There is a lot more to say about V2 on the empirical level, but these additional facts will be addressed in the next sections in tandem with relevant theory. As a preliminary summary, we observe that a felicitous theory should account not only for linear V2 per se (95), but also why it is sometimes still possible to have linear V3 (cf. the contrast between (11) and (13b)), as well as the main-embedded asymmetry (17–18).

2.2.4 Linear V2 and the ‘field model’ of the clause

As the evidence considered so far has shown, the V2 phenomenon is a salient feature of the Germanic languages, as reflected in the term ‘Germanic inversion’. Unsurprisingly then, verb-second has been recognized for a long time in Germanic philology (see for instance Erdmann (1886) for a quite modern description of the facts). In the topological model of the German clause, the so-called *Feldermodell* developed by Eric Drach (Drach 1963/1937) and elaborated further for Mainland Scandinavian languages by Paul Diderichsen (Diderichsen 1966/1944) the finite verb constitutes the left sentence bracket (*linke Satzklammer*) in main declaratives, and the domain in front of it, the prefield or *Vorfeld*, is restricted to a single constituent. This amounts to a word order constraint. The example in (9b), repeated below in slightly simplified form (dropping the PP adjunct) as (19), can be represented as follows in the topological model:

- (19) [Das Buch] **habe** ich im Auto vergessen.

Prefield	Left Bracket	Middle Field	Right Bracket
Das Buch	habe	ich im Auto	vergessen.

While the topological model has enjoyed great success in didactics and second-language acquisition and its terminology has become commonplace even in generative syntax, it has figured less prominently in the research literature on V2. A notable exception is Povl Skårup's important book on Old French syntax (Skårup 1975).

The drawback of the model is that it gives limited information on the structure of the clause and no information at all about its hierarchical organisation. It is descriptively successful in providing a word order constraint, but it does not explain why this constraint should hold or what it derives from. As for cases where two constituents are in fact allowed to precede the finite verb (cf. 13a), the topological model must simply add a 'pre-prefield' (*la zone annexé* in Skårup's terminology) to accommodate the extra constituent.

Nonetheless, when considering that the left sentence bracket corresponds to the finite verb in main clauses and to the complementiser/subjunction in embedded clauses, one cannot help but feel impressed by the striking parallels to later generative analyses. The traditional analysis presented in the next section already has a solid fundament; in this respect, Drach's work on the field model deserves to be recognized as the first major breakthrough in verb-second theory.

2.3 The 'traditional' analysis of V2

In the preceding sections I have referred to V2 both as a 'construction', a 'word order constraint', and 'a rule'.¹¹ All of these terms are commonplace in the literature on V2, for obvious reasons. But while they are useful descriptive labels, transformational generative grammar has been reluctant to grant them much theoretical status. This most obviously applies to the notion of 'construction', which is often seen as little more than a rather loose surface descriptive term.¹²

Work in generative syntax in the 1970s (Williams 1974; Koster 1975; Thiersch 1978) lead to a breakthrough in the form of the analysis of den Besten (den Besten 1983), which has subsequently been adopted as the standard analysis of verb-second in generative syntax. The success of this analysis is that it defines the structural properties of V2 precisely, thereby

¹¹I would like to emphasize that what is termed the 'traditional analysis' in this section is not a single concrete analysis provided by an individual researcher or even a small group of researchers, but rather represents a fusion of different analytical contributions as well as more general theoretical insights developed in the late 70's and the 80's, including, but not limited to Williams (1974), Koster (1975), Thiersch (1978), den Besten (1983), Travis (1984), Platzack (1986), deHaan & Weerman (1986), although the analytical 'core' of this analysis must be considered the work of den Besten (1983). The presentation in this section therefore takes the historiographic liberty of presenting a simple synopsis of a complex research story from a retrospective point of view.

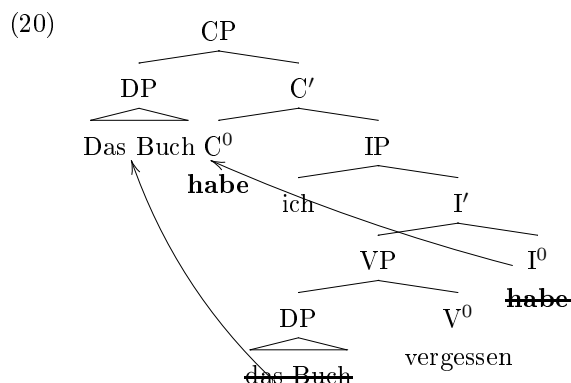
¹²As for rules/constraints, they were common in early transformational grammar but have since fallen into disrepute. Thus, in the early days of generative grammar, while PS-rules were still used to create Deep Structure from the lexicon, a central issue to research on German syntax was to decide whether the main clause V2 order (in particular the unmarked SVO order) or the embedded verb-final order should be considered the basic, underlying order; for arguments in favour of the former, see Vennemann (1972), and for the latter, Reis (1974). The latter position prevailed (for a discussion, see Thiersch 1978) and has enjoyed widespread consensus in derivational approaches to German syntax ever since, although a third position has emerged with a modern variant of the 'Universal Base Hypothesis' (Kayne 1994), according to which all word order patterns are derived from a universal SVO base. This assumption is adopted for Zwart's early Minimalist study of Dutch (1993).

reducing terms like ‘word order constraint/rule’ or ‘construction’ to epiphenomenal surface terms.

The centerpiece of den Besten’s analysis rests upon the main/embedded asymmetry of V2 illustrated in the previous section. Den Besten capitalized on the complementary distribution of complementisers and verb-second in Continental Germanic by claiming that the verb and the complementiser compete for the same position, COMP. This position is vacant in main clauses, allowing the verb to move into it, while it hosts the complementiser (or subordinations) in embedded clauses, thereby blocking verb fronting.

With the extension of the X-bar schema (Chomsky 1970; Jackendoff 1977) to functional categories (Chomsky 1986), the formerly exocentric clause was recast as a CP. Updating den Besten’s analysis to this new endocentric phrase structure, the structure of a sentence such as (19) is thus as follows: ¹³

- (19) [*Das Buch*] **habe** *ich* *vergessen*.
 The book have I forgotten.
 ‘I forgot the book.’

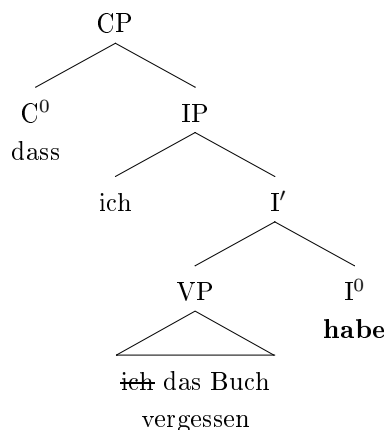


The claim embodied in this figure is that verb second is in fact a complex phenomenon consisting of two different movement operations: (1) Movement of an XP, in this case the direct object, from the core clause into the specifier of CP. (2) V-to-C movement of the finite verb to the head of the CP projection. In embedded clauses, on the other hand, the presence of an overt complementiser lexicalising C⁰ effectively blocks verb raising, and the verb stays lower in the structure, the exact position being open to variation across languages.

- (21) *Ich bereue, dass [ich] [das Buch] [vergessen] habe.*
 I regret that I the Book forgotten have.
 ‘I regret that I forgot the book.’

- (22) **Ich bereue, dass [ich] habe das Buch vergessen.*

¹³Note that this structure is minimal and only shows the core projections CP, IP, VP. It abstracts away from complications such as the question if the subject originates lower in the structure, for instance in SpecVP, which is omitted here. It also assumes that there is a head final IP-projection in German, although the empirical evidence for this projection is scant if existent. Nothing hinges on this.



V-to-C movement was later taken to be an instantiation of the more general operation of Head Movement (Travis 1984; Koopman 1984; Baker 1988) and considered to be subject to the Head Movement Constraint (Travis 1984), a locality constraint stating that a head X cannot move to a head Y by skipping an intervening head Z. In other words, head movement is considered to be cyclical. However, while there is strong evidence for cyclical head movement in English,¹⁴ the question is rather complicated in many other Germanic languages. In German, there is little if any empirical evidence that would allow us to decide whether the final verb in embedded clauses is in I⁰ or V⁰, as such movement would always be string vacuous.¹⁵ In Mainland Scandinavian, as we have already seen (cf. example (18) above), the finite verb is preceded by negators and various adverbs in embedded clauses. Since these elements are generally taken to occupy positions between IP and VP in these languages, this is commonly interpreted as direct evidence that the verb does not raise to I⁰ in non-V2 embedded clauses (Platzack 1986; Holmberg and Platzck 1995). If the verb does not move to I⁰ in non-V2 contexts in these languages, this raises the question why it should do so as an intermediate step in V-to-C movement, or how this intermediate step might ever be acquired by the children.¹⁶ I will return to the question of V-to-I later, as this is an important and still not completely resolved issue in Scandinavian syntax as well as in V2 theory in general.

2.3.0.1 Evidence in favour of the traditional analysis

There are several pieces of evidence to suggest that the hypothesis of a competition between finite verbs and complementisers might in fact be on the right track. First, observe that

¹⁴Crucially, the fact that only auxiliary verbs are acceptable in C⁰ in the context of polar questions and ‘residual V2’ constructions such as *wh*-questions, the very same verbs that are independently shown to undergo V-to-I movement. In other words, V-to-I ‘feeds’ I-to-C in English (for a discussion, see Roberts 1993:15-16).

¹⁵In fact, the matter is even more complicated. Basically, there are three options regarding the IP-projection in Modern German: 1. It is absent (Haider 1993). 2. It is head-final, like depicted in (20) (Grewendorf 1993). 3. It is head-initial, in which case it would have to be assumed that the verb does not raise at all in embedded clauses, like in (most varieties of) Mainland Scandinavian (Haider 2010; Vikner 2001).

¹⁶Platzack (1986) argues that V-to-C necessarily involves the intermediate step V-to-I and that both of these derivational moves must be acquired independently, while Holmberg & Platzack (1995) and Vikner (1995) suggest that V-to-C might take place in one fell swoop.

embedded V2 in German is triggered by dropping the complementiser (23a), since verb-final order (23b) is ungrammatical in such cases:

- (23) a. *Er hat gesagt, das [Buch] hat/habe er vergessen.*
 He has said the book has.IND/SUBJ he forgotten
 'He said that he forgot the book.'
- b. * *Er hat gesagt, er das Buch vergessen hat.*

Still, it is important to point out that embedded verb-second in Continental Germanic is not governed exclusively by the presence or absence of the complementiser, as one would expect from the traditional analysis. There are many verbs which quite simply do not accept embedded verb second, complementiser or not (Haider 1986:53):

- (24) a. *Ich bereue, dass [ich] [das Buch] [vergessen] habe.*
 I regret that I the Book forgotten have.
 'I regret that I forgot the book.'
- b. * *Ich bereue, dass [ich] habe das Buch vergessen.*
- c. * *Ich bereue, _ [das Buch] habe ich vergessen.*

In other words, embedded V2 in complement clauses in Continental Germanic is constrained by both the nature of the matrix verb as well as the presence or absence of the complementiser. This leads to a situation where embedded V2 is either impossible (if either of the two conditions *appropriate matrix verb* and *zero complementiser* is violated) or obligatory (if both conditions are met). As we shall see later, this picture is slightly caricatural, but it presumably captures the vast majority of cases.

Another piece of evidence sometimes adduced in favour of the traditional analysis comes from conditional adverbial clauses, which also feature V-to-C movement in the absence of a subjunction. This phenomenon seems to be found in all Germanic languages. In (25a), the subjunction lexicalises C⁰ and the verb stays in clause-final position, whereas in (25b) the verb assumes the clause-initial position in C⁰ in the absence of a subjunction; the paradigm

in (26) illustrates that the same applies to Icelandic.^{17 18}

- (25) a. *[Wenn] du heute Abend zu mir kommst, können wir uns einen Film anschauen.*
 If you today evening to me comes can we us a film
 on-look
 ‘If you come over to me tonight, we could watch a movie.’

b. **Kommst** du heute Abend zu mir, können wir uns einen Film anschauen.

- (26) a. *Jón verður góður ef hann æfir sig.*
 John becomes good if he practices self.
 ‘John will be good if he practices.’
- b. *Æfi Jón sig verður hann góður.*
 practise.SUBJ John self becomes he good.
 ‘If John practices, he will be good.’
- c. *Jón verður góður, æfi hann sig.*
 John becomes good practice.SUBJ he self.
 ‘John will be good if he practices.’

(Icelandic, taken from Thráinsson 2007:30.)

While there are clearly some indications that verbs and complementiser are attracted to the same structural position, the deeper reasons behind this correlation are not immediately

¹⁷Interestingly, the Germans I have consulted do not accept the variant where the conditional is placed after the matrix clause, that is the equivalent of the Icelandic (26c):

- (i) ?? Wir können uns einen Film anschauen, **kommst** du heute Abend zu mir.

One way of interpreting this is that even in German, where the complementarity between V2 and complementisers is generally very robust, it is not sufficient to just vacate the C-position in order to have V-to-C movement (but see Reis and Wöllstein (2010) for an analysis and a different conclusion). Another interesting question is why the verb is not interested in the presumably empty C⁰ in the case of relative clauses or embedded interrogatives. One might appeal to the ‘Doubly-filled COMP filter’ (Chomsky and Lasnik 1977:446), but there is sufficient cross-linguistic evidence to question the general validity of that notion. In fact, several Germanic varieties do optionally allow the filling of both the head and the specifier, but in these varieties, it is invariably the complementiser and not the verb that lexicalises C⁰ (Bayer 1984; Haegeman 1992; Schönenberger 2006). In Frisian, it is apparently the case that the ‘Doubly-filled COMP filter’ must obligatorily be violated (Hoekstra 1993). Similar claims are made for some Northern Italian dialects by Poletto (2000). In a cartographic approach to the LP (see section 2.4), the failure of the verb to raise in such contexts is even harder to explain, if one assumes with Rizzi (1997) that relative and embedded interrogative pronouns move to ForceP, since this opens up almost the entire LP, with neither Fin⁰ nor SpecFinP lexicalised. On the other hand, it should be mentioned that some German varieties do in fact feature V-to-C in relative clauses, see section 2.3.5.

¹⁸The same phenomenon can also be observed in adverbial clauses introduced by ‘als ob’ – ‘as if’ – in German:

- (i) a. *Es sieht aus, als ob es regnen wird.*
 It looks out, as if it rain will.
 ‘It looks as if it will rain.’

b. *Es sieht aus als wird es regnen.*

clear. Den Besten suggested it was due to the presence of inflectional features that attract the finite verb. Evidence in favour of this intuition has been adduced in the literature by referring to so-called ‘complementiser agreement’ in some Continental West Germanic varieties such as Bavarian or Flemish, where the complementiser carries overt inflectional morphology similar to the verbal agreement suffixes (Bayer 1984; Haegeman 1992; Zwart 1993). I give but one example, from Bavarian German (27). As this example clearly illustrates, the complementiser carries the same inflectional ending ‘sd’ (underlined) as the second personal singular subject:

- (27) *I frog’ me, obsd ned du des mocha kansd*
 I ask me, if not you this do can
 ‘I ask myself if you can’t do this.’
 (From Weiss 2005. Translation added.)

This intuition has also in the main been continued in some form in most generative work on verb-second (Platzack 1986; Vikner 1995), often formalised as a feature [Inf/Fin/Agr/ ϕ] in the C⁰ position.¹⁹

It is worth pointing out, however, that evidence from other languages shows that V-to-C movement is not logically dependent on finiteness. In Italian, today a non-V2 SVO language with considerable word order variation, V-to-C in fact turns up in a non-finite construction known as the ‘Aux-to-Comp’ construction, showing that V-to-C can be disjoint from finiteness in any concrete sense of the word. In (28), the non-finite auxiliary is clause-initial, preceding the subject which in turn precedes the participle, yielding an exceptional case of G-inversion in Italian. This is normally analysed as featuring V-to-C movement of the gerundive (Rizzi 1982; Belletti 2009).

- (28) ***Avendo** Gianni chiuso il dibattito, la riunione è finita prima.*
 Having Gianni closed the debate, the meeting ended early.
 ‘Gianni having closed the debate, the meeting ended early.’ (From Belletti 2009:77)

Perhaps one could argue that ‘non-finite’ is too crude a term and that what the verb in (28) really lacks is agreement morphology, not tense morphology. This could potentially suggest an analysis whereby the requirement or precondition for V-to-C is that the clause be ‘tensed’ in some sense. However, other languages show that V-to-C can occur even in the absence of tense; a prominent example is provided by non-finite adverbial clauses of purpose in Spanish; in (29) the non-finite verb is the infinitive, and yet this verb appears clause-initially while the subject surfaces between the fronted verb and a participle, a hallmark of V-to-C movement.

- (29) *Dáme su número para **poder** yo contactarlo.*
 Give-me his number so-that can.INF I contact-him.
 ‘Give me his number so that I can contact him myself.’

While this does not prove that V-to-C cannot be triggered by some finiteness feature in Germanic, it at least strongly suggests that this correlation is not universal. Furthermore, I also believe that these facts suggest another possible interpretation of examples like (25–26) above, where V-to-C was shown to take place in conditional clauses. Common to both (25–26) and (28–29) is that the verb raises in the absence of a complementiser or subjunction.

¹⁹See (Berman 2003) for an LFG approach without features.

This happens regardless of whether the language normally features V-to-C movement in declaratives or not. Rather, the function of verb movement here seems to be related to *clause typing*, marking the subordinate as a conditional, temporal or purpose adverbial clause, respectively. While these patterns only strengthen the hypothesis of some connection between the complementiser and the verb, at the same time they weaken the idea that verb fronting in (25-26) is necessarily related to the V2 grammar of these languages.²⁰

These observations notwithstanding, it is easy to see why the traditional analysis of V2 should have become a showcase of early P&P theory. Apart from the empirical evidence, V2 is neatly reduced to a matter of parametric setting (locus of inflection/AGR in the grammar) and phrase structure itself, as the restriction to a single XP in front of the verb falls out from the supposedly universal X-bar schema and the status of the clause as a CP. In other words, no need to invoke either PS-rules or word order constraints to account for the position of the verb; the verb-second ‘construction’ is reduced to a complete epiphenomenon.

2.3.1 Deriving the prefield: The EPP-feature and the ban on CP-adjunction

In fact, while the analysis of V2 as presented by den Besten seems to capture the V2 construction elegantly, it does not in fact derive the linear restrictions on the prefield entirely. Some additional machinery was needed to make sure the prefield is not allowed to stay empty, nor host more than a single XP.

Starting with the former point, it is a well-known fact that XP-fronting to the prefield is familiar from many languages as an operation intimately related to and possibly triggered by information structure, for instance to set up a topic-comment or focus-background structure. However, there must clearly be something more at work in the V2 languages, since even all-focus clauses must obey the V2 pattern, either by using the unmarked subject-initial order, or by employing a semantically vacuous expletive in the prefield (30a). It is not possible to start the clause with the verb (30b)–(30c), although these clauses should be perfectly well-formed with respect to both theta-theory and case-theory. (30d) shows that the expletive is only grammatical in the prefield, a fact which illustrates well the role of the expletive in ‘saving’ the V2 order:²¹

²⁰In fact, verb fronting as a clause-typing strategy seems to be a typical feature of the Indo-European languages (Watkins 1964). At an early stage, when clausal relations were predominantly paratactic and complementation/subordination was still underdeveloped, this strategy was used in main clauses to signal imperatives and polar questions. We could perhaps think of the patterns just reviewed as an extension of this clause typing strategy to embedded domains. Admittedly, a more developed version of this argument would also need to say something about the morphological correlates of the various verb-fronting operations, but since this matter is not of central concern here, I will not pursue this any further.

²¹The use of the expletive is constrained by more than just information structure. In the Mainland Scandinavian languages, the verb must be intransitive and the subject indefinite, meaning the construction is used presentationally to introduce new discourse referents, as in (ia). In high register Continental Germanic and Icelandic, it is possible to combine the expletive with transitive verbs.

- (i) a. [*Det*] **kom** en fyr inn på butikken.
 It came a guy in on shop-the.
 ‘There came guy into the shop.’
 b. ***Kom** en fyr inn på butikken.
 (Norwegian)

- (30) a. *[Es] ist ein Paket für dich gekommen.*
 It is a parcel for you come.
 ‘A parcel has arrived for you.’
 b. **Ist ein Paket für dich gekommen.*
 c. **Ist für dich gekommen ein Paket.*
 d. **[Ein Paket] ist es für dich gekommen.*

In other words, the V2 construction cannot be bypassed by particular information-structural configurations and must accordingly be assured by some grammatical principle. A common assumption is that C^0 carries an EPP-feature which causes the merger of an expletive in the prefield if no semantically motivated XP-fronting takes place. This EPP-feature has the effect of ruling out verb-initial declaratives.

A final assumption was necessary to derive the linear restrictions on the prefield. Since adjunction was frequently employed, in particular to represent non-argument constituents in various positions of the clause, a story was needed to account for the ungrammaticality of examples like (11), repeated here for convenience:

- (11) **[Gestern], [ich] habe das Buch im Auto vergessen.*
 Yesterday I have the book in-the car forgotten.
 (Intended: ‘Yesterday I forgot the book in the car.’)

In order to explain why this structure could not be generated by adjoining the initial adverbial to the clause, in parallel fashion to what was assumed at the time to be the case for the corresponding English translation, an assumption was made that UG provides a *universal ban on CP-adjunction* (Iatridou and Kroch 1992, Schwartz and Vikner 1996:12) which rules out structures like (11), where the initial adverb ‘gestern’ is adjoined to the maximal projection of the clause/CP.

However, since all V2 languages do in fact allow certain cases of linear V3, it was assumed that the ban on CP-adjunction could be circumvented in particular contexts. One such context was given above in (13b) above, repeated here as (31a); another and more frequent case of linear V3 is given in (31b). In both of these sentences, two constituents precede the finite verb, resulting in perfectly well-formed linear V3 constructions. The intuition is that such cases involve a left-dislocated constituent which is somehow ‘outside’ the clause, and that this constituent is picked up again by a resumptive element in the prefield. Thus, it is the presence of a resumptive pronominal that is co-indexed with the dislocated DP in (31a) that ‘licenses’ the linear V3 order. Likewise, in (31b) the initial subordinate clause is resumed by a co-indexed temporal adverbial in the prefield.

- (31) a. *[Meine Mutter]_i, [die]_i arbeitet im Krankenhaus. (German)*
 My mother, she works in-the hospital.
 ‘My mother, she works in the hospital.’
 b. *[Wenn du heute Abend zu mir kommst]_i, [dann]_i können wir uns einen*
 If you today evening to me come, then can we us a
Film anschauen.
 film watch.

‘If you come over to me to tonight, then we could watch a movie together.’

The generalization that emerges is thus that Germanic V2 languages allow linear V3 in contexts where the first element is left-dislocated and picked up by a resumptive element in the clause proper.²² Although the intuition is clear enough, it is presumably fair to say that the analysis whereby resumption and co-indexation can circumvent the alleged ban on CP-adjunction has remained somewhat underdeveloped from a formal perspective, and these assumptions have also largely been superseded by more recent theoretical developments, as we shall see in section 2.4.

Summary With this final theoretical assumption in place, the traditional analysis of verb-second can be considered complete. To sum up, it makes the following different theoretical claims:

The theoretical claims of the traditional analysis:

- (32) a. The finite verb moves to the head of the highest projection in the clause, CP.²³
- b. One XP moves to the corresponding specifier projection, Spec-CP. Alternatively, an EPP-feature on C⁰ will cause the merger of an expletive in Spec-CP.
- c. Complementisers and the verb compete for the same head position, C⁰.
- d. The trigger for verb movement is finiteness or inflectional features located in C⁰.²⁴

Notice that the different parts of (32) are not equal in motivation or theoretical necessity. While (32a–(32b) or something equivalent seems required to derive the basic word order facts, (32c–32d) are more tentative in nature and could be described as ancillary

²²It is sometimes claimed that the resumptive must itself appear in the prefield (Alexiadou 2006), but this is probably too strong. While this seems to be the case when the initial element is a subordinate clause, left-dislocated DPs (in the broadest possible sense of the term) can sometimes be linked to a pronoun in situ:

- (i) [*Broren hans*]_i, [*jeg*] **har** ikke hørt fra ham_i på mange år.
Brother-the his, I have not heard from him on many years.
‘His brother, I haven’t heard from him in many years.’

(Norwegian)

This example contains a resumptive co-indexed with the initial DP, but it does not occupy the prefield; for similar examples in German and Dutch, see Frey (2004), who demonstrates that the resumptive may even be embedded in a complement clause, and Grewendorf (2009). In other words, *non-inverted linear V3* is also possible in some cases in the Germanic V2 languages.

²³Müller (2004) develops an analysis of V2 that does not feature Head Movement at all, but rather obligatory (pied-piped) phrasal movement of the entire vP to Spec-CP. On this account, the verb is linearly second because the entire VP minus its ‘edge domain’, which contains exactly one constituent, has been evacuated prior to movement.

²⁴In the heyday of the P&P framework, (32d) was frequently considered the relevant parameter, such that languages where the ‘locus of inflection’ was in C⁰ rather than I⁰ would be V2 languages featuring an extra derivational step, namely Head Movement from I⁰ to C⁰. Another suggestion that has been explored is that V-to-C movement is somehow related to case assignment (Roberts 1993).

assumptions. At the same time, (32c–32d) represent an attempt at going beyond mere phrase-structural description to provide an explanation of this syntactic behaviour of Germanic V2 languages.

2.3.2 The controversy of subject-initial clauses

Before moving on to the next section, it is worth mentioning that a certain controversy has surrounded the status of subject-initial (non-inverted) clauses. Den Besten (1983) assumed that V-to-C movement and XP-fronting takes place in all main clauses without distinction, but if one assumes that the subject and the verb lexicalise the specifier and the head of IP at some point in the derivation, movement of the subject and the verb into the C-layer would be string-vacuous and only reproduce the same Spec-Head relation at a higher level. Some researchers assume V-to-C movement and XP-fronting in all clauses as a matter of conceptual and theoretical uniformity, but Travis (1984) and on several occasions, Zwart, have argued empirically for a regular IP-analysis of subject-initial clauses (Zwart 1991, 1993, 1997). Without getting into details here, the argumentation is based on grammaticality contrasts between subject-initial and inverted clauses with respect to phenomena such as preverbal expletives or clitics, complementiser agreement and coordination facts. This analysis is sometimes referred to as the *Split Hypothesis*, since linear V2 on this account would have two different structures depending on whether the clause is subject-initial (V-to-I) or inverted (V-to-C).

On the other side of the debate, Schwartz and Vikner have produced empirical evidence in favour of the traditional analysis (Schwartz and Vikner 1989, 1996) based on extraction facts (long distance wh-movement), prefield expletives, and asymmetries between subject-initial main and embedded clauses with respect to adjunction. This analysis has been termed the ‘symmetric analysis’, since it posits V-to-C in both subject-initial and inverted clauses. I refer the reader to these papers for details, but we will return later (section 2.6.3.1) to the issue of subject-initial main clauses from a slightly different perspective.

The examples we have been considering so far have been based almost exclusively on German. This was a deliberate choice, for the traditional analysis was motivated by the facts of Continental Germanic. As will become clear in this section and the following, verb-second theory is characterized to a very considerable extent by what could be called *theoretical path-dependency*. Concretely, much could have looked different if the analysis had been driven by the facts of V2 as they appear in other branches of Germanic,²⁵ an empirical domain we will turn to presently.

2.3.3 Outside Continental Germanic: some problems...

It soon became apparent that not all empirical facts relevant to the V2 phenomenon were accounted for by the analysis developed by den Besten and others. The most serious challenges came from outside Continental Germanic and revolve around the issue of the main-embedded asymmetry of V2. As for the Mainland Scandinavian branch, it was already illustrated in (18b) (repeated below) that such asymmetry exists in these languages as well. The position of the finite verb after the sentential negator ‘ikke’ and various IP-adverbs (if present) indicates that the verb does not raise out of the VP at all :

²⁵This point is expressed quite succinctly by Steiner (2014:3): ‘the predominance of Modern German in the V2 literature may have inadvertently skewed our understanding of what V2 is, as it ignores possible variation in the V2 grammar.’

- (18b) *Jeg angret på at [jeg] [ikke] kjøpte boken.*
 I regret on that I not bought book-the.
 ‘I regret that I didn’t buy the book.’

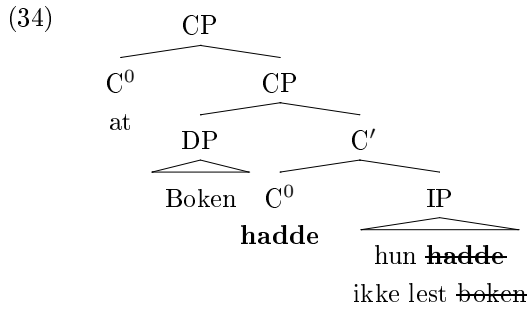
However, this asymmetry does not hold in full generality, since it is perfectly possible to find V2 in certain embedded clauses in Mainland Scandinavian (for an early discussion, see Anderson 1975). In (33a), the finite verb appears in linear second position of the complement clause, preceding the negator ‘ikke’. (4.3) shows that inversion is also possible in this context without dropping the complementiser, while (33c) shows that V2 is not obligatory in this context either.

- (33) a. *Hun sa at [hun] hadde ikke lest boken.*
 She said that she had not read book-the.
 ‘She said that she hadn’t read the book.’
- b. *Hun sa at [boken] hadde hun ikke lest.*
 She said that book-the had she not read.
 ‘She said that the book, she hadn’t read.’
- c. *Hun sa at [hun] [ikke] hadde lest boken.*
 She said that she not had read book-the.

Similar facts hold for all of the Mainland Scandinavian languages, showing that V2 is not only sensitive to the main-embedded distinction, but apparently also to certain properties of the matrix predicate. The verbs and copular predicates permitting embedded V2 have been shown to overlap well with the classes of predicates permitting embedded root phenomena in Hooper and Thompson’s (1973) influential study, and include verbs of strong assertion (generally *verba dicendi*, corresponding to class A in Hooper and Thompson’s typology), verbs of weak assertion (*verba cogitandi* denoting thoughts and mental processes, class B) and some semi-factive verbs like *see*, *find out*, *discover*, *understand*, *realize* and the like (class E). I will refer to these verbs as ‘viaduct verbs’, following the suggestion made by Walkden and Booth (to appear).²⁶

The evidence from Mainland Scandinavian poses a problem to the traditional analysis of V2, since the complementiser in C⁰ should effectively block V-to-C movement. The traditional analysis therefore faces a problem of under-generation. The standard analysis in the literature until Rizzi (1997) was to treat this as an instance of CP-recursion with the complementiser selecting a CP-clause as a complement instead of an IP (deHaan and Weerman 1986; Iatridou and Kroch 1992; Holmberg and Platzck 1995; Vikner 1995). On this account, the structure of an embedded V2 clause such as (4.3) would be as in (34):

²⁶The term ‘bridge verbs’ has sometimes been used collectively to designate this group of verbs (Vikner 1995), but this terminology is misleading, as the latter term is also used for verbs allowing extraction from their complements (Riemsdijk and Williams 1986:294); these verbs and those allowing embedded V2 do not overlap completely.



Such an approach solves the problem by providing an extra CP for the verb to move into, but arguably in a rather stipulative way. Since the existence of embedded V2 is constrained by properties of the matrix predicate, and since embedded V2 is never obligatory in Mainland Scandinavian, there is clearly need for some additional semantic explanation to accompany the syntactic 'CP-recursion' story.

2.3.3.1 Embedded V2 and the 'Assertion Hypothesis'

Since Hooper and Thompson (1973), a connection has been recognised as holding between the possibility of embedded root phenomena and *assertion*. Therefore, a prominent hypothesis holds that embedded V2 is associated with assertion or independent illocutionary force (Wechsler 1991; Holmberg and Platzck 1995; Truckenbrodt 2006; Heycock 2006; Julien 2007, 2009). This hypothesis can be stated as follows:²⁷

(35) *The assertion hypothesis:* (From Wiklund et al. 2009:1915)

The more asserted (the less presupposed) the complement is, the more compatible it is with V2 (and other root phenomena).

Assertion is generally taken to be the illocutionary force employed by a speaker when uttering a proposition and demanding of the addressee that this proposition be taken as part of the *common ground*. In a similar vein, Krifka claims an assertion implies a commitment on the part of the speaker to the truth of the proposition (Krifka 2014). Declarative main clauses by default carry assertions, then. Notice that the hypothesis in (35) views assertion as the exact opposite of presupposition, suggesting the two notions are the poles of a continuum.

While this explanation has a lot of intuitive appeal, it is not without its problems. First, as the minimal pair in (33b)–(33c) illustrates, embedded V2 is never obligatory in Scandinavian. Naturally, the option of not raising the verb does not turn the complement clause into a presupposition, but if (33c) is neither an assertion nor a presupposition, it must be something else. One possible answer is that the distinction between non-V2 and V2 in embedded clauses under verbs of assertion would correspond semantically to a mere 'recounting/reporting' of what was said (non-V2), as opposed to *asserting it as true* (Julien 2015).

However, it has been argued that assertion or independent illocutionary force is strictly speaking disjoint from V-to-C movement entirely. Reis objects to Truckenbrodt's (2006) analysis of embedded V2 as triggered by illocutionary force features in C⁰ by showing that assertions are available in a variety of verb-final dependent adverbial or relative clauses (Reis

²⁷The same basic idea is referred to as the *illocution hypothesis of V2* by Wiklund, who provides the following definition: 'V2 declaratives have illocutionary force, V-*in-situ* declaratives do not' (Wiklund 2010:81).

2006). Reis demonstrates that root phenomena like assertive discourse particles, declarative question tags as well as speech act adverbials are fully felicitous in such clauses even in the absence of V-to-C:

- (36) a. *Tom, den ich ehrlich gesagt/leider nicht ausstehen kann, hat...*
 Tom, whom I honestly said/unfortunately not stand can has...
 ‘Tom, who, to tell you the truth/unfortunately, I can’t stand...’
- b. *Das kann mich gar nicht ärgern, weil ich nämlich hiermit zurücktrete.*
 That can me PRT not anger because I namely hereby resign.
 ‘This can’t make me angry anymore because I hereby resign.’
 (From Reis 2006:373-374)

Parallel evidence and arguments have been adduced for Swedish by Wiklund, who claims that assertive verbs (classes A and B in the Hopper & Thompson typology) always take complement clauses that express assertions in Swedish, regardless of whether there is V-to-C in the complement clause (Wiklund 2010). These clauses also allow typical main clause phenomena like discourse particles, speech act adverbials and swear words, regardless of whether the verb stays in a VP-internal position or raises to C⁰. Wiklund therefore concludes that a strong version of the assertion/illocutionary hypothesis, postulating a biunique relationship between V2 and assertion, cannot be maintained, although a weaker version, according to which V2 order (in declaratives) always correlates with assertion, may be sustained.

However, even this weaker hypothesis faces problems, since embedded V2 is in fact possible in certain complement clauses which do not express assertions at all. These are clauses embedded under class E predicates, semi-factives which in fact presuppose the truth of their complements, seemingly in outright contradiction of the assertion hypothesis. Wiklund therefore points out that one must consider an ‘inclusive definition of ASSERTION in the sense that it has to cover also semi-factive verbs...’ (Wiklund 2010:87).

The question is if such a definition is workable, given that it is normally considered nonsensical or even impossible to assert presuppositions (Kiparsky and Kiparsky 1970). In Wiklund et al. (2009), a different analysis is developed, whereby the complements of class D predicates permit root phenomena since these predicates have the capacity to update the common ground between the speech partners; that is, their complement might constitute the *main point of the utterance* (MPU) in the sense developed by Simons (2007).

I will not go further into the debate on the semantics of embedded V2. To summarize briefly, it seem like the the relevance of notions like assertion, illocutionary force or main point of utterance is to *define the contexts where embedded V2 is in fact possible*, rather than saying anything about the semantics of V2 itself. It is still an open question if V-to-C movement per se is capable of ‘triggering’ anything semantically. Far from permitting root clause phenomena, embedded V2 is just itself one such root phenomenon, which can be employed or not in certain contexts – without noticeable semantic effect.

The lack of semantic effect has also been used to argue that V2 (and other head movement phenomena) may not be the result of syntactic processes at all. I will therefore quickly review some of these arguments before moving on.

2.3.3.2 V2 as PF-movement?

In recent years, the role and locus of Head Movement in the grammar has received novel attention. In Chomsky (1995:368), the question is raised in passing whether Head Movement

is really an operation in Narrow Syntax at all, and in Chomsky (2001:37-38) this argument is developed significantly.²⁸ There are several reasons that led Chomsky to hypothesize that Head Movement might in fact be a phonetic process altogether, but we will focus on only two of them; for a more detailed discussion, see Roberts (2011).

First, Chomsky points out that Head Movement has little if any effect on semantic interpretation. This is expected if Head Movement takes place in the PF-section of the derivation, which is without an interface to LF.²⁹ Secondly, Head Movement violates the Extension Condition, which stipulates that all instances of Merge, whether external or internal (Move), should extend the topmost node of the phrase marker.

Chomsky's broad assault on Head Movement in general is particularly relevant to V2, not only as a presumed instance of Head Movement, but as a phenomenon that has always been considered by some researchers as intimately related to prosodic processes such as cliticization. The potential relation between Germanic V2 and second position clitics was already discussed in the seminal paper by Wackernagel (1892), as has always persisted as a minority position in the literature on verb-second (Anderson 1993, 2000).³⁰

However, evidence has been adduced to demonstrate that V-to-C movement is not always without semantic effect. V-to-C movement is the only thing that syntactically distinguishes the minimal pair in (37); V-to-C movement (37b) interacts with the scope of the initial PP, an unexpected result for a movement which takes place at PF/in P-syntax.³¹

- (37) a. [In no clothes], [they] **would** look good.
 b. [In no clothes] **would** they look good.

(From Walkden 2016:7)

Roberts (2011) discusses various other alternatives to Head Movement that have been proposed, such as Remnant VP movement or 'reprojective movement'. I refer the reader to Robert's paper for details and references, but to summarize his general point one might say that, while these alternative approaches avoid some of the criticisms directed at Head Movement by Chomsky, they run into other problems such as the lack of an obvious trigger for Remnant Movement³² or a general complication of the theory of movement. Furthermore,

²⁸Chomsky makes an exception for *incorporation* in the sense of Baker (1988), since the creation of morphologically complex heads cannot be a purely PF-phenomenon, and since incorporation was assumed to be involved in grammatical function-changing phenomena (Chomsky 2001:37).

²⁹A partially similar explanation for the lack of semantic effect with V2 has been offered by Bayer (2008), who assumes that only the finiteness features of the verb must be realized in C⁰ in Narrow Syntax, and that the lexical stem of the verb is pied-piped along at PF. At LF, the verb is reconstructed in its base position.

³⁰Cf. also Dewey (2006) for the claim that Germanic V2 developed as a prosodic requirement. Among the evidence considered by Dewey is the frequent tendency for the finite verb in early Germanic verse to split the initial constituent, creating hyperbata of the Wackernagel type (Dewey 2006:31-33).

³¹The term P-syntax refers to movement that still takes place in syntax (after Spell Out), but which is somehow triggered by prosodic or other requirements imposed at the sensorimotor interface. Roberts points out that there is presently no theory about word order at PF that commands general consensus, nor do we even know if movement at PF as such even exists.

³²Remnant Movement is movement of an XP which contains a trace, in other words movement of a category that has been evacuated by some of its former members by some prior movement operation before movement of the XP itself takes place. In the case of V2, the alternative to Head Movement would most likely be Remnant VP Movement. But while such an operation may have some plausibility in languages like German, which generally allows middle field scrambling, it is harder to motivate in languages which lack independent evidence for the displacement operations that would need to evacuate the VP prior to movement. This also raises the question how this movement operation could be acquired.

given that none of these alternatives seem to be a global alternative to Head Movement, capable of explaining the range of phenomena traditionally attributed to this operation, it seems like there is still a role to play for Head Movement in current syntactic theory.

2.3.3.3 No competition between verbs and complementisers in Mainland Scandinavian

Returning to the traditional analysis of V2, I would like to point out at this point that alongside the well-known problem of under-generation, the traditional analysis faces a related problem of over-generation that to my knowledge has not attracted the same kind of attention in the literature. In the Mainland Scandinavian languages, it is perfectly possible not only to have V2 under overt complementisers, as we have just seen, but also to drop the complementiser without raising the verb, exactly what is ruled out in German (see 23). The presence or absence of a complementiser does not affect word order; rather, V2 is always optional under certain verbs.³³

- (38) a. *Han sa at han ikke hadde gjort det.*
 he said that he not had done it.
 ‘He said that he hadn’t done it.’

- b. *Han sa _ han ikke hadde gjort det.*

The latter example is completely normal and unmarked in spite of the lack of a complementiser, and this seems to suggest there is no competition between V2 and complementisers at all in the Scandinavian languages, and that the failure of the verb to raise in embedded contexts is due to something else entirely. could of course suggest that examples like (38b) involve some kind of PF deletion of the complementiser, but in the absence of evidence, this is not an attractive solution.³⁴

³³It has even been claimed (Reinholtz 1989, Vikner 1995:84-84) that embedded V2 is impossible in Mainland Scandinavian when the complementiser is dropped, suggesting the situation here is the exact mirror image of what is found in Continental Germanic. This is at least not the case for all varieties; for instance, dropping the complementiser is perfectly fine in my (northern) variety of Norwegian. Interestingly, a slight degradation is felt when inversion is used:

- (i) a. *Han sa _ han hadde ikke gjort det.*
 He said he had not done it.
 ‘He said he hadn’t done it.’

- b. ? *Han sa _ det hadde han ikke gjort.*

This contrast might be interpreted as indicating that (ia) only features V-to-I movement. If this is the case, *Subject-verb-negation*-strings are in fact an insufficient diagnostic for embedded V2, although they have frequently been used this way (Wiklund et al. 2009). One could make an argument in favour of the *Split Hypothesis* from these facts, (see section 3.1.1.1). Still, the contrast with Continental Germanic is stark, since overt complementisers in the Mainland Scandinavian languages, to the extent they play a role, rather facilitate V2.

³⁴Observe also that it does not work to suggest that V-to-C in (38b) takes place in covert syntax due to weak feature strength and Procrastinate, since this would leave unexplained why the verb does in fact overtly raise to C sometimes in embedded clauses and always in main clauses, unless one would like to postulate two different C-heads, one with a strong inflectional feature and one another with a weak feature (and both optionally null) and then postulate that the latter is only available in embedded clauses (although similar analyses have in fact been suggested, cf. the account of wh-phrases in Brazilian Portuguese in Hornstein et al. 2010:42-44). I can see no motivation for going to such lengths only to save the hypothesis of a competition

2.3.4 More problems: ‘symmetric’ V2 languages

In a sense, the Mainland Scandinavian languages demonstrate the same kind of main-embedded asymmetries as German or Dutch. The difference is that in the former group, these asymmetries apparently follow from the choice of matrix verb, rather than from the presence or absence of complementisers, while in the latter group, both of these factors play a crucial role.

However, it has been argued that there exist V2 languages which lack this asymmetry between main and embedded clauses altogether. This claim was made for Icelandic by Rögnvaldsson and Thráinsson (1990)³⁵ and for Yiddish by Santorini (1989) and Diesing (1990).³⁶ These authors argued that verb-second operates in main and embedded clauses alike. The following examples feature inversion under what seems to be non-assertive (39a) and factive verbs (39b-39c), classes C and respectively D in Hooper and Thompson’s typology (1973), something which is generally not possible in either Continental Germanic or Mainland Scandinavian languages.³⁷

- (39) a. *Jon efast um [à morgun] fari María snemma à fætur.*
John doubts about that she has not met this man.
‘He doubts that she has not met this man.’
(From Rögnvaldsson and Thráinsson 1990:23)
- b. *Jón harmar að [Þessa bók] skuli ég hafa lesið.*
Jon regrets that this book shall I have read.
‘John is sorry that I’ll read this book.’
(From Rögnvaldsson and Thráinsson 1990:23)
- c. *Es iz a shod vos [hayntike tsaytn] kenen azoy fil mentshn afile nit*
It is a shame that today’s times can PRT many people even not
leyenen.
read.
‘It is a shame that nowadays so many people can’t even read.’
(From Diesing 1990:44)

Based on such examples, it might seem like Icelandic and Yiddish do not feature any asymmetries at all between main and embedded clauses with respect to V2. At least two possible analyses were proposed to account for this state of affairs. One suggestion was that such cases involve CP-recursion (Platzack 1986; Vikner 1995). This is the same analysis that was proposed for embedded V2 in Mainland Scandinavian, but if Icelandic and Yiddish allow V2 in all kinds of embedded clauses, this analysis is forced to assume that CP-recursion is freely available in these languages. Such ‘generalized CP-recursion’ was perceived to be

between the verb and the complementiser, a hypothesis which, as far as I can tell, finds little if any empirical support in Mainland Scandinavian.

³⁵But see also Thráinsson 1986.

³⁶See also Biberauer 2002, who claims a similar ‘symmetric’ trend is developing in Afrikaans.

³⁷On closer scrutiny, (39b) cannot possibly be semantically equivalent to Mainland Scandinavian ‘angre’ or German ‘bereuen’, since these verbs normally mean something like ‘wish that one had not made the choice X’, something which it is nonsensical to say of the actions of others. In Wiklund et al. (2009, p.1922), it is argued that ‘harma’ is closer in meaning to English ‘regret’ in the sense ‘feel sorry’, and that this verb does not presuppose the truth of its complement in the strict sense, since it can represent new information to the hearer. In this respect, the authors argue, it resembles semi-factive verbs (class E), which are known to permit embedded root phenomena, in some respects, while yet differing in other respects (the factivity of its complement is content-sensitive).

theoretically unattractive, and an alternative analysis was developed, according to which verb-second operates at a lower clausal level in the ‘symmetric V2’ languages, namely the IP (see Rögnvaldsson and Thráinsson 1990 for Icelandic and Diesing 1990 for Yiddish). This would explain why embedded V2 is generally possible, since the complementiser and the verb never compete for the same node. On this account, the surface similarity of the Germanic V2 languages masks an important structural difference. This analysis has also been proposed for historical stages of Germanic and Romance, notably by Pintzuk for Old English (Pintzuk 1991, 1995), by Fontana for Old Spanish (Fontana 1993) and by Lemieux and Dupuis for Old French (Lemieux and Dupuis 1995).

2.3.4.1 The traditional typology of Germanic V2

In order to take stock of the evidence we have been reviewing so far, as well as of the research situation around the mid 90’s, it is instructive at this point to consider the typology of verb-second languages as it was perceived at that time. On analogy with the *traditional analysis*, I will call this the *traditional typology of V2*.

Based on observations of the kind we have been examining, Vikner (1995) developed the following three-way typology of V2:³⁸

(40) **The traditional typology of verb-second:** (based on Vikner 1995)^{39 40}

- *Well-behaved V2 languages:* V2 takes place in complementiserless clauses. German, Dutch, Afrikaans.⁴¹

³⁸In Vikner’s study, it was assumed that Faroese patterned with the Mainland Scandinavian languages because it lacks independent V-to-I in embedded clauses (i.e. the order *subject-negation-verb* is quite frequent), the assumption being that independent V-to-I is a prerequisite for generalized/symmetric V2 of the Icelandic kind. Subsequent research has nuanced this picture somewhat. Jonas (1996) claimed there is dialectal variation and made a distinction between ‘Faroese A’, which patterns like Icelandic in allowing generalized embedded topicalisation, while ‘Faroese B’ rather patterns with the Mainland Scandinavian languages, therefore requiring an asymmetric analysis. Heycock et al. (2010) confirm that Faroese patterns rather like Icelandic with respect to embedded V2, although they did not find any evidence for a dialect split. Interestingly, the authors also found that Faroese rather pattern like Mainland Scandinavian with respect to embedded V-to-I, thereby raising some doubt on the often postulated link between embedded V-to-I and embedded V-to-C (Vikner 1995; Koeneman 2000). Angantýsson (2011) even finds a *positive* correlation between accepting the order *Adv-V* in embedded clauses and embedded topicalisation.

³⁹We might also have included the group *Residual V2 languages*, where V2 is restricted to certain specific constructions, as is the case in modern English, French and Spanish.

⁴⁰The term *well-behaved V2 language* was not used by Vikner, but has gained some currency in the literature. In general, the terminology employed to describe this variation has been less than constant. Gärtner (2016) uses the terms *broad* and *narrow embedded V2* to describe languages like Icelandic/Yiddish and the Mainland Scandinavian languages, respectively. Another, more-theory laden terminology is used by Holmberg (2015), who distinguished I-V2 languages (Icelandic/Yiddish) from C-V2 languages (the rest). This terminology presupposes that the observed empirical differences are due to a difference in the locus of verb-movement, an assumption which is more and more challenged in the research literature.

⁴¹One might raise the question if Dutch and German should be grouped together. Although both are SOV V2 languages, Zwart claims that the complementiser is never left out in embedded clauses in Dutch, while V2 is still possible in colloquial language:

- (i) *Tasman zei dat [hij] had er geen zin in.* (From Zwart 2011:107)
 Tasman said that he had LOC NEG appetite in
 “Tasman said that he didn’t feel like it.”

According to Holmberg, embedded V2 is rare in (standard) Dutch (Holmberg 2015:358).

- *Limited embedded V2 languages*: V2 occurs with overt complementiser, but only in a definable subset of embedded clauses. The Mainland Scandinavian languages, Faroese, Frisian.
- *General embedded V2 languages*: V2 occurs in main and embedded clauses without distinction. Icelandic, Yiddish.

While this typology still retains some descriptive utility today and is in essence reproduced in a recent overview paper on V2 by Holmberg (2015), the picture has been relativized somewhat since Vikner’s study, with the overall message being that verb-second languages presumably are more similar to each other than previously assumed. Furthermore, although this has not been explicitly stated, it does not seem unreasonable to say that a certain *convergence* has taken place, in the sense that both the ‘well-behaved’ and the ‘general embedded/symmetric’ V2 languages have turned out to be more like the ‘middle group’ represented by the Mainland Scandinavian languages. Let us briefly examine why this is the case.

2.3.4.2 Do ‘symmetric V2’ languages exist?

In recent years, ‘symmetric V2’ languages have received novel attention, often with the outcome that the validity of the very notion has been called into question. Subsequent research on Icelandic has shown that there is both dialectal and generational variation and that judgements even vary considerably between speakers (Angantýsson 2008). Jónsson (1996) suggests there is a dialectal split in Iceland; ‘Icelandic A’ would be as described in Rögnvaldsson and Thraínsson (1990), while ‘Icelandic B’ would be more like the Mainland Scandinavian languages and need an asymmetric analysis. Others have gone one step further and rejected the symmetric analysis for Icelandic (Bentzen 2007), or even questioned the very existence of symmetric V2 languages outright (Hrafnbjargarson and Wiklund 2009).⁴² In Wiklund et al (2009), sentences featuring inversion under non-assertive and factive verbs (classes C and D) were generally not accepted by the Icelandic or Faroese informants.⁴³ It

⁴²In his comparative study of six different Medieval Romance varieties, Wolfe came to the same conclusion, rejecting previous hypotheses about the potentially symmetric nature of V2 in Old French or Spanish: ‘... the widely-assumed class of truly ‘symmetrical’ V2 languages may not exist at all.’ (Wolfe 2015c:149)

⁴³There is also, to my mind, some outright confusion over the issue of embedded V2. For instance, Holmberg cites the following minimal pair from Wiklund et al. 2007 as evidence that V2 is ‘optional in relatives and adverbial clauses’ (Holmberg 2015:357):

- (i) *Ég veit um ena bók som Jón (hefur) ekki (hefur) lesið.*
 I know of one book that Jon (has) not (has) read.
 “I know about one book that Jon has not read.”

This is not evidence for optionality of verb-second (understood as V-to-C), however, only optionality of V-to-I. Indisputable evidence for V-to-C, in relatives and adverbials as in other clauses, would have to feature inversion:

- (ii) ?? *Ég veit um ena bók som hefur Jón ekki lesið.*

Since the word order *Subject-Verb-Negation* is apparently accepted under all predicates in Icelandic, contrary to what is the case with embedded topicalisation, this suggests that Icelandic and (some variants) of Faroese feature independent V-to-I movement. Minimal pairs like the one illustrated in (i) may plausibly be interpreted as meaning that V-to-I is optional in some contexts. It should be mentioned that Wiklund et al. (2009) argue on the basis of word order facts in non-finite complement clauses that *Subject-Verb-Negation*

is also worth pointing out that Hrafnbjargarson and Wiklund (2009) found that adjuncts are more easily fronted to the prefield in embedded clauses in Icelandic, a finding that suggests that for some languages, *inverted V2* may have to be split into *argument inversion* and *adjunct inversion*. To the best of my knowledge, such a distinction is not relevant in the other Germanic languages, although similar effects have been reported for the Rhaeto-Romance V2 variety of St. Leonardo by Benincà and Poletto (2004:60-61).⁴⁴

As for Yiddish, it was already pointed out by den Besten and Moed-van Walraven (1986) that topicalisation in embedded interrogatives and relative clauses was not optimal, a judgement which is echoed in Diesing (1990), although the author points out that context can mitigate this. Walkden and Booth (to appear) conducted a corpus search into historical Yiddish and found little evidence for IP-V2 syntax, nor for the claim that embedded V2 is generally available in Yiddish.

Turning back to the so-called ‘well-behaved’ Continental Germanic languages, it was already observed above that at least German seems to show the same distribution as in the Mainland Scandinavian languages with regard to embedded V2, with the important difference that the former group normally also requires the complementiser to be dropped. However, even in these languages it is possible to come across embedded V2 in the presence of complementisers (Zwart 1997; Biberauer 2002; Freywald 2008), a fact that indicates that, while the difference between the group of languages is real enough, it is not quite as categorical as perhaps once assumed.

To summarize, the evidence accumulated over the last couple of decades cast some doubts on the accuracy of the traditional three-way typology of V2 languages. To put the matter more succinctly: the empirical differences between the three groups are at the very least smaller than previously assumed, and the differences which still remain – near-total complementarity between complementisers and V2 in Continental Germanic, more frequent embedded XP-V_{fin} orders in Icelandic (and possibly Yiddish) – may well be unrelated to the V2 syntax of the languages, but rather follow from other, independent differences, the same conclusion reached by Walkden and Booth (to appear).

2.3.5 Embedded V2 in non-complement clauses

We will round off this section with an examination of embedded verb-second in non-complement clauses, in particular various adverbial clauses. Since it is necessary to keep this phenomenon distinct from the issue of generalized embedded/symmetric V2, which has been discussed and problematized already, the focus here is on embedded V2 in non-complement clauses in Mainland Scandinavian and Continental Germanic.

V2 and more generally embedded root phenomena are attested in a wide variety of adverbial clauses. Haegeman (2007, 2010) makes a distinction between *central* and *peripheral* adverbial clauses. Root phenomena are generally permitted in the latter group, which

strings also feature V-to-C movement. The authors do not provide an explicit analysis of the contrast in grammaticality judgements between what they call *subject-initial* and *non-subject-initial* (i.e. *inverted*) V2 for class C and D predicates, but conclude on that ‘...[n]one of the Scandinavian languages can therefore be said to display generalized embedded V2...’ (Wiklund et al. 2009:1922). My general point here is that, no matter the clause type, only inversion is strong evidence for V-to-C, or alternatively – as pointed out to me by George Walkden (p.c.) – postverbal material that is necessarily adjoined to IP (or higher).

⁴⁴In a derivational framework, it seems natural to search for an explanation of such asymmetries between arguments and adjuncts in the potential distinction between moved/internally merged and base-generated/externally merged elements, perhaps along the lines of Haegeman’s suggestion that operator movement may count as interveners for argument fronting to the LP (Haegeman 2012).

semantically express independent propositions and syntactically allow argument fronting and expressions of epistemic modality. First, V2 is attested in various Germanic languages in consecutive adverbial clauses, as illustrated in (41–43). The embedded clause in (41) features topicalisation of the direct object and inversion, while (42–43) features verb-movement across sentential negation.⁴⁵

- (41) *Jeg var så stresset at [boken] glemte jeg i bilen.*
 I was so stressed that book-the forgot I in car-the.
 ‘I was so stressed that I forgot the book in the car.’
 (Norwegian)

- (42) *Hy is sa siik dat [hy] kin dy hjoed net helpe (kin).*
 He is so sick that he can you today not help can.
 ‘He is so sick that he cannot help you today.’
 (Frisian, from Holmberg 2015:359)

- (43) *Han er så sjuk så/att [han] kan inte (kan) hjälpa dej.*
 He is so sick so/that he can not (can) help you.
 ‘He is so sick that he can’t help you.’
 (Swedish, from Holmberg 2015:359)

Since adverbial clauses are not arguments, one cannot approach this issue via the C-selectional properties of their matrix predicate. Consecutive clauses are particular, however, generally being quasi-arguments of a reinforcing adverb (‘I was so stressed, that...’). Unlike other adverbial clauses, they do not provide a reference point (temporally, locally, causally, etc.) for the eventuality contained in the preceding main clause, but rather express the consequence of the latter. The logical relationship between adverbial clause and main clause is therefore exactly the opposite of what is the case in other adverbial clauses. This is also reflected in the fact that consecutive clauses always follow the main clauses of which they express the consequence, seemingly some principle of iconicity in syntax.

Another context where embedded V2 is well attested is in adverbial clauses of reason, or rather adverbial clauses introduced by subordinators of the ‘because’-kind. In fact, ‘reason’ is somewhat misleading here, since a crucial observation is precisely that embedded V2 does not seem to express the reason for the eventuality contained in the accompanying main clause.⁴⁶ Antomo and Steinbach suggest that German ‘weil’-V2 clauses feature V-to-C movement, which enables independent illocutionary force, and back up this claim by showing how they exhibit various pragmatic and semantic properties that set them apart from their

⁴⁵It is unfortunate that the practice of using these *subject-verb-negation* strings as evidence for embedded V2 has become so established, since the assumption it builds on, that there are only two possible positions for the finite verb (V⁰ or C⁰) in the relevant languages, is not beyond doubt.

⁴⁶We might say that the *semantic type* is still that of *reason*, only the conjunction operates at a higher level of discourse. This distinction between *external* conjunction (holding between states of affairs) and *internal* conjunction, holding at the level of discourse, is common; see Verstraete (1999) for references and for a discussion if internal conjunction should be further split into *epistemic* and *speech act* conjunction. The Germanic languages do in fact possess such specialised ‘internal’ conjunctions which operate at the level of discourse, such as ‘for’ in Norwegian or ‘denn’ in German. This means that ‘because’-clauses across Germanic are in fact trespassing into the domain of these paratactic conjunctions, as has been pointed out (Pasch 1997).

non-V2 (i.e. clause-final) counterparts and bring them into line with V2 complement clauses (Antomo and Steinbach 2010).⁴⁷ At the level of syntax, their most salient characteristic is a reduced distributional flexibility: ‘weil’-V2 clauses must always follow the main clause to which they belong, as illustrated by the contrast in (45):

- (44) *He koe net kommer omdat [hy] moast Teake helpe (moast).*
 He could not come because he must Teake help (must).
 ‘He couldn’t come because he had to help Teake.’
 (Frisian, from Holmberg 2015:359)
- (45) a. *Peter kommt zu spät, weil [er] hat keinen Parkplatz gefunden (hat).*
 Peter comes to late, because he has no parking-place found (has).
 ‘Peter comes to late, because he hasn’t found a place to park.’
 b. **Weil er hat keinen Parkplatz gefunden, kommt Peter zu spät.*
 (German, from Antomo and Steinbach 2010:4)

Other adverbial clauses which exhibit V2 order in several Germanic languages include adversative clauses introduced by ‘while’-subjunctions (46) as well as adversative concessive clauses⁴⁸ introduced by ‘although’-subjunctions (47):

- (46) a. *Foreldrene jobber, mens [i stua] sover barna deres.*
 parents-the work while in living-room sleep children-the their.
 ‘The parents are working, while in the living-room their children are sleeping.’
 (From Bentzen 2009:18)
- b. *Tagsüber sind Berlins Straßen immer verstopft, während [nachts] gibt es eigentlich nie Stau.*
 Through-the-day are Berlin.GEN streets always jammed while at-night
 is there really never jam.
 ‘In daytime the streets of Berlin are always jammed, while in the night there is really never any jam.’
 (From Freywald 2016)
- (47) a. *Hun bestod eksamen, skjønt [noen toppkarakterer] fikk hun ikke akkurat.*
 She passed exam-the although any top.marks got she not really.
 ‘She passed the exam, although she didn’t exactly get top marks.’ (From Bentzen 2009:18)
- b. *Ich will keine Kekse mehr, obwohl [ich] nehme noch einen.*
 I want no crackers more; although I take yet one.
 ‘I don’t want any more crackers, although I’ll take another.’
 (German, from Frey and Masiero 2018:69)

⁴⁷Reis (2013) takes issue with their analysis. Although ‘weil’-V2 clauses clearly feature to V-to-C movement and syntactic parataxis, Reis claims that neither of these syntactic properties are responsible for the illocutionary force and the particular semantic properties, since the latter are available in the absence of the former. Cf. the ‘assertion debate’ in section 2.3.3.1.

⁴⁸In German, verb-second is used in clauses introduced by ‘wobei’ \approx ‘although, whereas’ (Günther 2000).

We may conclude that V2 order is possible in a definable subset of adverbial clauses exhibiting considerable similarities across Germanic. Without entering into the details of how these V2 adverbial clauses differ semantically from their non-V2 counterparts, the rough generalization seems to be that V2 is optionally possible with conjunctions which permit a ‘high’ interpretation, in other words where the logico-semantic relationship between the clauses expressed by the conjunction does not pertain between *states of affairs* in the world, but at the level of discourse organization, what Haegeman refers to as *peripheral adverbial clauses* (Haegeman 2007, 2010). This means that a syntactic parataxis is established which allows V2 order. Crucially, it seems to be the case that all instances of V2 in adverbial clauses require that the adverbial clause follow its head clause. V2 seems to be impossible in purely temporal adverbial clauses, perhaps because a high attachment is not available in such cases.

Finally, it is worth pointing out that some varieties of German also exhibit V2 order in relative clauses. I am not aware of the existence of V2 in this domain outside of German, although other embedded root phenomena such as speech act adverbials have been reported (Wiklund 2010:87).⁴⁹ The examples in (48a)-(48b) illustrate V2 order in a subject-initial and inverted relative clause, respectively. Example (48c) shows that the same type of distributional restriction that was noted above for adverbial clause V2 is valid here also, since the V2 relative clause can not modify a DP which precedes the matrix clause verb (i.e. which appears in the prefield).⁵⁰

- (48) a. *Das Blatt hat eine Seite, [die] ist ganz schwarz.*
 The sheet has a side that is completely black.
 ‘The sheet has one side which is completely black.’
- b. *Ich suche jemanden, [den] nennen sie Wolf-Jürgen.*
 I search someone who call they Wolf-Jürgen.
 ‘I’m looking for someone who they call Wolf-Jürgen.’
- c. * *Jemanden, [den] nennen sie Wolf-Jürgen, suche ich.*

(From Gärtner 2001:98-99)

2.4 Cartographic approaches to the Left Periphery

With the advent of cartography (Rizzi 1997 *et seq.*), many different phenomena involving the clausal left-periphery had to be reconsidered. On the basis of word order facts from mainly Italian and other Romance languages, Rizzi proposed to split the erstwhile unitary CP into a layer of different A’ projections, as depicted in (49):

- (49) [ForceP... [TopP*... [FocP... [TopP*... [FinP... [IP]]]]]]

On Rizzi’s account, the C-system is to be considered ‘the interface between a propositional content (expressed by the IP) and the subordinate structure (a higher clause or,

⁴⁹On the other hand, the increased awareness of word order instilled in me over the last three years have allowed me to observe a truly remarkable frequency of the order *subject-verb-negation* in relative clauses in spoken Norwegian. Not a single case of inversion has come to my attention, though. Again, this suggests that V-to-I may be an option in some embedded clauses that reject V-to-C.

⁵⁰It is worth noting that Gärtner argues that these clauses are not really relative clauses from a syntactic perspective, although that is how they are interpreted (Gärtner 2001).

possibly, the articulation of discourse. . .’ (Rizzi 1997:283). The link between the Left Periphery and the core clause (IP) is expressed by the lowest node, Fin^0 , while the connection to the articulation of discourse is expressed by the higher node, Force^0 , assumed to be responsible for encoding the speech act or the illocutionary force. In between these ‘formal’ heads, there is a focus position surrounded on either side by a layer of recursively nested topic positions.

While Rizzi’s original structure for the LP has been the object of revision and elaboration, resulting in several updated ‘roadmaps’ of the LP, it is fair to say that cartography has shown an unusual capacity to scale up, since cross-linguistic research has uncovered quite robust word order patterns along the lines of (49). As we shall have occasion to discuss in section 2.6.4, these patterns raise the question of the exact status of cartographic hierarchies in the grammar. Anticipating that discussion somewhat, it is worth noting that Rizzi emphasizes the role of functional projections in providing transparency at the C-I interface, suggesting that ‘the syntactic computation hands over to the interpretive component representations transparently indicating dedicated positions for certain discourse functions. . .’ (Rizzi 2004:7; cf. also Belletti 2004b:4). This is very much in line with the central Minimalist desideratum of seeing derivations as driven by interface requirements (Chomsky 1995).

2.4.1 Cartography and V2

The advent of cartography has had profound consequences for the understanding of the V2 phenomenon. And yet, I believe it would be misleading to say that the theory itself has been substantially altered by cartography. Rather, the traditional analysis has been continued and reinterpreted in the light of the new phrase structural reality uncovered by cartographic research. In consequence, some problems faced by the traditional analysis can now receive a more satisfactory solution.

As a case in point, consider the phenomenon of embedded verb-second. In the traditional analysis, this was strictly speaking not supposed to exist at all, and the numerous counterexamples were solved by stipulating CP-recursion, allowing an extra projection for the verb and the topicalised element to move into. In a system like (49), it is possible to say that certain verbs (classes A, B and E in Hooper and Thompson’s system, so-called ‘viaduct verbs’ (Walkden and Booth to appear)) select a ForceP , a clausal complement carrying independent illocutionary force and thereby permitting embedded root phenomena, while other verbs (classes C and D) select a smaller complement, a FinP that does not carry independent illocutionary force.⁵¹ The same approach can extend to V2 in non-complement clauses (see section 2.3.5), with the important difference that the availability of an embedded left periphery cannot be derived from the c-selectional properties of the matrix verb, but must

⁵¹This hypothesis receives strong empirical support from the phenomenon of *recomplementation* or *double complementisers*. In this construction, an embedded left peripheral phrase is sandwiched in between two overt complementisers. While recomplementation is particularly frequent in the history of the Romance languages (Poletto 2000), it is also attested in Germanic, as witnessed by (i):

- (i) *Peter glaubt, dass [den Studenten], dass den keiner gelobt hat.*
 Peter thinks that the.ACC student that him.ACC nobody.NOM praised has.
 “Peter thinks that, as for the student, nobody has given him any praise.”
 (From Grewendorf 2009:68)

This receives a natural explanation if we assume that the higher complementiser lexicalises Force^0 and the lower Fin^0 , with the left dislocated phrase occupying some intermediate left peripheral position such as TopicP . See also Roberts (2004) for a similar case in Welsh.

be made to follow somehow from parataxis.

Cartography therefore permits a maximally simple theory of embeddability: embedded root phenomena are the result of an embedded left periphery, again the result of a high complementiser in Force⁰.⁵² This fits nicely with a truncation account to word order facts such as has already been a common assumption for various non-finite clauses (Haegeman 2012); a FinP is then a truncated clause, crucially lacking the left peripheral topic and focus positions necessary for V2.

While cartography has attracted much attention from all camps, it seems to have found particular favour with Romanists. Apart from possible sociological factors (the cartographic research program having been developed by linguists in the Romance tradition (Rizzi 1997; Cinque 1999)), there is a good reason for this, namely the fact that the Romance languages generally exploit the left periphery to a far greater extent than for instance the Germanic languages. Compare a case of multiple fronting as in (50) with the corresponding German clause (51). According to Rizzi, (50) involves a focus, *questo*, preceded and followed by a topic. The equivalent being starkly ungrammatical in a Germanic V2 language, one would have to choose which of these three constituents to front to the prefield, for instance the focus:

- (50) [*A Gianni*], [*QUESTO*], [*domani*], *gli* ***dovrete*** *dire*.
To Gianni, THIS, tomorrow, him.CL should.2.PL say.

Literally: ‘To Gianni, this, tomorrow, you should tell him.’

(From Rizzi 1997:291)

- (51) [*DIESES*] ***solltet*** *ihr morgen Gianni erzählen*.
THIS should.2.PL you tomorrow Gianni tell.

As for the older stages of the Romance languages, cartography has provided a possible solution to an observation made in the wake of den Besten’s traditional analysis of V2. It was pointed out by Benincà (1983) that Old French and Northern Italian varieties featured high amounts of linear V2 and a widespread use of a construction strongly reminiscent of the modern Germanic V2 languages, namely the inversion structure where the subject intervenes between a finite auxiliary and a non-finite main verb (G-inversion), a construction that has all but died out in declarative clauses in the modern Romance languages. At the same time, the Old Romance languages display numerous exceptions to linear V2 of a kind that is not possible in the modern Germanic V2 languages. It was therefore suggested that these Old Romance varieties were ‘relaxed V2’ systems; in a sense, they feature some ‘Germanic’

⁵²It is important to emphasize that this is not a miracle solution to all problems of embedded V2, even in the so-called asymmetric systems. In Mainland Scandinavian languages, embedded root phenomena are available under the viaduct verbs, but V2 order is still just one option, as it is also fine to leave the verb in the VP. A natural analysis would be to say that viaduct verbs do not have to select ForceP; they can also select a FinP, and this syntactic difference would then correspond semantically to the difference between asserting something (requiring independent illocutionary force encoded in Force⁰) and just reporting something. However, this analysis encounters problems both on the syntactic and the semantic side. Syntactically, it has been shown that other standard root phenomena, such as discourse particles or interjections, may co-occur with a non-raising verb. These elements are not even located in the left periphery, but clause-internally. If we assume that they are still licensed by the selection of ForceP, this immediately raises the question why the verb does not raise to the vacant Fin⁰ position. In other words, this is additional evidence that, at least in the Mainland Scandinavian languages, there is no inherent feature on Fin⁰ that attracts the verb. On the semantic side, it has been argued that complement clauses without verb raising represent assertions just as much as their V2 counterparts, cf. the debate on assertion vs. main point of utterance (MPU) in section 2.3.3.1.

properties (widespread use of G-inversion), while already making substantial use of the left periphery like their modern descendants.

In a cartographic approach to the left periphery, these two properties can be reconciled analytically. The assumption is that the finite verb raises to a head in the left periphery, and that one or more XPs are allowed to appear in front of the verb, for instance a sequence *Topic-Focus-Verb*. We will return to the technicalities of the derivations that have been proposed, but first we will focus on some conceptual consequences of such an analysis.

2.4.2 Verb second = V-to-C?

It should be immediately clear that a theory of verb-second that allows substantial amounts of V3 and V4 entails some kind of definitional drift, since there is no linear restriction on the prefield equivalent to what is found in the Germanic V2 languages; in fact, V-to-C movement alone is retained as the sole criterion for defining V2 (Poletto 2002; Ledgeway 2008; Wolfe 2015b). I believe there is good reason to question the appropriateness of this conception of V2.

For starters, this definition seems *unintuitive* in light of the fact that the phenomenon, in the original sense of the word, owes its name to the perceived linear restriction, not to inversion. Secondly, and much more importantly, an equation between V-to-C movement and verb-second would potentially extend the label ‘V2’ to an unknown, but presumably very high number of languages in the world. For instance, all languages where VSO is derived by V-to-C movement will be V2 languages on this definition. VSO is the third-most common word order among the languages of the world (Langus and Nespors 2015:142), and it does not seem unreasonable to venture that a considerable share of them might necessitate a V-to-C analysis.⁵³ In this scenario, V2 languages would pass from being a typologically very rare phenomenon, attested in only a handful of geographically scattered languages outside the Germanic family (Holmberg 2015), to a cross-linguistically widely attested phenomenon.⁵⁴ Once again, this is unintuitive and leaves unexplained the salient and typologically rare pattern found in the Germanic languages, which in my view merit a distinct label beyond that of being ‘strict V2 languages.’

A proponent of this extended V2 definition⁵⁵ might object at this point that there does not per se exist any linear restriction on the position of the finite verb in any language, neither in Germanic nor in ‘relaxed V2’ systems; linear order is a superficial trait of variation, a surface phenomenon related to deeper grammatical properties such as the locus of verb movement and phrase structure. At first, this seems to be true; we have seen that linear $V \geq 3$ orders are indeed possible in Germanic V2 languages as well, and one of the major achievements of the traditional analysis was precisely to reduce the linear restriction to a surface effect of phrase structure, a pure epiphenomenon of the X-bar schema and the status of the clause as a CP - plus an assumed universal ban on CP adjunction. Obviously,

⁵³Thanks to George Walkden (p.c.) for pointing out this to me.

⁵⁴This typological extension might be limited significantly by saying that V2 languages feature V-to-C movement and an EPP-feature on the relevant left-peripheral head targeted by the verb, but no linear restriction in any form, such that V2 languages feature linear $V \geq 2$. This is not the approach adopted in for instance Wolfe (2015), where the verb-initial grammar of Old Sardinian is considered a verb-second grammar with V-to-Fin movement, but without an EPP-feature on Fin⁰, nor in the essentially similar analysis developed for Late Latin by Ledgeway (Ledgeway 2017).

⁵⁵I am addressing something of a straw man here; my point is that this is in essence the core of the argumentation, as I read it, provided by those researchers who are concretely in favour of broadening the definition of ‘V2’ to comprise all Romance varieties, old and new, which feature V-to-C movement in declarative clauses (Poletto 2002; Benincà 2004; Benincà and Poletto 2004; Ledgeway 2008; Wolfe 2015b).

this analysis must be rethought in light of the new phrase structural reality uncovered by cartographic research.

Given a layered CP consisting of many different functional projections, the following questions are crucial: A) how should the Germanic V2 systems be analysed? B) How should the so-called ‘relaxed V2’ systems be analysed? C) If the so-called ‘relaxed V2’ systems do indeed feature V-to-C movement, what are the factors responsible for the difference between the former and the latter group?

We will now explore some possible answers that could be given, and come concrete answers that *have* been given to these questions. I will argue that none of them straightforwardly reduce linear order to a superficial trait of variation. This leads to the conclusion that at least some of the Old Romance languages should not be lumped together with the Germanic languages under the common denominator ‘V2’, and that surface linear order restrictions, although not part of the theory itself, should still somehow be guaranteed in the definition of a V2 language.

2.4.3 ‘Relaxed’ and ‘strict’ V2 languages

As already mentioned above, the structure of the left periphery has been the object of revision after Rizzi’s (1997) paper, and this revision has generally gone in the direction of more elaborate structures. Benincà and Poletto (2004) take issue with Rizzi’s hypothesis of a topic position on either side of the focus position (see (49), and claim there is no topic position below FocusP. They also suggest that recursion of a projection is generally not possible, and that both FocP and TopP are really just shorthand notations for fields of different topic and focus positions with slightly different information structural values. The authors also come to the conclusion that there is a ‘Frame’ field above the ForceP, able to host scene-setting adverbial expressions and Hanging Topics. This gives a layered left Periphery along the lines in (52), where it is important to notice that every projection may potentially be a shorthand for even finer sub-structures:

$$(52) \text{ [FrameP [ForceP [TopicP [FocusP [FinP [IP]]]]]]]$$

Given such a finely structured left periphery, it is clear that the notion of V-to-C movement is no longer precise. We are therefore led to ask exactly what projection is targeted by the finite verb in Germanic V2 languages.

The first intuition would be to assume that it raises to the highest projection, in order to account for the fact that only one constituent is normally allowed to precede it. However, this hypothesis would suffer the same problems of undergeneration as the traditional analysis, predicting embedded V2 to be non-existing. Furthermore, it would falsely predict that linear V2 is completely exceptionless.⁵⁶ Even worse, this hypothesis makes the wrong predictions regarding the information structural properties of the initial XP, which can in fact be either topic, focus, scene-setter, etc. And finally, it would also predict that the lower part of the left periphery should be able to appear between the verb and the subject.⁵⁷ In other words, this theory would fare significantly worse than the traditional story.

⁵⁶The reason why this prediction arises is that cartography does not easily allow adjunction, with some cartographers even claiming that there is no adjunction at all in the grammar (Benincà and Poletto 2004). On such an account, the universal ban on CP-adjunction has become a universal ban on adjunction *tout court*, and co-indexed resumptives (see section 2.3.1) cannot circumvent this. The relationship between phrase structure and cartography will be discussed in section 2.6.4.

⁵⁷This has in fact argued to be correct by Frascarelli and Hinterhölzl, based on examples like (i):

Suppose we assume instead that the verb can move to any head in the left periphery, as long as the relevant head is activated by an XP in its specifier, in a kind of criterial approach to V2 (Samo 2018). While this approach correctly predicts that the initial constituent does not have a dedicated function, it does seem to predict that it must have *some* kind of semantically meaningful function. But again, this is wrong. Not only can the prefield in all Germanic V2 languages host topics, foci and scene-setters, it can also host semantically vacuous, phonologically unstressed expletive elements, as shown in (53):

- (53) [*Es*] **hat** heute eine Frau angerufen.
 it has today a woman called.
 ‘A woman called today.’

This is perhaps not a serious problem, since one can assume that, in the absence of active criteria in the LP, an EPP-feature triggers some kind of ‘Last Resort’ merger of expletives in Spec-FinP. More problematically, in any case where the relevant criterial projection hosting the verb and the XP is a low one, one must wonder why it is not possible to have more than one XP in front of the verb. This problem cannot be brushed off lightly by appealing to locality restrictions, since locality only constrains movement; on the reasonable and widely held assumption that at least scene-setters can be base-generated in the LP (Poletto 2002; Wolfe 2015b), the prediction is that these should be able to precede the V2 construction. We already saw in (11) that this prediction is not borne out:

- (11) *[*Gestern*], [*ich*] **habe** das Buch im Auto vergessen.
 Yesterday I have the book in-the car forgotten.

The problem of the ‘linear restriction’ is shared by any full-fledged cartographic approach that postulates verb movement to a low left peripheral head, whether this happens in specific criterial constructions, or quite generally. The restriction to a single preverbal constituent, a fact which fell out altogether naturally from the monolithic CP of the traditional analysis, is now all of a sudden left unexplained and must be accommodated by adding something to the theory. As we have already seen, it does not work to just move the verb higher in the left periphery.

What follows from this discussion, in sum, is that a cartographic approach has a hard time accommodating the following facts about Germanic V2 systems:

- (54) a. Embedded verb second exists in all Germanic languages and must be accounted for (suggesting the verb cannot move too high).
 b. All the different IS values associated with the LP are available to the initial XP (suggesting the verb cannot move too high).

-
- (i) [*Gestern*] **hat** der Hans die Maria getroffen.
 Yesterday has the.NOM Hans the.ACC Maria met.
 ‘Yesterday Hans met Maria.’

(German, from Frascarelli and Hinterhölzl 2007)

According to F&H, the initial adverbial moves to a the specifier of the high projection FrameP and the verb moves to the corresponding head position, while the subject ‘der Hans’ lexicalises a lower left-peripheral (shifting) topic position, ShiftP.

- c. The left-peripheral sequence does not appear to the right of the verb, i.e. between the verb and the core clause (suggesting the verb cannot move too high).
- d. The prefield is (generally) restricted to one single constituent (suggesting the verb cannot move too low).

The crucial point to note is that from a cartographic perspective, there is an inherent tension between (54a–54c) and (54d). In the next section, I will review so-called ‘bottleneck’ approaches that have been designed to overcome this problem. I will show that these are still not able to derive the linear restriction, while at the same time suffering from serious conceptual drawbacks.

2.4.4 Bottleneck approaches to V2

We have seen that the nature of the evidence from the Modern Germanic languages is such that it is hard to embed the theoretical claims of the standard analysis in the cartographic model of the left periphery. The crucial difficulty arises from property (54d) above, the highly restricted prefield. As mentioned, this property is absent from some other languages for which a V2 analysis has also been proposed in the literature. In these so-called ‘relaxed V2’ languages, the prefield may host several constituents, giving rise not only to linear V3, but also linear V4 and even $V > 4$. A cartographic analysis might therefore prove more workable for these varieties, due to the absence of property (54d). What follows therefore applies first and foremost to ‘strict’ V2 languages. For a more detailed discussion of some of the empirical problems related to the ‘relaxed’ V2 languages within the ‘bottleneck’ approach, see Hsu (2017).

In order to derive (54d), an analysis has been developed which is considerably more involved than the traditional story. Relying on the theory of locality known as Relativized Minimality, (Rizzi 2001) Roberts (2004) suggests that the relevant head is the lower CP head Fin^0 . In V2 languages, this head must be overtly filled by movement or merge. In embedded clauses, the complementiser is merged here, and in main clauses, the verb raises. Next, an EPP feature on Fin^0 requires an XP to fill SpecFinP .⁵⁸ Roberts goes on to argue as follows:

‘XP movement to Spec-Fin in full V2 clauses is movement caused only by Fin ’s EPP feature. . . . The moved XP is thus of no particular type in terms of the typology of potential interveners, and so is able to block any type of movement.’
(Roberts 2004:316)

In other words, the element in SpecFinP blocks all further XP-movement to the left-periphery by Relativized Minimality. This approach has come to be known as the ‘bottleneck’ approach (see also Haegeman (1996) for the original idea and Mohr (2009) for an application of this theory to German).⁵⁹

⁵⁸Poletto suggests the EPP must be ‘. . . conceived as a general requirement on having a predicative structure as the highest relation in the clause.’ (Poletto 2002:216) A similar notion, the ‘subject of predication’ is entertained by Mohr (2009). However, this generalization is problematic and breaks down both at the IP and CP levels in the case of preverbal expletives, which cannot establish a predicative structure with the rest of the clause. The same strictly speaking applies to most adverbials.

⁵⁹It has also been suggested (Roberts 2012; Wolfe 2015b) that the bottleneck condition might be linked to the status of Fin^0 as a phase-head in the derivation. Phases are the minimalist heirs to Barriers from the GB-framework and are invoked to account for (among other phenomena) locality effects by assuming a derivational theory of multiple spell-outs; a phase head is a point in the derivation where the complement of

This is not the end of the story, but let us stop for a moment and observe that this analysis does not at all follow from the basic assumption underlying Relativized Minimality. In fact, the rationale behind RM is that a more specific XP (presumably to be understood as richer in features) blocks crossing movement of a less specific XP, and evidence to support this view has been adduced in the literature (see for instance Starke 2001 and Abels 2012). It is hard to see how Relativized Minimality could be invoked to account for the alleged blocking effect of the EPP-feature, if the theory is to maintain any predictive power.⁶⁰

Even if one accepts that the EPP in Fin^0 might work as a bottleneck and stop further movement to the left periphery, this clearly cannot be the end of the story. If the EPP just needs the specifier of FinP to be filled, why does it sometimes attract topics, sometimes foci, sometimes scene-setters, regardless of their position in the core clause? A central idea behind cartography is that the functional heads should be read off as transparently as possible at Logical Form. Accordingly, we need topics to go to TopP and foci to go to FocP , and so on.

The idea is that a functional head in the left periphery with an uninterpretable feature attracts an XP with a corresponding (unvalued) feature to its specifier.⁶¹ On its way to the relevant specifier, the XP moves through SpecFinP , thereby checking the EPP feature and closing the bottleneck of the left periphery for further movement (presumably by leaving a copy), before reaching its final destination.⁶²

The central idea is that the bottleneck in SpecFinP only blocks movement; it is possible to base-generate additional XPs in the LP after the V2 constraint has been satisfied, potentially giving rise to V3 and V4 orders. As we know by now, this is not possible in strict V2 languages of the Germanic kind. Therefore, an additional assumption has been developed by Wolfe (2015). In fairness, Wolfe's theory is not developed to account for the Germanic V2 languages, but rather for the relatively strict linear V2 requirements of certain Old Romance varieties such as Old French, Old Spanish and Old Venetian. I will consider

the relevant head is sent off to the A-P and C-I interfaces for articulation/interpretation. Syntactic objects with unchecked features must therefore move to an 'escape hatch' in the (outer) specifier of the phase head before the complement is shipped off to the interphases (see Chomsky 2000 for the conceptual foundation and Gallego 2010, 2012 for elaboration). It is worth noting, however, that the bare phrase structural framework with multiple specifiers adopted in phase theory does not match with the Kaynean X-bar template adopted in cartography. The reason why this is relevant is that it is not entirely unproblematic for cartography to just avail itself of an extended X-bar schema with multiple specifiers, since this would interfere with the Principle of Transitivity: if A precedes B, how can we know if A is in a higher projection or in an outer specifier of B? On the other hand, if cartography maintains the traditional X-bar schema, it might not have the relevant escape hatches needed in phase-driven derivation at all, unless one assumes that outer specifiers are created derivationally on the spot at the phase edges.

⁶⁰In other words, this is Relativized Minimality 'upside-down', as Abels put it during the conference 'Word Order in the Left Periphery' (Abels 2018) - a sentiment which Rizzi agreed with. For empirical evidence against the notion that less specific can block more specific, again see Starke (2001).

⁶¹Note here that it is not at all obvious how features like 'Topic' or 'Focus' enter the derivation in the first place, as they are not lexical features; an XP is not a topic or a focus per se, but rather becomes so due to the configurations it enters at some later stage (necessarily in syntax in T-model based generativism). A related question is how the derivation of such sentences can respect the Inclusiveness Condition (Chomsky 1995:228-229), which states that the linguistic object at LF must only contain features of the lexical items already present in the Numeration. See Aboh (2010) for the hypothesis that IS-features are indeed present in the Numeration.

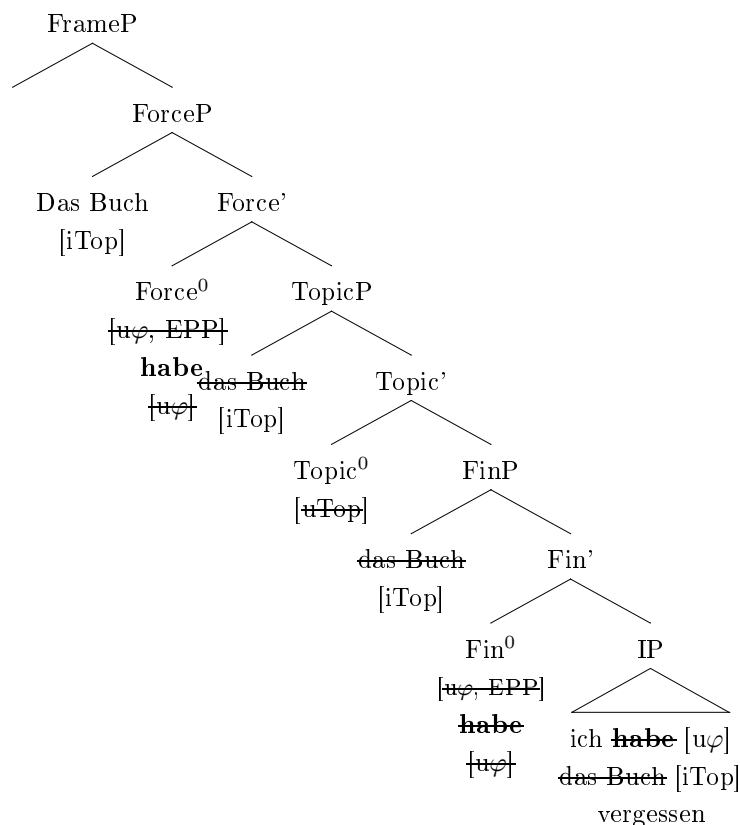
⁶²If no appropriate feature is available that can move to the LP, it is commonly assumed that the EPP can be satisfied by merger of an expletive in SpecFinP . Yet another possibility suggested in the literature on Continental Germanic is 'formal movement', whereby the EPP feature of Fin^0 attracts the closest XP in the IP/middle field. This movement is also considered to be semantically vacuous as it does not on its own give rise to extra pragmatic effects; it may however interact with a prior scrambling operation which has moved another XP above the subject. (Fanselow 2002; Frey 2004b)

Wolfe’s hypothesis for Old French in detail in chapters 3 and 4. However, since Wolfe hypothesises that this might be the system underlying strict V2 languages in general, it is worth exploring its potential to account for Germanic as well.

2.4.4.1 The ‘double bottleneck’ approach

Wolfe’s idea is that strict V2 languages are endowed with a second set of EPP and finiteness features, located in a higher left-peripheral head, namely Force^0 . The derivation proceeds in identical fashion to what is assumed for the ‘relaxed V2’ systems, but crucially involves an extra step. For an example like (55) below, involving an initial topic, this gives the following derivation: after the IP has been constructed, the finite verb moves to Fin^0 to check its φ -features. The topic projection in the left periphery ‘probes’ its complement and finds a DP with an unchecked topic feature, *das Buch*. This DP is then attracted by internal merge (move) to the specifier of the (phase) head Fin^0 , setting up the Spec-Head relation necessary to check the EPP-feature on Fin^0 . The DP moves on to SpecTopP to check its topic feature, but crucially leaving a copy in SpecFinP which acts as a bottleneck, closing the left periphery for further movement operations. The verb is then moved from Fin^0 to Force^0 to check the ‘second set’ of φ -features, and the DP, having checked its topic feature in SpecTopP , is attracted to SpecForceP by the higher EPP. Now that the verb and the single element have reached their final destination in Force^0 , the verb-second constraint is derived. Omitting the FocP , which is irrelevant to this example, this gives the following derivation for an example like (55):

- (55) *[Das Buch] habe ich vergessen.*
The book have I forgotten.
‘I have forgotten the book.’



It is fair to say that this analysis is somewhat less than minimal. It employs a questionable interpretation of Relativized Minimality to derive the lower bottleneck in FinP. Since this is not enough to derive the restricted nature of the prefield, it must employ a final operation which is no more than a pure repetition of the same procedure.⁶³ Also, one must ask how a verb can move twice to check finiteness features, but no explanation is provided. Furthermore, the constituent in the prefield ends up in the specifier of ForceP, which does not tell anything about its semantics. The dual semantics of the moved phrase is therefore not associated with the base position and the position it occupies at PF, meaning that the interpretation at LF, barring any final readjustments in covert syntax, must be read off the base position and the highest copy in SpecTopP. In the words, there is no transparency at the interface of the kind cartography strives to acquire. This is not an analysis driven by interface requirements.

However, a more fundamental problem with this analysis is that it still does not work empirically. Since the whole derivation up until ForceP is rendered opaque by the dual bottlenecks and EPP-features, the predictive power left is mainly related to the area above ForceP, where first-merger is predicted to be possible. Wolfe, building on Benincà and Poletto (2004) assumes that the Frame field includes at least a projection for hanging topics and scene-setters. In order to test this hypothesis, one needs an understanding of what counts as a scene-setter. Benincà and Poletto do not provide a definition beyond saying that the FrameP encodes ‘the ‘where and when’ of the sentence’ (Benincà and Poletto

⁶³Note that the movement from TopP to ForceP also violates Criterial Freezing (Rizzi 2007).

2004:71), but Wolfe provides a more detailed characterisation:

‘The pragmatic characteristics of this group of elements is homogeneous. They have adverbial characteristics, scope over the entire clause and anchor the speech-act either temporally, spatially or aspectually’ (Wolfe 2015b:14).

If one concretely considers what kind of elements are proposed by these authors, they involve various adverbial expressions of time and place, in addition to connectives and some others; examples include ‘yesterday’,⁶⁴ ‘in 1999’ (Benincà and Poletto 2004:66-67), ‘now’, ‘then’, ‘thus’, and adverbial *when*-clauses (Wolfe 2015b). Without taking any stance on whether they qualify as scene-setters on the definition provided by Wolfe, it seems to be the case, as far as I have been able to gather, that not a single one of these is acceptable with V3 in any Germanic standard V2 variety.⁶⁵ It matters little that some of them might be rendered grammatical by dislocating them and adding a resumptive in the prefield, for they are predicted to be grammatical without such resumption strategies, as they are indeed in the ‘relaxed V2’ systems and according to Wolfe even in the relatively stricter V2 varieties Old French, Old Spanish and Old Venetian. In modern Germanic V2 languages they all trigger inversion and linear V2, and they are furthermore all freely embeddable, contrary to what one would expect if they belong above ForceP and the highest possible complementiser is merged in Force⁰. In other words, in Germanic, these kind of adverbial expressions either do not qualify as scene-setters (and it does not seem reasonable to assume that the same elements can be consistently analysed as scene-setters in some languages and as something else in other languages), or the FrameP is situated below ForceP.

On the other hand, it is possible to have certain left-dislocated DPs to the left of the V2 construction in Germanic. These also require a resumptive in the following clause, but this is presumably an independent principle of left-dislocated DPs (LDs), since they must be linked to the clause somehow. It is commonly assumed that there are various types of LDs with different pragmatic properties, and this has led cartographers to the assumption that they occupy different structural positions (Benincà and Poletto 2004). In Germanic, a distinction is minimally recognized between hanging topics (HTs, also called *nominativus pendens*) and so-called contrastive left dislocations (CLDs).⁶⁶ The former has the pragmatics of an aboutness topic, is prosodically detached, and does not correspond in case with its resumptive correlate (unless accidentally, if both are nominative) in the following clause. The following examples illustrate that HTs are allowed to precede the V2 construction in both German and Icelandic, giving rise to linear V3.

- (56) [*Der Hans*], [*ich*] *kenne ihn schon seit zwölf Jahren*.
the.NOM Hans I know him.ACC already since twelve years.
‘As for Hans, I’ve known him for twelve years.’

(German, from Riemsdijk 1997:5.)

⁶⁴Holmberg also refers to the adverbial expression ‘today’ as a scene-setter (Holmberg 2015:348).

⁶⁵This analysis might work better for V3 in urban vernaculars of the kind reported in Walkden (2017), where exactly these kind of initial elements, whether they qualify as scene-setters or not, are frequently followed by SVO-orders. Walkden also assumes that these lexicalise a high FrameP.

⁶⁶The term Contrastive Left Dislocation goes back to Thrafnsson (1979) and is really something of a misnomer, since the phrases so designated do not necessarily carry any contrastive reading at all. This is pointed out by Frey, who suggests the term ‘German left dislocation’ (Frey 2004a). To this it might be objected that the construction exists in the other Germanic languages as well, so ‘Germanic’ would probably be a more appropriate epithet. I will retain the term CLD.

- (57) [*Þessi hringur*], [*Ólafur*] **hefur** lofað *Mariú honum*.
 this ring-NOM Olaf has promised Maria it.DAT.
 ‘This ring, Olaf has promised it to Maria.’
 (Icelandic, from Thráinsson 2007:358.)

In other words, the hypothesis that hanging topics occupy a very high position in the left periphery receives support, and the preceding examples are compatible with the ‘double bottleneck’ approach or any other analysis which postulates verb movement to Force⁰. It has also been suggested that this construction is unembeddable (Grewendorf 2009:69). Others have claimed the opposite (Villa-García 2012) but if we for the sake of the argument assume that it is correct, this would constitute more evidence in favour of a high position for HTs, above ForceP.

However, the other and far more frequent LD construction in Germanic, contrastive left dislocation, also precedes the V2 construction. Its most salient difference from HTs is that the dislocated phrase agrees in case with the resumptive correlate.⁶⁷ Much cartographic work on the left periphery coincides in assuming a lower position for CLDs than for HTs, somewhere inside the topic field situated below ForceP (Benincà and Poletto 2004; Frascarelli and Hinterhölzl 2007). But if this is the case, the verb cannot be in Force⁰ in examples like (58):⁶⁸

- (58) [*Diesen Frosch*], [*den*] **hat** die *Prinzessin* gestern geküsst.
 this.ACC frog it.ACC has the.NOM princess yesterday kissed.
 ‘This frog, the princess kissed (it) yesterday.’
 (From Boeckx and Grohmann 2005:1)

Furthermore, this construction is embeddable, both in German and the other Germanic languages. Although this may at first sight seem like counterevidence to the analysis developed by Wolfe, since the verb clearly cannot be in Force⁰ – a position which presumably is lexicalised by the complementiser – this is in fact not the case. The point is that Wolfe assumes two different loci for verb-movement, splitting the old V-to-C movement into a twofold process of V-to-Fin and Fin-to-Force, both triggered by EPP-features and φ -features. The latter derivational step is excluded from embedded contexts completely in the presence of an overt complementiser, unless we assume a third, even higher complementiser than the one in Force⁰, while V-to-Fin movement is predicted to be possible under viaduct verbs.⁶⁹

⁶⁷Since case marking of the left dislocated phrase is the primary criterion for distinguishing between HTs and CLDs, these constructions are hard to tell apart in other Germanic languages than German and Icelandic. Note, however, that the resumptive is a clause-internal regular pronoun in the HT example (56), whereas it appears in the prefield as a D-pronoun in (57). This might be considered additional morphological and distributional evidence; if the latter is worth anything as a possible criterion, it is clear that HTs are available in the Scandinavian languages as well, cf. section 2.3.1. However, Boeckx and Grohmann (2005) argue that this is an unreliable criterion and even suggest that it might not be possible to distinguish between the two constructions in the absence of case. Grewendorf (2009) claims of CLDs that ‘there is no pause between this element and the following clause’, a claim which seems somewhat too strong to me.

⁶⁸The same problem is pointed out in Salvesen (2013).

⁶⁹Bayer (2001) also gives an example from German:

- (i) *Ich glaube [den Hans], [den] kennt er kaum.*
 I think the Hans him knows he barely
 ‘I thank that he barely knows Hans.’
 (From Bayer 2001:24.)

- (59) *Jón segir að [Þessum hring], [honum] hafi Ólafur lofað Maríu.*
 John says that this ring(D) it(D) has Olaf promised Mary(D)
 ‘John says that Olaf promised this ring to Mary.’
 (Icelandic, from Thráinsson 2007:359.)

But this raises another problem for the double bottleneck analysis, namely the fact that it predicts a more liberal V2 syntax in embedded clauses. After all, almost the entire left periphery apart from the area above ForceP is available, since the verb lexicalises the very lowest head. At the very least, one would expect the order *Topic-Focus-Verb* to be available. One could try appealing to locality effects, if foci can block topics by RM, but this leaves unexplained why these orders are in fact quite liberally attested in the ‘relaxed V2’ systems, as Wolfe (2015) himself demonstrates.

One way or the other, there is an unresolved issue here. It would be very unattractive to stipulate differences in the way locality works in different languages. And even if one did choose to say that foci block topics in Germanic but not in Old Romance, we still would not have an explanation for why it is not possible to base-generate scene-setters of the kind discussed by Beninca and Poletto and Wolfe in the left periphery of embedded clauses, since we have already seen that these belong below ForceP in Germanic. As non-selected adverbial elements with a sentence-wide scope, it should be possible to base-generate them in a high position, avoiding any locality interventions imposed by the phrase in SpecFinP. But the order *Scene-setter-Topic/Focus-Verb* is not grammatical in embedded clauses, either.

As it stands, the double bottleneck approach does not make the right predictions for Germanic V2 languages. It escapes (most of) the problems of a simple V-to-Fin analysis (2009) in main clauses by making some theoretically costly assumptions, only to face them again in embedded clauses, this time without a remedy.

2.4.5 Feature scattering and the bundled-CP approach

There exist yet another approach to V2 which is worth considering, namely the approach adopted by Hsu (2017), which is based on the theory of feature scattering proposed by Giorgi and Pianesi (1996). The central idea of this theory is that languages might share a common inventory of morphosyntactic features, but that they differ as to whether these features head their own projection in syntax or are bundled on one or more heads. While this model in fact predates cartography, it might be considered a kind of ‘cartography light’ version. While abandoning the cartographic tenet of ‘one feature, one head’,⁷⁰ it is still compatible with the cartographic idea that features are strictly ordered through the adoption of a Universal Ordering Constraint:

- (60) *UNIVERSAL ORDERING CONSTRAINT:*

The features are ordered so that given $F_1 > F_2$, the checking of F_1 does not follow the checking of F_2 .

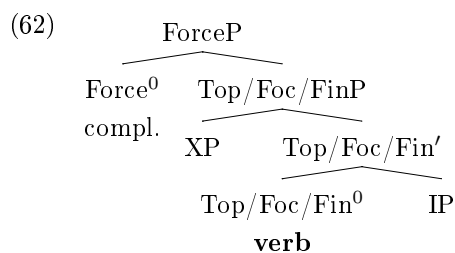
- (61) (Giorgi and Pianesi 1996, from Hsu 2017:18).

⁷⁰Note that Cinque and Rizzi stress the ‘heuristic’ value of this principle and explicitly admit that complex heads might arise in syntax. However, they also hypothesize that complex heads can only arise through head movement, so that complex heads ‘cannot be ‘atoms’ of the syntactic computations’ (Cinque and Rizzi 2009:14). Bundled heads are at odds with this hypothesis.

When employed in the left periphery, this approach constitutes a compromise between the traditional approach based on a unitary CP and the full cartographic model. While a head may carry several features such as [Frame] [Topic], [Focus], only the features associated with an EPP will trigger movement or merge of a phrasal category in the corresponding specifier.⁷¹

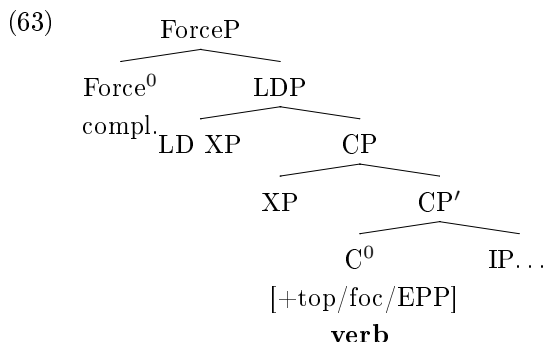
While Hsu (2017) is a nice demonstration of how this approach is descriptively successful in capturing the diverse facts of several so-called ‘relaxed V2 languages’, I will show that this model faces some of the familiar problems when extended to the strict V2 languages of the Germanic type. Let us start with the assumption, hinted at by Hsu himself, that these languages simply bundle all the left peripheral features on one single head. This amounts, of course, to nothing less than the traditional den Besten analysis recast as a ‘bundled CP’. Accordingly, it makes the same predictions: (1) it predicts that embedded verb-second under overt complementisers should not be possible and (2) it predicts linear V2 to be exceptionless in main clauses. While the first assumption might in fact hold for some speakers of Continental Germanic, it is generally incorrect for Germanic as a whole, and the second prediction is not correct at all.

This means that we need to divide this feature bundle and spread the features over more than one head. But where is the correct cut-off point? All positions that are available in main clauses are available in embedded clauses as well (with the possible exception of Hanging Topics), when embedded under an appropriate lexical verb, and this even includes linear V3 orders with left-dislocated elements. This means that it does not work to just split out Force⁰, since this only predicts embedded V2, not embedded V3:



It is therefore necessary to make room for another projection above the V2 construction, but still below ForceP, since it is embeddable. This projection can host various left-dislocated elements (LDs), at the very least CLDs (and these can in turn have quite diverse IS properties) and subordinate clauses, provided there is a resumptive somewhere inside the core clause. Since the information-structural properties of left-dislocated phrases are quite diverse, this projection must itself host a feature bundle related to all possible left-dislocated elements, and I will therefore just call it LDP. As for the prefield in V2 constructions, it is able to host both topics and foci, scene-setters, expletives, etc. Rather than giving it very cumbersome name like ‘FrameP/TopP/FocP/FinP’, I will just call it by its traditional name ‘CP’. Omitting irrelevant positions, this gives the following representation:

⁷¹In fact, this model does not have to make a distinction between moved and first-merged elements like the ‘bottleneck’ approach.



This structure seems to give the right description of the facts in Germanic V2 languages. To the extent that there is any predictive power in this model, this derives indirectly from cartography, or from the ordering of the bundled features, as it is precisely this ordering that permits certain ‘cut-off points’ and certain left-peripheral co-occurrences while banning others. The theoretical claims are still the same as in the traditional analysis: the complementiser itself is a head in the clausal projection, and verbs and complementisers compete for this same node, which is identified as Fin^0 , not Force^0 , in this model. The unavailability of verb-second in embedded clauses is still treated as a purely syntactic fact. This means that the model will face some of the familiar problems when faced with the evidence from the Scandinavian languages, since the hypothesis of a competition between complementisers and the verb is not really well supported by the evidence (see section 2.3.3.3). Still, this model retains descriptive adequacy, not a small feat when compared to the difficulties faced by full-fledged cartographic models.

2.5 Linear non-V2 orders in Germanic V2 languages

All of the Germanic V2 languages permit certain deviations from the linear V2 pattern. Some of these are common to all of the languages, others are particular to one language or group of languages. We will consider some of these in this section, bearing in mind that the intention is not to review all linear non-V2 patterns that exist in Germanic, but rather to dispel any illusion that linear V2 is almost exceptionless, or even that the deviations are limited to a few isolated cases. Furthermore, while many of the exceptions may still be accommodated within the general theory of V2, in the sense that they online deviate from linear V2 without violating structural V2, others even provide direct counterevidence against the structural V2 mechanism itself.

2.5.1 ‘V3 adverbs’

In the Scandinavian languages, certain adverbs may give rise to linear $V \geq 3$ in main clauses. The first class involves a set of adverbs that regularly appear as the second constituent of the clause, ‘sandwiched’ between the initial XP and the finite verb. Following Thrainsson (2007), I will refer to these adverbs as ‘V3 adverbs’. Their exact distribution is quite complex: while all of the V3 adverbs can appear in second position whenever the initial XP is the subject (64a), apparently only some of them can do so when the initial XP is a non-subject (64b); on the other hand, none of them can appear as the first constituent

of the clause themselves (64c).⁷² Note also that the adverbs may be combined, forming strings like in (64d), where the verb is relegated to linear 5th position. There is also some micro-variation across the Scandinavian languages in this domain and perhaps also between dialects; note that while Icelandic displays the same phenomenon, the class of adverbs is not identical to that found in Norwegian, as the cognates of ‘kannski/náttúrulega/sennilega’ are not allowed to intervene between the subject and the verb in Norwegian. These are exactly the same adverbs that are in fact permitted in clause-initial position in Icelandic (as are they in Norwegian).

- (64) a. [*Han*] [*bare/ nærmest/ omtrent/ rett* *og slett/ fullstendig/ totalt*]
 He just almost circa straightly and plainly completely totally
ignorerste *beskjeden*.
 ignored message-the
 ‘He just/almost/straightforwardly/completely/totally ignored the message.’
- b. [*Beskjeden*] [*bare / nærmest / ??omtrent / ??rett og slett / ??fullstendig / ??totalt*]
ignorerste *han*.
- c. [**Bare *nærmest *omtrent *rett og slett *fullstendig* *totalt*] *ignorerste* *han*
beskjeden.
- d. [*Han*] [*bare*] [*nærmest*] [*fullstendig*] *ignorerste* *beskjeden*.
 (Norwegian)
- (65) a. [*Jón*] [*bare/ einfaldlega/ kannski/ náttúrulega/ sennilega*] *ljúkur Þessu*
 John just simply maybe naturally probably finished this
einhhvern daginn.
 some day.
 ‘John will just/simply/maybe/naturally/probably finish this one day.’
- b. [**Bare/ *einfaldleg/ kannski/ náttúrulega/ sennilega*] *hefur Jón lokið*
 just simply maybe naturally probably has John finished
Þessu.
 this.
 (Icelandic, from Thráinsson 2007:39-40.)

Nilssen (2003) argues that V3 adverbs are derived by Remnant VP-fronting. We will not evaluate that claim here, but since V3 adverbs occur with inversion strings as well, they do not necessarily provide evidence against V-to-C movement; another option could be to consider them somehow incorporated with the verb. This hypothesis receives some support from semantics. To the extent that it is possible to provide a general characterization of the semantics of V3 adverbs, they seem to express an evaluation of the degree to which the event/action indicated by the verb took place, or alternatively, the degree to which it is appropriate to use the verb in question to describe the event;⁷³ Since our main concern

⁷²Thanks to Karen Dahl Hovind (p.c) for bringing this fact to my attention.

⁷³All of the V3 adverbs can also appear in a postverbal position in the middle field. Sometimes there is a semantic difference between a V3 adverb in postverbal and preverbal position. In (ia), there is a potential ambiguity between two readings which can be paraphrased as “He almost shouted (but did not)” vs. “He spoke so loudly that it could almost be described as shouting.” In (ib), only the second reading is available. If the adverb has incorporated here, (ib) would mean something like: ‘He almost-shouted’.

here is simply to document instances of linear non-V2 orders, we will not pursue the matter any further.

2.5.2 V3 with ‘maybe’-adverbs

There is another group of adverbs that also induce linear $V \geq 3$, but which are not V3 adverbs, since they appear in clause-initial position, exactly the position where V3 adverbs cannot appear (see section 2.5.1). The most widely discussed is the adverb *kanskje/kanske* – ‘maybe’ – which is attested in this pattern across Scandinavian (Platzack 1986; Rögnvaldsson and Thráinsson 1990; Faarlund et al. 1997; Thráinsson 2007). In some languages, like Icelandic and Swedish, this adverb is also frequently employed as a V3 adverb, while this pattern is more restricted in Norwegian (Bentzen 2014). In all of the languages, *maybe* can also be followed directly by the verb in accordance with the general V2 schema.⁷⁴ Interestingly, when used in linear non-V2 contexts, these adverbs demonstrably do not feature V-to-C movement at all in some languages, since the addition of negation and IP adverbs demonstrate that the verb does not move out of the VP.

- (66) [*Kanske*] [*Markus*] [*inte*] **vill** *ha* . . .
 Maybe Marcus not wants have . . .
 ‘Maybe Marcus does not want any.’ (Swedish, from Josefsson 2003:166)
- (67) [*Kanskje*] [*han*] **ikke** *vet* *det*.
 Maybe he not knows it.
 ‘Maybe he doesn’t know it.’ (Norwegian)

The failure of the verb to raise means that the word order is the same as in subordinate clauses. It has been pointed out that an explanation for this may be sought in diachrony, since the adverb ‘kanskje/kanske’ is really a contraction of a modal verb *kan* – ‘can’ – and a full verb *skje* – ‘happen’. This origin as a subordinate clause reveals itself in the optional addition of the complementiser after the initial adverb. The same applies to *kanhende*, another adverb meaning ‘maybe’ and also the result of a contraction between a modal and full verb.

- (68) a. [*Kanskje*] (at) [*han*] [*ikke*] **vet** *det* *det*.
 b. [*Kanhende*] (at) [*han*] [*ikke*] **vet** *det*.

Regardless of the diachronic explanation of these word order facts, it seems hard to argue that these clauses are subordinate clauses in a synchronic perspective, and we are therefore led to conclude that some main clauses in Scandinavian do not feature verb movement at

-
- (i) a. [*Han*] **ropte** *nesten*. (*Norwegian*)
 He shouted almost.
 ‘He almost shouted.’
 b. [*Han*] [*nesten*] **ropte**.

⁷⁴In fact, the matter is somewhat more complicated. In Bentzen (2014), speakers from four different locations in Eastern Norway (generally) accepted V2 after ‘kanskje’ (‘maybe’) with pronominal subjects, but rejected it when the subject was nominal. In Danish, on the hand, V2 order seems to be the only order generally accepted.

all. Moreover, such diachronic ‘islands’ might also be reinterpreted as a productive pattern. Interestingly, the pattern observed in (68) seems to extend to certain other adverbials with similar epistemic semantics, such as ‘muligens’ \approx *possibly*. This adverb is also attested without verb movement, and also optionally features an added complementiser, although it does not derive diachronically from the intersection of a main clause and a complement clause.⁷⁵

- (69) *Jeg ser ikke poenget med å oppføre seg slik hun gjør. [Muligens] [hun]*
 I see not point-the with to behave REFL.CL as she does. Possibly she
[ikke] ser det selv.
 not sees it herself.
 ‘I don’t see the point in behaving the way she does. Perhaps she doesn’t realize herself.’⁷⁶

There are two general points to bring from this discussion. First, that verb-second languages may not only feature deviations from the linear V2 pattern, but also main clause constructions which lack V-to-C movement altogether.⁷⁷ Second, that diachronic islands may persist for a protracted period of time without yielding to the more general pattern, and possibly even expand in some cases.

2.5.2.1 V3 with biscuit conditionals and other adverbial clauses

It was noted in section 2.3.1 that V3 orders in Germanic may arise through left-dislocation of an XP, but that this possibility is generally only possible on the condition that a resumptive element inside the core clause be coindexed with the dislocated phrase. In the case of initial subordinate clauses, a light adverbial of some kind may often fulfil this function.

⁷⁵One might attempt to derive these structures synchronically as embedded clauses by assuming ellipsis of the entire main clause except the adverbial (ia). This analysis would explain the word order and the possibility of adding a complementiser, but it leaves unexplained why this ellipsis is restricted to particular lexical items such as ‘muligens’, since the pragmatic recovery of the main clause should be just as straightforward with any other epistemic adverb (cf.(ib)). The semantic, and in consequence, syntactic extension from the ‘maybe’-class therefore seems more plausible.

- (i) a. *(Jeg tror) muligens (at) [hun] [ikke] ser det selv.*
 I think possibly that she not sees it herself.
 \approx “Maybe she doesn’t realize herself.”
 b. **(Jeg tror) definitivt (at) [hun] [ikke] ser det selv.*
 I think definitely that she sees it not herself

⁷⁶(Taken from <https://forum.kvinneguiden.no>)

⁷⁷There are also other main clause constructions, both in Scandinavian and Continental Germanic, which do not feature V-to-C. A common Scandinavian case is illustrated in (ia), while (ib) provides a similar example from German. To the extent that these are exclamative and do not really count as pure declarative clauses, they fall outside the focus of this thesis.

- (i) a. *[Bare] [hun] [ikke] kommer for sent!*
 Only she not comes too late
 “If only she doesn’t come to late!”
 b. *[Was] [der] [nicht] [alles] erzählt!*
 What he not everything tells.
 “All the things he tells!”

However, in some Germanic varieties like Standard German, there are also cases where an initial subordinate clause may lack a resumptive in the core clause. This is the case with so-called biscuit conditionals, which are semantically different from normal conditional constructions in that the speaker commits to the truth value of the proposition contained in the matrix clause regardless of the truth value of the antecedent conditional clause. In (70), the speaker asserts that there is juice in the fridge, an assertion which naturally holds regardless of whether the addressee is thirsty or not. It should also be noted that V3 is not only possible in German in such cases, but strongly preferred to V2 (Krifka 2017). This differs markedly from the situation in for instance Standard Norwegian, where V3 after biscuit conditionals are generally ungrammatical (71):

- (70) [*Wenn du Durst hast*], [*ich*] **habe** Saft *im* *Kühlschrank*.
 if you thirst have I have juice in.the fridge
 ‘If you are thirsty, I have juice in the fridge.’
 (From Csipak 2018:1)

- (71) * [*Hvis du er tørst*], [*jeg*] **har** saft *i* *kjøleskapet*
 if you are thirsty I have juice in fridge-the
 (Norwegian, intended meaning as in (70))

Furthermore, the same kind of resumptiveless V3 constructions are also encountered after other initial adverbial clauses, for instance after certain ‘because’-clause introduced by *weil*. Semantically, there is a clear parallel to the biscuit conditional case, since (72) expresses that the speaker does not believe there is a causal relationship between the antecedent ‘because’-clause and the proposition expressed by the matrix clause:

- (72) [*Weil du den Schlüssel nicht findest*], [*er*] **ist** *in der Schublade*.
 because you the key not find he is in the drawer
 ‘Since you cannot find the key, it is in the drawer.’
 (From Csipak 2018:2)

Csipak analyses such cases as involving modification of the speech act rather than the proposition contained in the matrix clause. Syntactically, this is expressed by construing the initial adverbial clause in a high projection ‘ActP’ which dominates the CP. In other words, this is another instance of the *peripheral* adverbial clauses (Haegeman 2007, 2010) which permit high attachments syntactically and which have greater scope in semantic terms. The difference from the cases which were observed earlier (cf. section 2.3.5) is that in this particular constellation, they appear before the matrix clause and induce linear V3 orders of a kind that is unexpected under the traditional analysis of verb second, since they violate the general ‘resumption’ condition that was held to be a fundamental constraint on CP-recursion.

2.5.3 Miscellaneous other linear $V \geq 3$ orders

In this section, we will review various kinds of $V \geq 3$ orders which are somewhat different from the ones reviewed in the previous sections, since it is unclear if all of them feature several constituents in front of the verb or not. In other words, the term ‘linear non-V2’ is potentially somewhat misleading here. Still, even if these phenomena might involve complex constituents of some sort, it is important to have an understanding of the contexts where

such complex constituents might arise, since we may expect to encounter similar cases in written corpora of dead languages like the ones we will be examining in subsequent chapters. In the examples in this section, the bracketing indicates the maximal amount of constituents one could imagine for any given structure.

Some of the cases to be considered are seemingly common to all Germanic V2 languages, others are restricted to a subgroup or even a particular language. Starting with the former case, it seems that all Germanic languages accept certain *focus particles* like ‘only’, ‘also’, ‘even’ to precede the constituent in the prefield. These can plausibly be analysed as involving a complex constituent with the focus particle somehow modifying the constituent to which it attaches.⁷⁸ However, other researchers have in fact postulated distinct syntactic focus positions in front of the prefield (Büring and Hartmann 2001).

- (73) [*Bare/også/til og med*] [*de yngste*] **kom** på forestillingen.
 Only/also/to and with the youngest came on play-the.
 ‘Only/also/even the youngest came to the play.’ (Norwegian)

- (74) [*Nur/auch/sogar*] [*die Harten*] **kommen** in den Garten.
 Only/also/even the hard come into the-ACC garden.
 ‘Only/also/even the hard make it into the garden.’
 (German, adapted from Müller 2018:56)

A related case is temporal adverbs like ‘never’, which may also precede the constituent in the prefield, acting as a modifier. This is not unsurprising when the constituent in the prefield is itself a temporal expression, but ‘never’ may also modify locative expressions or combine to create even more complex constituents:

- (75) [*Aldri*] [*før*] [*i Norge*] **har** det vært målt en høyere temperatur.
 Never before in Norway has it been measured a higher temperature.
 ‘Never before in Norway has a higher temperature been measured.’
 (Norwegian)

- (76) [*Nie*] [*zuwor*] [*in Deutschland*] **hat** sich jemand für eine Fernsehserie so
 Never before in Germany has REFL some for a television-series so
kopfüber in die Vergangenheit gestürzt . . .
 headlong in the past plunged.
 ‘Never before in Germany has anyone dived so headlong into the past because of a television series.’

⁷⁸This analysis receives support from the fact that these particles cannot modify pronominal arguments unless these are prosodically stressed and receive some particular IS prominence such as contrast. Furthermore, there is a clear parentage between these constructions and the V3 adverbs in that several of the latter group reappear here (such as ‘bare’/‘til og med’ in Norwegian). This suggests that what differentiates between V3 adverbs and focus particles is really only what they can modify semantically; V3 adverbs can modify the verb itself, while focus particles cannot. Some can do both; this is unsurprisingly the case for ‘bare’ – *only* – and ‘til og med’ – *even* – while ‘nesten/nærmest’ – *almost* fail to modify a normal DP since this would be nonsensical in most cases: *Almost the youngest came to the play. However, in exceptional cases these adverbs can in fact do the job of focus particles, providing the DP allows such modification:

- (i) [*Nærmest et barn*] **var** hun.
 Almost a child was she.

‘She was almost a child.’

(German, from *die Welt* online, 10.10.2017)⁷⁹

Apparent cases of multiple constituents in the prefield seem to be particularly frequent in German; for an instructive overview, see Müller (2018). More striking yet than the actual quantitative dimension is the heterogeneous nature of the phenomenon, as the possible combinations of different constituents are very high. Among the combinations reported by Müller are subject-adverb, adverb–subject, accusative object–PP, accusative object–adverb, dative object–PP, dative object–accusative object, PP–PP, support verb constructions–idiom chunks (Müller 2018:58–71). It is also possible to come across what seems like more than two constituents in the prefield:

- (77) [*Zum ersten Mal*] [*ein Trikot*] [*in der Bundesliga*] **hat** *Chen Yang* *angezogen*...
For-the first time a jersey in the Bundesliga has Chen Yang put-on...
'Chen Yang puts on a jersey for the first time in the Bundesliga...'
(German, from Müller 2018:72)

In spite of the plethora of such (apparent) multiple frontings, Müller claims there are clear restrictions on the phenomenon in terms of the relative order and secondly, the various XPs must always be clause-mates; if they originate in different clauses, the result is ungrammatical. Müller is therefore led to conclude that these examples really involve the fronting of some kind of complex verbal projection headed by a silent verbal head, (Müller 2018:81) and he goes on to present an analysis couched with the framework HPSG, which we will not review here. Let it suffice here to say that the hypothesis of some complex preverbal constituent seems plausible, in which case these examples are presumably compatible with the traditional analysis as well, involving neither a deviation from structural nor linear V2. The relevance of this section therefore lies in the message it sends (or should send) to a corpus linguist, namely that first appearances can be deceptive and that one must be wary not to dismiss all apparent cases of multiple preverbal constituents as incompatible with verb second.

2.5.4 V1 orders and empty prefields

All linear non-V2 orders considered so far have been $V \geq 3$ orders. In this section we will focus on V1 orders, which in a sense constitutes the opposite kind of problem to any theory that circumscribes an 'idealised V2 language' without deviations from linear V2. While much of the attention in the previous sections was on Mainland Scandinavian, this section will focus particularly on German, since V1 is a particularly prevalent option in that language.

There is a kind of linear V1 which seems to be available across all Germanic languages, namely 'topic drop' contexts where a deictic pronominal subject (78) or a continuity topic (79) is dropped in preverbal position in continuous discourse, yielding linear V1.

- (78) *Jeg tror ikke jeg kommer på jobb i morgen. Har litt feber.*
I think not I come on work in tomorrow have some fever.
'I don't think I'll come to work tomorrow. I've got some fever.'
(Norwegian)

⁷⁹<https://www.welt.de/kultur/plus169444371>

- (79) *A: Wat heb jij met die boeken gedaan?*
 what have you with those books done?
 ‘What have you done with those books?’
B: heb ik aan Marie gegeven.
 have I to Marie given
 ‘(Those), I gave to Marie.’ (Dutch, from Thrift 2001:63)

It is clear that such cases pose no great problem to the general theory of V2, since we may plausibly assume that these are instances where the initial constituent is simply left phonetically unexpressed, since it is easily recoverable from the situation or the immediate context.⁸⁰

However, there are other constructions where the idea of a silent initial constituent is not equally straightforward. A prominent case is the so-called *Narrative Inversion* construction of Icelandic. This word order pattern was particularly frequent in Old Icelandic, particularly after the initial conjunction *ok* – ‘and’ (Platzack 1985):

- (80) *Gengu Þeir inn og heilsuu fólkinu.*
 Walked they in and greeted people-the
 ‘They walked in and greeted the people.’
 (Icelandic, from Thráinsson 2007:349. Translation added.)
- (81) *ok kam hann Þangat, ok var Hoskuldr uti, er reið í tún.*
 and came he there and was Hoskuldr.NOM outdoors when rode into field.
 ‘And he came there, and Hoskuldr was outdoors when (he) rode into the field.’
 (14th century Icelandic, from Sigurðsson 1989:154)

It is unclear if Narrative Inversion should be seen as featuring some kind of null element in the prefield and whatever that null element might be; Holmberg claims ‘it is not inconceivable that the initial position is filled by a covert temporal adverbial particle ‘then’...’ (Holmberg 2015:353). Zwart, discussing verb-initial structures which he considers cases of Narrative Inversion in Modern Dutch, postulates an empty operator in SpecCP, adding in a footnote: ‘I will not be concerned with the question what the empty operator binds.’ (Zwart 1993:205, fn.20).

The most in-depth contribution to the question of verb-initial constructions in Germanic is presumably Önnarfors’ (1997) analysis of V1 declaratives in German. Önnarfors shows that V1 declaratives are not at all restricted to situations of topic drop or joke-telling, but rather range over a wide spectrum of different discourse functions, such as *narrative V1*(82a), which is the type used in jokes, *enumerative V1*(82b), *deontic modality V1*(82c), *content-explaining V1*(82d) and *exclamative V1*(82e).

- (82) a. *Kommt ein Mann ins Kaufhaus: “Ich hätte gern einen Regenschirm.”*
 Comes a man.NOM in-the store I had.COND willingly an umbrella.
 ‘A man comes into the department store: ‘I’d like an umbrella, please.’

⁸⁰That is not to say that the conditions governing topic drop are simple, or that there are no restrictions besides pragmatic recoverability; for a discussion of the facts in Dutch, see Thrift 2001.

(From Önnorfors 1997:101)

- b. *Wir müssen Fritz unbedingt im Krankenhaus besuchen. Hans ist betrunken. Anne hat kein Auto. **Bleibst** also nur noch DU.*
We must Fritz.ACC unquestionably in-the hospital bring Hans
ist drunk Anne has no car remains therefore only still you
'Wir need to visit Fritz in the hospital immediately. Hans ist drunk and Anne has no car. You remain the only option.'

(From Önnorfors 1997:132)

- c. *Die nächsten Jahre gammle ich. Mein Vater rechnet damit. **Soll** sich der Junge doch austoben', sagt er...*
The next years waste I my father.NOM counts therewith should
REFL the kid.NOM PRT indulge says he...
'The next years I will squander away. My father is expecting it. Let the kid have his fling, he says...'

(From Önnorfors 1997:136)

- d. *Ich begab mich auch nicht mehr gern ins Lehrerzimmer, **wusste** ich doch, wie Rolf dort den Ton angab.*
I proceeded REFL also not more willingly in-the teachers-room knew
I PRT how Rolf there the tone set.
'I didn't feel like venturing more into the teachers room, knowing how Rolf set the tone there.'
- e. *Mann, **haben** wir gelacht.*
Man have we laughed.
'O boy, how we laughed.'

(From Önnorfors 1997:171)

Önnorfors argue that these constructions are very old, predating the V2 stage of Germanic, and possibly even a relic of the marked V1 declaratives which have been reconstructed for Proto-Indo-European.⁸¹ If this is correct, they have lived alongside the emerging and ever-expanding V2 construction for over a thousand years without succumbing to it. Moreover, unlike what purportedly is the case with Narrative Inversion in modern Icelandic, these constructions are all typical of spoken language, and therefore highly relevant to generative theories of grammar. Önnorfors goes to great lengths to demonstrate that these constructions are not derived from the V2 construction by omission of an element in the prefield, since the addition of all candidate elements such as expletives or light adverbs or particles

⁸¹See also Miller (1975) for a hypothesis that VSO was in fact the older Indo-European word order which survived as a marked word order along the emerging SOV order.

either change the discursive appropriateness of the construction or in some cases even render it unacceptable.⁸² Furthermore, the discursive uses of the construction are much too varied to warrant the assumption of an operator in the prefield, at least on the reasonable assumption that an operator should contribute something beyond that of saving the V2 analysis. Therefore, Ötnerfors concludes, the construction features V-to-C movement, but no preverbal specifier at all.⁸³

Ötnerfors' solution is completely compatible with current Minimalist conceptions of phrase structure as projected from lexical items as well as the general Minimalist desideratum of keeping structures at a minimum. Furthermore, I would like to suggest, in line with Ötnerfors' conclusion, that it presumably is *crucial that these constructions lack a specifier*, and that the inappropriateness resulting from the addition of a semantically light element in the prefield results precisely from the failure to respect this condition. In other words, the construction as a whole seems to be intimately connected to the (admittedly somewhat vague) notion of clause-typing; the discursive dependency between the clauses is ensured by verb movement without concomitant XP movement.⁸⁴

If one accepts Ötnerfors' conclusion, the relevance of these V1 declaratives lies particularly in the fact that they not only feature surface linear non-V2 orders, but that they also do not share the syntactic structure of V2 declaratives. More concretely, they seem to lack the EPP-feature of the normal declarative clause of a V2 language.⁸⁵ While this construction has been largely lost in the other branches of Germanic, it seems to live happily along the V2 construction in modern German.

2.6 The approach adopted in this thesis

In this section I will clarify the general framework and leading assumptions that will be adopted in this thesis. Section 2.6.1 offers a discussion of the problems in finding a theoretically satisfactory definition of a V2 language, while section 2.6.2 clarifies the definitions that will be adopted in this thesis. The general framework and in particular the assumptions regarding the acquisition of phrase structure will be the topic of section 2.6.3. Section 2.6.4 discusses the role accorded to cartographic hierarchies in this thesis.

2.6.1 Why it is hard to define a V2 language

The preceding sections should have made clear that there are still several unresolved issues in verb-second theory. In particular, the linear restriction has not yet been derived in a

⁸²Compare for instance a variant of (82b) with expletive *es* in the prefield:

- (i) Wir müssen Fritz unbedingt im Krankenhaus besuchen. Hans ist betrunken. Anne hat kein Auto.
*Es **bleibst** also nur noch DU.

Likewise, in (82d) there is presumably no element that could be added without rendering the clause unacceptable, and in (82c) and (82e) the addition of an adverbial like 'dann' or 'so' – *then, so* – or similar elements would alter the meaning significantly.

⁸³The same conclusion is also reached by Brandtner (2004), who sees V-to-C movement as intimately related to Force-marking, or clause-typing.

⁸⁴It is very tempting to see verb-initial embedded clauses like conditionals or 'als ob'-clauses as intimately connected with the declarative V1 construction rather than the V2 construction, see section 2.3.

⁸⁵This is perhaps possible to integrate with the V2 grammar if we assume, as is common in current theorizing, that the EPP is not a feature itself, but a property of a feature. On this view, one could assume that the verb may occasionally raise to check some formal 'clause-typing' feature in C, and that this feature is not associated with an EPP-feature.

satisfactory manner, nor has it been shown to be a superficial trait of variation. In a sense, V2 is less epiphenomenal than it used to be. For this reason, we will try to reassert the importance of the linear restrictions that have both given the name to the phenomenon and made it a typologically rare thing. At the same time, we want to avoid making explicit reference to linear order in the definition, since this is both theoretically unattractive and empirically problematic.

We must distinguish very clearly between two different notions when providing a definition of V2, namely the notion of a *V2 construction* and a *V2 language*. I believe the latter notion is theoretically more problematic than commonly assumed. The problem is as follows: a theoretically stringent definition of a V2 language would have to be explicit enough to make it possible, at least in principle, to decide by purely empirical means if a given language L is a V2 language or not. As it turns out, the problem is not to provide such a stringent definition; in fact, several definitions are imaginable. The problem is rather that on every such stringent definition, the resulting class of ‘V2 languages’ is either an empty set, as all Germanic V2 languages simply fail to qualify, or the class does not include all languages intended. Let us explore briefly why this is the case.

We already reviewed and rejected a completely stringent definition in section (2.2), repeated here for convenience:

- (83) In a verb-second language, the finite verb must appear as the second constituent of all declarative main clauses.

This definition does not work at all,⁸⁶ since all V2 languages feature perfectly grammatical linear non-V2 orders (both V1, V3, and others). As a consequence, linear order will not be part of the definition of a verb-second language. We will therefore make reference to structure, rather than linear position. This opens up the possibility for another stringent definition:

- (84) In a verb-second language, the finite verb moves to C^0 (or its equivalent in a cartographic LP) in all declarative main clauses.

This is in essence the definition adopted by many linguists working on the Old Romance languages (Benincà 2006; Ledgeway 2008; Wolfe 2015b), but as already pointed out, it might potentially also capture a great amount of VSO-languages. As for the Germanic family, it is not obvious that this definition captures any language. Even when disregarding the fact that the structural status of subject-initial clauses is far from resolved, there are main clauses which ostensibly do not feature V-to-C movement in at least all the Scandinavian languages, where clauses introduced by the adverbials *kanskje*, *kanhende* – ‘maybe’ – may freely leave the verb in a VP-internal position. In addition, this definition only makes reference to V-to-C, providing no restrictions on the prefield.

Since we want to bring restrictions on the prefield into the definition, we could adapt our definition as follows:

- (85) In a verb-second language, the finite verb moves to C^0 (or its equivalent in a cartographic LP) and an XP moves to a left-peripheral specifier position in all declarative main clauses.

⁸⁶It could be that this definition captures other languages which are not considered V2 languages, such as Warlpiri, where the only obligatory word order rule according to Hale is that the auxiliary must be the second constituent (Hale 1983).

This definition excludes VSO languages with V-to-C movement, but unfortunately, it also excludes the Germanic languages. The Scandinavian V2 languages are excluded for the same reason as before (lack of V-to-C in some main clauses), while Continental Germanic, or at least German, presumably does not feature XP movement to a left-peripheral specifier position in all cases, as convincingly argued by Ötnerfors (1997). But since the word order patterns that fall outside of this definition somehow constitute ‘marked’ constructions, one might try another definition:

- (86) In a verb-second language, the unmarked word order features V-to-C movement (or its equivalent in a cartographic LP) and an XP moves to a left-peripheral specifier position in all declarative main clauses.

This definition is potentially disastrous, since the unmarked word order must be considered subject-initial clauses, and as already mentioned, it is not beyond doubt that these feature V-to-C movement in all (or even any) Germanic V2 languages.

It should be clear that a V2 language is strictly speaking more an ideal than a reality, but before concluding, it is worth mentioning a definition provided by Holmberg (2015). He first gives the following definition of the ‘V2 property’:

- (87) a. A functional head in the left periphery attracts the finite verb.
 b. This functional head wants a constituent moved to its specifier position.
 (From Holmberg 2015:375)

Holmberg suggests that property (87b) should be formalised as a ‘generalized EPP-feature’. He then explicitly raises the question if *verb-second language* is a well-defined notion, answers the question in the affirmative, and claims a V2 language is a language which has the two properties (87a–87b), adding that ‘the EPP feature can only prevent V3 (V4, V5, etc.) order derived by *movement*. It does not prevent V3 order derived by external merge.’ (Holmberg 2015:376)

Holmberg’s definition is theoretically quite involved. It does not provide strong restrictions on the prefield, since it allows $V \geq 3$ orders to arise through base-generation. In other words, it opens up for the inclusion of ‘relaxed V2’ systems into the definition. This is a valid definitional move, of course. A more problematic aspect of Holmberg’s definition to my mind is that it incorporates several assumptions about V2 languages which are not beyond doubt. First, it is not at all clear that the verb moves to the left-periphery in subject-initial clauses across Germanic, whereas it is clear that it does not always move to the left-periphery. Second, even in the cases where the verb does move to the left-periphery, it is not obvious that it always carries an EPP-feature, as demonstrated by the work of Ötnerfors (1997). Thirdly, the assumption that linear V3 orders can only arise through base-generation is also questionable, since a common analysis of contrastive left dislocation (CLD) constructions is in fact that of a movement dependency. Finally, if Holmberg only means to say that a V2 language is a language which *features* these properties, we would have to include ‘residual V2’ languages as well, and presumably many others.⁸⁷

⁸⁷Holmberg also consistently refers to Icelandic and Yiddish as ‘I-V2’ languages. Although he emphasizes that this does not entail a commitment to a specific analysis (p.356), his definition of V2 on the other hand does commit to a specific analysis. In case it should turn out that the V-to-I analysis of Icelandic and Yiddish is correct, these languages would in consequence no longer qualify as V2 languages on Holmberg’s definition.

2.6.2 The definition of V2 adopted in this thesis

On the whole, it does not seem possible to provide a completely stringent notion of a verb-second language that captures the languages we want to include while excluding the ones we wish to exclude. We shall have to make do with a less than stringent definition that still seems intuitive and reasonably adequate. In light of these considerations, I will henceforth adopt the following definitions:

- (88) *A verb-second construction is a construction where:*
- a. The finite verb lexicalises a head position in the C-domain, and
 - b. there is a single fully productive A' projection preceding the verb which can and must host a single XP.⁸⁸
- (89) *A V2 language is a language where:*

inversion structures are V2 constructions as defined in (88). V2 languages may occasionally feature inverted linear non-V2 orders which are licensed by particular lexical items or in particular constructions.⁸⁹ If these constructions are reanalyzed as involving another productive left peripheral projection, the V2 status of the language is lost, although V-to-C movement may persist.

The most important thing to notice about this definition of a verb-second language is that it does not say anything about the structural status of subject-initial clauses. This is necessary if we want to maintain that linear order does not play a direct role in the phenomenon and to avoid incorporating a premature assumption into the theory itself, namely the assumption that all main clauses are necessarily CPs in all Germanic V2 languages. Subject-initial clauses can either involve movement of the subject and the verb to the CP-domain, in other words be V2 constructions, or not; this does not affect the status of the language as a V2 language. In the case where subject-initial clauses are mere IPs, the linear second position of the verb is accidental and theoretically extrinsic to the V2 syntax of the language. From this follows that verb-second languages may in principle be verb-final in subject-initial clauses, as would be the case in German or Dutch if these languages did not feature any kind of verb movement in subject-initial clauses. In such a scenario, where the projection hosting the verb in subject-initial clauses is head-final (or the structural equivalent in approaches that adopt the Universal Base Hypothesis), the amount of linear V2 in main clauses might be expected to be relatively low, although I am familiar with no such language.⁹⁰ If that projection is head-initial, on the other hand, the linear V2 output of the language will be very high.

⁸⁸This definition is deliberately stated in neutral terms with regards to the representational/derivational divide. The term A'-position is used descriptively to mean a position that does not impose any grammatical function on its occupant, without any deeper accompanying theoretical claim.

⁸⁹Admittedly, it is not straightforward to draw a line between productive patterns and 'particular constructions'. This just further underscores the difficulty in providing a stringent and relevant definition of a V2 language. For instance, it is unclear if the 'V3 adverbs' of the Scandinavian languages are a productive or a lexically specified class.

⁹⁰In principle, the situation could also arise where a main-clause V2 syntax produces significantly less linear V2 than a non-V2 embedded syntax. This could happen if subject-initial clauses are head-final while embedded clauses are head-initial. Once again I am not familiar with such a language, but as will be seen in chapter 3 and 4, Old French is indeed a language where the embedded V-to-I syntax produces more linear V2 than the main clause syntax (without anticipating the conclusion regarding the status of that syntax).

It is clear that on a strict definition like this, there is no such thing as a ‘relaxed V2’ language; all V2 languages are by definition strict. This need not blind us to the fact that there is a typologically and theoretically relevant difference between languages which feature widespread subject-verb inversion of the Germanic kind without obeying the linear restrictions of V2 languages, and languages which do not normally license subject-verb inversions of the Germanic kind at all. This is the difference between most Old Romance languages and their modern descendants. But rather than saying that the former are ‘relaxed V2 languages’, we will simply say that they feature V-to-C movement – if and only if that can be demonstrated to be the case. Thus, verb-second languages are a sub-group of V-to-C languages. I believe that one positive outcome of this definition is that it makes it possible to quantify the notion of a verb-second language rather than argue over a theoretically spurious dichotomy. The more V2 constructions a language contains, or the more domains/clause types that have been conquered by the V2 construction, the ‘more V2’ the language is.⁹¹

As far as the distinction between ‘symmetric’ or ‘asymmetric’ V2 languages is concerned, this is not something that should be part of the definition. Rather, it is an empirical question to be solved through data analysis. However, I believe the theoretical developments of recent years and in particular the empirical investigations into alleged ‘symmetric’ languages like Icelandic or Yiddish call for a reassessment of the typology of V2. I therefore fully concur with those researchers who are skeptical of the existence of symmetric V2 languages; not because it is conceptually implausible that such a language should exist, but simply because the evidence in favour of that hypothesis seems rather weak. I believe the null-hypothesis is that V2 languages are by nature asymmetric and that V2 constructions are root phenomena. This is exactly the same conclusion reached by Walkden and Booth, who raise the following interesting question:

‘Should the typology of V2 be rethought? A natural and restrictive hypothesis would be that there is only one type of V2 language, with variation – insofar as it exists – being attributable to universal properties of the mapping between syntax and information structure, and to idiosyncratic properties of individual lexical items such as complementisers and complement-taking verbs...’ (Walkden and Booth to appear)

I concur with this reasoning, but at the same time, I am concerned that this hypothesis in actual practice might potentially run into some circularity. The strong and interesting hypothesis of an isomorphic relationship between syntax and information structure is relativized by two possible loci where ‘lexical idiosyncrasy’ might counter the effects of the former, namely the lexical properties of complementisers and verbs. This leads us to ask the following question: given a language L whose embedded word order patterns differs from the ruling assumption about the universal relationship between syntax and IS, how can we know: A) If the lexical idiosyncrasy is in the verb or the complementiser (or both)? B) If the deviant word order is in fact due to lexical idiosyncrasy at all, and not rather direct counterevidence to the syntax-IS-isomorphism hypothesis? In short, the problem with C-selectional properties (or other idiosyncratic lexical properties) is that, once taken as a primitive, they become virtually impenetrable to further analysis.

In order to avoid the lure of sweeping-under-the-carpet of lexical idiosyncrasy, we will

⁹¹This also means that the diachronic stability of V2 languages is expected to correlate with their ‘strength’ as a V2 language, since fewer exceptions gives less fertile ground for reanalysis. It seems intuitive to imagine that such factors may at least be partially responsible for the greater diachronic stability of the Germanic V2 languages than the Old Romance inversion systems.

adopt a stronger hypothesis that runs in the other direction. The foundation is pleasantly solid, namely the semantics of lexical expressions. While it is not empirically trivial to establish complete synonymy of two expressions in different languages, it is still an empirical question. Therefore, I will assume that the semantics of verbs, their s-selectional features, determines their c-selectional features.

There is too much counterevidence to this hypothesis to leave it without further modification. In fact, some of the viaduct verbs are known to take non-finite complements like AcI constructions in many languages, clauses which are normally severely truncated. It might be that a distinction must first be made between finite and non-finite complements; the hypothesis would then be that finite complement clauses under viaduct verbs are structurally larger than finite complement clauses under non-viaduct verbs.

The general theory of V2 may seem to suggest that even this hypothesis is too strong, in light of the fact that viaduct verbs can show lack of V-to-C in all Germanic languages. However, such a conclusion would probably be misguided. We have seen considerable evidence that lack of V-to-C does not necessarily mean that ForceP is not activated, since other root phenomena like discourse particles or interjections may co-occur with non-raising verbs. The selection of ForceP, in cartographic terms, is a prerequisite for V-to-C, but does not automatically lead to it (apart from in German, if the complementiser is dropped). The relationship between S-selection and C-selection is clearly not isomorphic, but we can still maintain the hypothesis that C-selection of Force and embedded root phenomena depends on the semantics of the matrix verb.⁹² Concretely, viaduct verbs (classes A, B and E) universally permit embedded root phenomena because they universally select high complementisers, provided the language allows finite complementation at all, presumably an independent property. The complementisers themselves are mere conduits, having at most rudimentary lexical content which is overwritten or specified by the properties of the selecting verb. Under these strong assumptions – and they must of course be supported by independent evidence – the syntax-IS-isomorphism is established as an empirical testable hypothesis rather than a credo.

It is important to emphasize that if we find counterevidence, this does not necessarily prove that the syntax-IS-isomorphism hypothesis is wrong, since we must also envisage the possibility that languages differ with respect to phrase structure. Although this flies in the face of certain strong cartographic tenets, I believe a difference in phrase structure is the preferable theoretical locus to account for cross-linguistic variation, far better than assuming that languages differ with respect to the way locality works, or that this variation is due to lexical idiosyncrasy of various sources, an hypothesis which is bound to be very hard to test empirically. This also allows us to test if truly symmetric V2 languages exist or not.

2.6.3 The general framework

The general framework adopted here is a strongly empiricist, non-parametric version of Minimalism.⁹³ As a consequence, we will be fundamentally concerned with the following

⁹²The empirical challenge is of course greatly exacerbated in working with dead languages. Still, it should not be insurmountable, as long as the direction of the analysis is clear: s-selection determines c-selection, which means that one cannot use the observed c-selectional properties of a verb to draw conclusion about the semantic class of the verb unless there is independent evidence for it.

⁹³For an accessible introduction to current Minimalist theorizing, see Adger (2003). I will not go into details about the reasons for rejecting a parametric approach. Let it suffice to say that they are essentially the same as those offered by Newmeyer (2005) and Boeckx (2014). See also Gallego (2011) for a useful discussion.

question: what kind of evidence do children need to construct a V2 grammar?

I will adopt the hypothesis that children *build structure*, and furthermore that they are conservative structure builders, only adding projections and expanding the clause when faced with clear evidence. Concretely, I will assume that children behave according to a principle which I will dub *The String-Structure-Assignment-Principle* (SSAP) and which can be defined as follows:

String-to-Structure-Assignment-Principle (SSAP): Children assign the minimal structure that is consistent with the global string input in a maximally economic way.

The qualification *minimal structure* is important, as I will assume that children generally only assign the minimal structure that is consistent with the global input.⁹⁴ This means that the child will only construct a V-to-C analysis when the word order facts dictate or at least strongly suggest such an analysis; other things being equal, a V-to-I analysis is preferable to a V-to-C analysis. The notion of *global input* is meant to express the hypothesis that children, if possible, assign a single structure that accounts for all of the word order patterns in the input. On this view, the acquisition process is considered a bottom-up-process in the most literal sense, a constant revision of a single structure that will start out in minimal form and then gradually expand. The fundamental question of the following chapters, then, is quite simple: how far did the Old French and Late Latin declarative main clause extend?

2.6.3.1 The role of quality vs. quantity in acquisition

A natural question to ask at this point is how much weight the SSAP places on the role of quality vs. quantity in the acquisition process. It should be clear from the definition given that quality plays a crucial role; in the specific case of V2, inversion strings – and in particular G-inversion – are highly important, although they do not in and of themselves constitute unambiguous evidence for V-to-C movement. In fact, much of chapters 3 and 4 will be devoted to a discussion of the proper and most natural interpretation of G-inversion structures in Old French.

This emphasis on quality does not mean that children are unaffected by the quantitative dimensions of the input. Strings which are rarely heard, for instance due to production mistakes, will be dismissed due to their low quantity. It might be tempting to believe that such strings are still disregarded because of their quality, for instance if they deviate from an otherwise consistent input, but this only begs the question: what makes some strings consistent and others deviant in the first place? If one and the same ‘mistake’ is produced quite consistently in the presence of a child – say an ungrammatical violation of linear V2 in modern German – then this string is not deviant at all, but rather seems to conform to the norm of the language from the perspective of the child and will certainly be internalized accordingly. In other words, it seems that quantity must play some role.

There is no agreement on the magnitude of frequencies that constitute a triggering experience for V2 in the acquisition literature. Lightfoot suggested on the basis of corpus studies of Continental Germanic that 30% of main clauses should feature inversion strings for the child to deduce a V2 grammar (Lightfoot 1999:41). Other studies have suggested

⁹⁴As pointed out to me by George Walkden (p.c.), this is essentially an acquisitional counterpart of the Principle of Economy of Structure in LFG. There is one important difference, though, namely that I adopt the more conventional view that the X-bar schema is somehow basic (or at least generalized from cases where there is full evidence for it), so when children discover evidence for, say, a head position, they postulate the existence of the specifier as well, or vice versa.

that the threshold is significantly lower; Yang (2003) reports a figure of 23% of the relevant inversion strings in his study of L1 acquisition of Dutch, while Westergaard finds only 13.6% inversion in her study of child-direct speech in Norwegian (Westergaard 2009:67). Since Yang and Westergaard's data come from acquisitional studies, they seem more directly relevant than Lightfoot's corpus-based conclusions. It is worth emphasizing that all these authors consider inversion strings crucial.

Rather than adopting some arbitrary threshold figure, it seems plausible to assume that the required frequencies of a particular construction will correlate to some extent with the global input. Concretely, the threshold for constructing a V2 grammar on the basis of a type of evidence such as inversion will presumably be lower if the non-inverted input is also easily consistent with a V2 grammar. On the other hand, if there is much 'noise' or apparent inconsistency in the non-inverted input, it does not seem unreasonable to suppose that the child needs higher frequencies to construct a V-to-C parse, since this might entail that children must work with several competing hypotheses at the same time. Any given construction has a maximally low quantity at the threshold at which children first manage to parse clauses. At this point, every construction is new and probably by default considered grammatical. As the child increases its understanding of the hierarchical structure of the clause in the face of growing evidence, former constructions which fail to be reproduced regularly come to be degraded from 'grammatical' to 'marginal' or eventually even 'ungrammatical.'

Observe that there is a potential tension inherent to the SSAP as stated above which is completely intended. This is the potential tension between the notion of *minimal structure* and *maximally economic way*. Consider concretely the case of subject-initial (i.e. non-inverted) main clauses in the Germanic languages and recall the debate of their structural underpinnings (V-to-I as argued by Travis (1984) and Zwart (1993, 1997) or V-to-C as argued by Schwarz and Vikner (1989, 1996)). The SSAP is in fact consistent with both analyses. A priori, the central tenet of *minimal structure* suggests that the default is the simpler V-to-I parse. On the other hand, the proviso *maximally economic way* suggests that, once the V-to-C parse is required from inversion strings, the V-to-I parse of subject-initial strings is relinquished, since the latter strings are also consistent with a V-to-C parse. V-to-C is therefore consistent with the *global input* in a way that V-to-I is not, and it is possible that this leads to the elimination of the rule of V-to-I from the grammar. But then again, it is equally possible that both rules co-exist side by side and are called upon individually to create inverted and non-inverted structures. At our current level of knowledge, we presumably do not want to preclude the issue. Moreover, this is a prime candidate for a locus where different acquirers might make different decisions, with some selecting a split parse and others adopting a unified parse. Such *underspecification* of the input provides an interesting opportunity for subsequent reanalysis, and it therefore seems beneficial to keep this slight tension in the SSAP.

2.6.4 Cartographic hierarchies and phrase structure

The SSAP is only a hypothesis about the behaviour of language acquirers, not a principle that can actually guide children in the acquisition process itself. However, it seems necessary to assume that children are guided by some principles that constrain the possible grammars they can construct from the input. One natural and highly relevant question is what role is played by cartographic hierarchies or 'functional sequences' in this process.

There is no single, clear answer to this question forthcoming from the cartographers themselves. Rizzi has pointed out on several occasions that it is possible to consider carto-

graphic hierarchies as both *explanans* and *explanandum* of linguistic theory and furthermore made clear that the latter option seems more natural, emphasizing that the actual mapping of the projections through empirical investigation (drawing up the proper ‘roadmap’) is merely a precursor to more explanatory analysis. This makes sense: in order to explain something, one needs to know *what to explain*.⁹⁵ For this reason, Rizzi objects to the use of the word ‘template’ by Abels (2012) to describe cartographic sequences, stressing that nobody has suggested that the sequence itself is a primitive without need of further explanation, stressing the potential role of ‘*interface conditions or independent grammatical properties*’ in this respect (Rizzi 2013:213, fn.4 - italics added). Still, one cannot help feeling that there is some curious tension between this reasoning and the assumption that the cartographic sequence is innate (in any possible sense of the word); if ‘independent grammatical properties’ (such as locality) can explain the sequence, we effectively do not need to assume it to be innate, it will arise *in syntax* itself. As for *interface conditions* (these must necessarily be the C-I interface), this would suggest that our external cognitive systems are incapable of processing and interpreting anything less than a rigidly ordered sequence of XPs, including various kinds of adverbials in the IP field. This is possible, but to my mind not probable.

We will therefore keep cartographic roadmaps of the LP in the back of our minds, but it will be assumed that children need evidence for it. This leads us to ask exactly what constitutes evidence for a syntactic position. The answer adopted here is simple: only word order facts lead children to construct syntactic positions. This means that prosody cannot count as evidence to *create* a syntactic position. As an illustration, consider again the prefield in Germanic V2 languages. The prefield can host various adverbials, topics, foci and expletives. Children are of course sensitive to the different interpretive properties of these elements, and their intonational properties are far from identical; foci come with a different intonation contour than topics. Yet none of these elements can co-occur in the prefield, meaning there is no evidence that they lexicalise different projections.⁹⁶ Since we have seen that full-fledged cartographic approaches, aided by locality assumptions of various kinds, are incapable of getting the word order facts of Germanic V2 languages straight, I will assume that, confronted with such a situation, children start *syncretising* a position, meaning they add different features to the same projection (Giorgi and Pianesi 1996; Hsu

⁹⁵This does not mean that explanatory processes should not run in parallel with the purely empirical work; Abels (2012) argues that most of the co-occurrence pattern in the Italian Left Periphery can be deduced from locality effects through Relativized Minimality. Rizzi (Rizzi 2004, 2013) has suggested that the uniqueness of the left peripheral focus position is due to interpretive clashes that would arise at LF if focus were allowed to reiterate, since this would force the lower focus to be part of the presupposition of the higher focus.

⁹⁶There is evidence for positions to the left of the prefield, though, as in the case of various LDs. This will of course lead to the creation of a projection.

2017).⁹⁷ ⁹⁸⁹⁹ This means that the cartographic tenet of ‘One-Feature-One-Head’ will not be adopted as an a priori assumption.

I will borrow some insights from cartography, though, notably the Principle of Transitivity which is the cornerstone of cartographic reasoning. However, I will only consider it to be an innate *parsing preference*, a third factor in the sense of Chomsky (1995), rather than something hardwired into UG or something that follows from universal properties of phrase structure. To this principle is also added a general theory of embeddability:¹⁰⁰

- (90) *Principle of Transitivity*: If A precedes B and B precedes C, then A precedes C.
 (91) *Theory of embeddability*: If X is embeddable and X precedes Y, then Y is embeddable.

As has been pointed out in the literature, the principle in (90), which is the foremost day-to-day tool of a cartographer in action, must be assured by some restrictive theory of phrase structure. The weapon of choice for most cartographers has become the Linear Correspondence Axiom (LCA) of Kayne (1994), which holds that linear precedence is determined by asymmetric C-command, in turn the result of the X-bar schema being universally constrained in such a way that heads precede complementiser and specifiers precede heads. Furthermore, only one specifier per projection is permitted, and adjunction is generally not possible. Under these assumptions, the Principle of Transitivity follows without further assumptions.

I will not adopt the LCA here. The price to pay for this strong uniformity assumption is a considerable increase in movement operations, including ‘roll-up’ operations which often

⁹⁷Any C-head selected in a given derivation must not carry several inconsistent features at the same time, of course.

⁹⁸Note that there might be a more limited role to play for prosody, in the sense that prosody might potentially identify a syntactic position in case of ambiguity. Let us say that a child has parsed subject-initial clauses into an IP and inverted clauses into a CP, also noting that foci in the prefield come with a particular intonation. Then they are confronted with a subject-initial clause like (i), where the subject carries focal intonation.

- (i) *MOREN min jobber på skolen, (ikke faren min).*
 mother-the my works at school-the not father-the my.
 “My mothers works at the school, not my father.”

It is conceivable that the child interprets such clauses as involving the (independently established) focus position in the CP based on the prosodic cue. However, this matter is complex, since it involves the more general ‘focus-in-situ’ property of the Germanic languages, and since we will be concerned with written corpora of dead languages where prosodic information is not available, we will leave the matter aside.

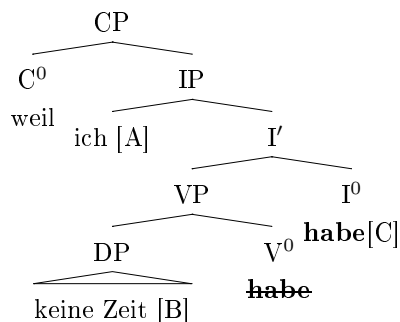
⁹⁹What about the role of morphology in the acquisition process? This does not play a role in the debate of Germanic V2, but suppose for the sake of argument that left-peripheral foci in Germanic V2 languages were associated with a C-particle, as is the case in for instance Gungbe (Aboh 2004). In this case, the approach adopted here assumes that children still syncretise foci into the same position as other preverbal elements, with the important difference that they assign additional morphology to the focus feature in C⁰.

¹⁰⁰There is some counterevidence to such a theory of embeddability. For instance, discourse particles in the German languages appear in the middle field, not the prefield (where they are generally banned). Under a simple theory of embeddability like the one in (91), the prediction is that discourse particles should be free to appear in any finite embedded clause. This is not the case, since many discourse particles are not acceptable under non-viaduct verbs. If we assume that discourse particles depend on notions like *independent illocutionary force*, again dependent on the projection of a ForceP in syntax, this is not surprising. A possible explanation is that Force⁰ must enter in some kind of long distance relation (perhaps Agree) with a projection in the middle field, or alternatively, one may postulate covert movement. Neither of these solutions seems very satisfying to me. In either case, the principle of embeddability is slightly too simple in such cases. I will simply disregard such cases and consider (91) to be valid in most cases.

lack obvious triggers and are employed only to derive surface word order, including the most common word order in the world, SOV.¹⁰¹ Rather, I will adopt the common Minimalist assumption that movement is costly and should preferably add something significant to the derivation. At the end of the day, such considerations boil down to the most natural way of accounting for cross-linguistic differences; the stance adopted here is that the preferred locus is phrase-structure, both by allowing some leeway in the linearization properties of the X-bar schema (notably by reintroducing the Head Parameter) and in the clausal architecture itself.¹⁰²

As a consequence, the Principle of Transitivity does not follow entirely. That is intended, since it represents nothing more than a parsing preference, presumably derived from the universally linear nature of the input. To take the hierarchy of adverbial positions in the IP (Cinque 1999 *et seq.*), one might assume that these XPs will be parsed automatically into left-branching specifier projections, since this represents the default parsing by PT and since there is no counterevidence. I will therefore adopt the assumption that adjunction is banned or at least severely restricted. On the other hand, I will assume that children have sufficient resources to override this principle, such as recognition of scope. An embedded German clause illustrates the point. In the surface form of (92), A precedes B and B precedes C, yet B is an argument of C. I assume that children construe the verb in a head-final projection in such cases, meaning B is c-commanded by C, not the other way around.

- (92) ... weil [ich] [keine Zeit] **habe**.
 since I no time have.
 ‘... because I don’t have the time.’



In other words, hierarchical phrase structure cannot be directly read off linear order in the approach adopted here. This will be particularly important when discussing the Latin

¹⁰¹Abels and Neeleman (2012) point out that unless coupled with a concrete and restrictive theory of movement, the LCA does not lead to an interesting theory of word order. Furthermore, they show that the LCA does not derive the X-bar schema, let alone the specific Kaynean instantiation of it, *Specifier-Head-Complement-Hypothesis* (SHCH). The authors demonstrate that an approach without the SHCH is equally successful in deriving the cross-linguistically attested and unattested word orders in the extended projection of the nominal projection, only by adopting a ban on rightward movement and allowing more base generated orders in accordance with the traditional X-bar schema.

¹⁰²These comments might suggest that a full-scale revision of the syntax-morphology interface is needed as well. This does not have to follow, though; the point is that the different projections of a clause have different roles to play, and as a consequence, it is perfectly possible to assume that the formal projections of the clause responsible for assigning morphology to the verb and its arguments are universal (and even universally ordered, if one is so inclined), while the so-called A-bar projections have more leeway. This has been common practice for decades in non-cartographic research anyway.

evidence in chapter 5, which is characterised by a very high degree of *structural ambiguity*.

2.7 The methodology of the annotation

The texts that were analysed in this thesis will be presented in the respective chapters on Latin and Old French. All of the texts were available in electronic corpora, permitting them to be extracted without the risk of error inherent to manual transmission. The following is a short description of the general procedure of the annotation; for a full description of every aspect or technicality, the reader is invited to consult the user manual which resides along with the data files in the TROLLing Repository at the University of Tromsø (Klævik-Pettersen 2018).

2.7.1 Annotated categories

The annotation was made using Excel. Two different files were created, one for Latin and one for Old French. For each file, the annotation of the two texts were kept as a separate worksheet carrying the name of the text and a colour code. In addition, each file contained a third worksheet, called ‘Tables’, where various quantitative data were calculated and presented for each text. The tables and cell areas were marked with the same colour code as the corresponding text, and the tables were marked with the same number as they carry in this thesis; for instance, table 3.1 in the next chapter is found with the corresponding number in the Excel-file for Old French.

In the annotation, the most important categories were the following: clause type, linear sequence of grammatical functions, string type, type of predicate, linear position of the verb, discourse status of the subject, embedder. The clause types were : main clause, main clause interrogative, main clause imperative, complement clause, adverbial clause, embedded interrogative, relative clause, conjunct, fragment. For Latin, two-non finite clause types were also added: participial clauses and absolutive clauses.¹⁰³ Only declarative main clauses, complement clauses, adverbial clauses, embedded interrogatives and relative clauses were included in the quantitative data presented in this thesis. A couple of particular expressions were also excluded, these will be mentioned in the relevant chapters.

The *string type* is an important category, used among other things to calculate the rate of inversion in the texts. It is established by representing the finite verb with the letter ‘V’, nominal subjects with ‘S’, pronominal subjects with ‘Sp’, and any other single constituent with the letter ‘C’. In order to reduce the amount of string types, a symbol ‘X’ was added to represent a (potentially empty) string of constituents other than the finite verb and the subject, such that every string type includes an ‘X’. For Old French, ‘X’ can only be the last symbol of the string and is used when both the verb and the subject (if overt) have made their appearance: a subject-initial string with a nominal subject is accordingly *SVX*, an inverted linear V2 string with a null-subject is *CVX*, a string consisting of the subject, then a constituent plus the verb is *SCVX*.

For Latin, the ‘X’ can also appear in the beginning or in the middle of the string type. In the beginning, it is only used for the string type *XV*, which means a null-subject, verb-final string. In the middle, it is only used for the string type *SXV*, which is the same string, but with an overt initial subject. These two string types are used to track the amount of verb-finality in the Latin data, a precaution which is not necessary for Old French, since

¹⁰³ Absolutive clauses correspond to the ‘Ablative Absolute’ construction.

verb-final strings in Old French can never represent a head-final VP/IP. For Latin, these issues remain open and it is completely unwarranted to make the same assumptions as for Old French; in other words, the *structural ambiguity* of the Latin data makes it necessary to keep track of verb-finality as a separate category.

2.7.2 The *bona fide* principle

When annotating linear word order, one often encounters many practical problems that must be solved, preferably according to some consistent principle. One such problem is to decide what counts as a single constituent. In many cases, a given sequence of words can be annotated as one, two, and sometimes even more constituents. In the contexts of the current investigation, which focusses on word order, this is not so important if the constituents follow the verb. If they on the other hand precede it, the choice has clear quantitative consequences for the linear word order facts. Since the procedure adopted should be consistent, this is not just a trivial detail: a consistently ‘inclusive’ bracketing (count as few constituents as possible) may give quite different results than a consistently ‘separative’ bracketing (count as many constituents as possible). Other problems are related to *inheritance* or the scope of coordination between conjunct clauses, how to treat clitics and clitic-like elements, how to decide if something is a parenthetical or not, how to treat stacking of constituents in front of the verb, complex predicates, discontinuous structures and in particular notoriously problematic constructions like the Latin AcI, which sometimes acts like a cohesive syntactic unit, sometimes like a loose association of constituents scattered all around its selecting verb, how to treat interrogatives and relative clauses (which had to be treated differently in Old French and in Latin), and many other details. I will not go through the technical choices made in all these cases, as this will be documented and made available in a user manual that goes along with the data files in the TROLLING repository. Some particularly relevant points will also be addressed in footnotes at different moments in later chapters.

Here, I will rather explain the logic of the annotation. The corpus was annotated according to a principle which I have dubbed the *bona fide* principle, and which basically consists in *separating annotation from analysis* as much as possible. Concretely, this means that, whenever a situation arose where several choices were possible, that choice was made which makes the least assumptions, and a tag was added to signal that the clause in question contains such a problematic case. Let me give a couple of examples. In (93), the finite verb is preceded by three heavy clausal constituents: two ablative absolute constructions followed by an embedded adverbial clause (*iam ut exiremus...*).

- (93) *[Lecto ergo ipso loco omni de libro*
 Read-*PST-PTCP-ABL* thus same passage-*ABL* all from book-*ABL*
Moysi et facta oblatione ordine suo], [hac
 Moses-*GEN* ad made-*PST-PTCP-ABL* oblation-*ABL* order-*ABL* REFL there
sic communicantibus nobis], [iam ut exiremus
 thus communicate-*PRS-PTCP-ABL* us-*ABL* now as go.out-*IPFV-SBJV-1PL*
de aeclesia], dederunt nobis presbyteri loci ipsius
 from church-*ABL* give-*PRF-3PL* us-*DAT* priests-*NOM* place-*GEN* same
eulogias... (Egeria 3.6)
eulogiae-ACC

‘Having read that entire passage from the book of Moses and made oblation as customary, then communicating there, just as were about to leave the church, the

priests of the place gave us *eulogiae...*’

One could argue that this is a case of ‘stacking’ and that what really precedes the verb is a single slot for a complex temporal adverbial expression, and that the clause should therefore be annotated as a V2 clause in linear terms. However, this is far from beyond doubt. The clause is therefore annotated as a linear V4 clause, and a tag – ‘stacking’ – is added to a separate column. This makes it easy for future users of the dataset to make up their own mind on the matter, since it is not too complicated to remove certain sequences of heavy constituents and check the resulting impact on the quantitative data. If the clause had been annotated directly as linear V2, this would have incorporated an analysis which is far from beyond doubt directly into the quantitative data, and it would not be possible to undo the effects of this assumption on the figures without manually going over the entire annotation, a very time-consuming process.

Another example: in (94), there are no less than five conjoined main clauses (two of them with asyndetic/covert coordination). Since all but the very first have their own subject, there is no reason to exclude them, as might be the case in many instances of coordination. Yet the first main clause is preceded by an ablative absolute which functions as a temporal adverbial expression, plausibly a kind of scene-setting element. The question is if this element is shared or inherited by all of the other conjuncts. In terms of interpretation, this is quite plausible, since the temporal adverbial expression seems to scope over all conjuncts. However, one cannot really tell for sure, and in either it is not entirely clear that this would mean that the first constituent is *syntactically* shared. Accordingly, the first clause is annotated as linear V2, and the four following clauses as linear V1, and a tag – ‘inheritance’ – is added to a separate column:

- (94) [*Facta ergo missa Martyrii*] *uenitur* *post Crucem*,
 made ergo mass-NOM Martyrium-GEN come-PASS-3SG past cross-ACC
dicitur *ibi* *unus ymnus tantum, fit* *oratio* *et*
 say-PASS-3SG there one hymn-NOM only happen-3SG prayer-NOM and
offeret *episcopus* *ibi oblationem et communicant omnes.*
 offer-3SG bishop-NOM there oblation-ACC and communicate-3PL all-NOM
 (*Egeria*, 35.2)

‘After the dismissal at the Martyrium, one goes past the Cross, a single hymn is said, a prayer is made, and the bishop offers the oblation there...’

Many other examples could be given. Most of them come from the Latin texts, as the syntax of Old French is much clearer and less ambiguous (although there are some complex cases there too). However, I consider that the general procedure is sufficiently illustrated. Needless to say, many choices simply impose themselves during annotation, and it is not advisable to forego reasoning in favour of any kind of slavish principle. On the whole, however, I consider the *bona fide principle* a well-founded procedure that allows for a clear separation of data collection and analysis.

Chapter 3

Old French: the main clause

3.1 Introduction

In this chapter and the following, a detailed quantitative and qualitative analysis of two Old French texts will be undertaken with a view to uncovering the phrase-structural organization that generates the different surface word order patterns. The data comes from a manual annotation of two prose texts from the first half of the thirteenth century, although I will also to some extent rely on evidence adduced by other researchers. I will be primarily concerned with the leftmost or structurally highest portions of the clause, with a particular emphasis on the question whether and to what extent Old French had developed a verb-second syntax featuring V-to-C movement. The current chapter takes us through the syntax of main clauses, while chapter 4 is devoted to embedded clauses.

The remainder of the introduction provides some general comments on Old French and how to delimit the object of study properly (section 3.1.1), and also introduces our corpus texts (section 3.1.2). Sections 3.2 to 3.7 deal with various aspects of the syntax of main clauses; sections 3.2–3.5 are devoted to an analysis of linear V2 strings, while linear V1 and V3 strings are the object of sections 3.6 and 3.7 respectively. Section 3.8 addresses some additional issues that were left unanswered in the previous sections, summarises and seeks to unite all of the findings into a formal analysis. This analysis must however remain *underdetermined* until the embedded data from chapter 4 has been presented; only then will we be able to provide a more complete picture of the nature of Old French syntax.

The two following chapters exclusively focus on late Old French of the 13th century as a synchronic system without consideration of time; for a sketch of the diachrony of Old French and Old Romance in general, see chapter 6.

3.1.1 Old French and verb-second; some preliminary remarks

‘Old French’ is the name traditionally given to the French language from its first written manifestation in the Oaths of Strasbourg (842) to around the mid-fourteenth century, after which period the same language is referred to as ‘Middle French.’ While the former date is justified by the simple fact that *les Serments* are the first French, or indeed first Romance, text written in the new vernacular scripta¹ the latter date is clearly more conventional.

¹Quite likely the new vernacular scripta arose at least partially because of the growing chasm in Carolingian France between reformed ecclesiastic Latin and spoken language and the concomitant desire to give

Periodization in diachrony always contains some degree of arbitrariness, and labelling, apart from being anachronistic, generally carries with it a notion of reification which is unjustified on purely linguistic grounds. The case of French is no exception, as the transition from ‘Old’ to ‘Middle’ French does not imply any profound rupture or discontinuity in the evolution of the language. For instance, the inversion structures which are the central focus in this chapter do not vanish abruptly around 1350, but rather shows a gradual decline over the century and the following, such that it is impossible to state any date or even short interval where G-inversion ‘was lost’. For a discussion of the problematic nature of labelling and periodization in Romance diachrony, the reader is invited to consult the many relevant contributions in Wright (1991) and also the chapter by Wright in the more recent *Cambridge History of the Romance Languages* (Wright 2013).

In spite of these considerations, I have chosen to retain the traditional term Old French (abbreviated OF) with a capital letter in the epithet ‘Old’. The reason for this lies in the fact that the Old French prose texts of the thirteenth and fourteenth centuries – barring some diatopic morpho-syntactic variation which is of little concern to us – show a considerable degree of internal cohesion in most aspects of grammar but orthography. As we shall see later in this chapter, this applies in particular to word order and the position of the verb, which is the focus of the current investigation. The reader should bear in mind, however, that the term ‘Old French’ as employed in this and the following chapter is strictly speaking only referring to the aforementioned prose texts of the thirteenth and fourteenth century, not to the entire period between 842-1350. When used in this restricted sense, it is possible to make interesting generalizations about the French language of the period.

3.1.1.1 Old French and the V2 controversy

For several reasons, Old French enjoys something of a privileged position in the debate on Old Romance verb-second. Firstly, awareness of the quite consistent second position of the verb in Old French dates back at least to the late 19th century. The Swiss philologist Rudolf Thurneysen is generally accredited with the discovery (Thurneysen 1892).² Perhaps not entirely correctly; although Thurneysen was presumably the first to elaborate on Old French V2 in a paper, the frequent subject-verb inversions had not escaped earlier philologists, as the following observation by Le Coultre demonstrates:

‘Does it follow from what we have just stated that the verb must necessarily occupy the second position in the clause like in German, and that wherever the subject is not at the head of the clause, it must be after the verb? The rule is not absolute, but takes place approximately in the proportion of 80%’ (Le Coultre 1875:17)³

Secondly, it is a recurrent claim that Old French verb-second was somehow ‘stricter’ than that of the other Romance languages, allowing less exceptions at the surface level of linear word order (Benincà 1983, Vanelli et al. 1985:167, Benincà 2006, Vance et al. 2009, Wolfe

the latter a written form of its own (Banniard 1992; Wright 1982). For a recent critique of the ‘logographic’ theories of Wright and Banniard and a very different view on the role of the Carolingian reforms, see also Varvaro (2013).

²It is less often observed that Thurneysen himself more than hinted at the pan-Romance character of verb-second (Thurneysen 1892:302). See also Diez (1877:463).

³‘Résulte-t-il de ce que nous venons de dire que le verbe doit nécessairement occuper la seconde place dans la phrase comme en allemand, et que partout où le sujet n’est pas en tête, il doit se trouver après le verbe ? La règle n’est pas absolue, mais elle se réalise environ dans la proportion de 80%’ (Le Coultre 1875:17)

2015b). It might partially be due to this privileged position that the research literature on Old French V2 is relatively large compared to that on the other Romance varieties.

As was noted in the introduction, there is also another reason why Old French is special with regard to verb-second. The Germanic superstrate brought about by the Frankish conquest of Gaul in the fifth century and the subsequent bilingual character of the (early) Merovingian state is an undeniable historical fact, although it is impossible to assess with certainty the diatopic and diastratic details of this bilingualism (see section 6.3.1 for a short discussion). Nonetheless, this state of affairs means that the idea of verb-second as a result of language contact and bilingualism holds a great deal of initial plausibility. This hypothesis has a rather long tradition in French philology (Meillet 1931:37; von Wartburg 1958:107) and is regularly reasserted (Thomason and Kaufman 1988; Posner 1996:53), most recently by Mathieu (2009) and Hänsch (2014). We will not directly assess that hypothesis in this chapter, but rather defer it to the discussion in chapter 6, where we can benefit from the hindsight of previous chapters. In this chapter, we will be concerned with a more fundamental question, namely whether Old French was ever truly a V2 language in the medieval period.⁴

Many researchers have answered this question in the positive, (Benincà 1983, 1995; Vanelli et al. 1985; Adams 1987a,b, 1989; Roberts 1993; Vance 1997; Labelle 2007; Salvesen 2013; Steiner 2014; Wolfe 2015b)⁵ but at the same time, there also exists a non-negligible countercurrent of researchers who call into question the validity of the V2 analysis for Old French, claiming the similarities between Old French and Modern Germanic are superficial and not reflective of the same underlying structure. These researchers have on the whole focused more on demonstrating that Old French and modern Germanic – generally exemplified by German with some ancillary arguments taken from Icelandic – were different than actually developing explicit models of the syntax of Old French. Still, they uniformly reject the V2 status for Old French (Kaiser 2002; Ferraresi and Goldbach 2002; Rinke 2003; Kaiser 2009; Rinke and Meisel 2009; Rinke and Elsig 2010; Kaiser and Zimmermann 2011; Elsig 2012; (see also Becker 2005 for a more cautious conclusion), generally by rejecting a V-to-C analysis in favour of an analysis with the verb in I^0/T^0 . We will return to some of the models that have been proposed; let it suffice here to say that this dissension is in itself a justification for more research on the syntax of Old French and to some extent provides the *raison d'être* of the current chapter.

Although the question if Old French should be considered a V2 language is not without interest, it was argued in chapter 2 that the notion of a V2 language is not entirely stringent from a theoretical point of view. The more fundamental issue in this chapter is therefore the question of what kind of syntactic model most appropriately captures the observed word order patterns. Recent theoretical developments couched within a cartographic model of the left periphery have spawned analyses that make novel and interesting claims about the locus of verb movement and the nature of root-embedded asymmetries (Poletto 2002; Wolfe 2015b). We will engage with this recent literature and explore whether the patterns observed favour a traditional analysis with a unitary CP-projection, or whether adopting an articulated CP-layer consisting of several projections might yield a better insight into the

⁴I will not be concerned with the loss of V2 in a diachronic perspective since it is not the focus of this dissertation; see Adams 1987a, 1989; Kroch 1989; Roberts 1993; Côté 1995; Platzack 1995; Vance 1997. For a general discussion, see also Kaiser 2002.

⁵We could also include non-generative work such as Foulet 1930, Lerch 1934 and Skårup 1975 here, but since the controversy around verb-second in Old French revolves around the structural analysis of the phenomenon, that is V-to-C movement, rather than the linearization, it would not be entirely fair to rally these researchers to either side of the debate.

nature of Old French syntax.

3.1.2 The texts

Unlike the case with texts written in Latin, which will be discussed and problematized in chapter 5, we do not have strong reasons to distrust the written testimony of Old French prose texts of the 13th century beyond the caveats which always apply when using written texts to draw conclusions about spoken language. Naturally, in the case of OF as well, one must avoid the temptation of seeing prose as a simple codification of spoken language; prose too is clearly constrained and shaped by stylistic factors, as has been demonstrated among others by Cerquiglini (1981). Still, it does not seem unreasonable to suppose that prose texts are *constrained* by the syntax of the spoken language, such that the artistic aspirations of the author must unfold within the limits imposed by syntax. This is at least largely the case for prose literature today, and in particular as regards word order.⁶ I believe this gives us reason for optimism on the part of our corpus.

The same can not be said of verse, and we therefore take care to avoid texts written in verse. Already Thurneysen pointed out (Thurneysen 1892:296) that the word order of OF verse differs heavily from prose and is not suitable as evidence of the spoken language (see also de Kok 1985:4).⁷ A priori, a similar problem is attached to translations, as the word order of the source text may exert influence upon the translation. However, in cases where both the source text and the translation are available for comparison, this potentially distorting effect may be controlled for. Again, comparison with the historical prose corpus as a whole should also reveal whether the word order of a given translation is idiosyncratic or consistent with the evidence from other texts. In the case of *La Vie de Saint Eustace* the translation is stylistically very free and no attempt whatsoever has been made to follow the original, doubtlessly because the Latin word order of the original was quite simply ungrammatical in thirteenth-century French. In a more general vein, the word order of translations from Latin might be expected to deviate in some instances from a native verb-second syntax in favour of the word order of the source text, but when no such deviation is found, we may reasonably safely surmise that we are dealing with vernacular word order.

3.1.2.1 Le Roman de Tristan en Prose

The legend of the passionate and adulterous love between Tristan, nephew of King Marc of Cornwall, and Iseult, the wife of the latter, enjoyed great popularity from the High Middle Ages and well into the Renaissance (Radwan 2011:28). The fact that the Prose Tristan is transmitted in no less than 82 manuscripts or fragments of manuscripts (Ménard 1987:8) gives witness to this popularity. At the same time, this plethora of manuscripts, combined with the enormous proportions of the work (some of the complete manuscripts contain around 500 folia) had for a long time the effect of dissuading philologists from the task of editing the work. To this day no single edition of the whole text exists. It has been recognized since Löseth (1891/1970) that the manuscripts contain at least two different versions of the text, both of which contain elaborations and episodes not found in the other.

⁶Of course, morpho-syntax can be heavily influenced by a conservative normative standard, as is the case in modern French prose. Word order is not left entirely unaffected, since there is for instance a clear tendency to use inversion more frequently than in spoken varieties, but on the whole, prose does not concoct its own word order or fall back on word order patterns that fell completely out of use centuries ago.

⁷In chapter 6, section 6.3.1, it will be suggested that the almost complete lack of prose texts before the 13th century complicates our understanding of the earlier phases of Old French.

For our corpus, Curtis' (1963) edition was chosen. This edition is based on the ms. Carpentras 404, dated to the second half of the 13th century. This choice of manuscript is called into question by Ménard (Ménard 1987:21-24), but since his concerns primarily regard the latter part of the manuscript, published in tome III (1985), this criticism need not detain us; of far greater importance is the fact that Curtis' edition is practically available in electronic form from the PROIEL corpus (Haug and Jøhndal 2008). The first 2000 clauses of tome I, main and embedded combined, were extracted, thereby eliminating the danger inherent in manual transmission. These clauses were annotated manually according to the principles laid out in section 2.7.

The text dates from the first half of the 13th century, presumably as early as 1200-1230. The Carpentras ms. contain regional features which suggest an origin in the south-eastern reaches of the *langue d'oïl* area, but the editor concludes that, on the whole, the language of the scribe is based on the Francien dialect (Curtis 1963:24). For a discussion of the multiple authorship and the relationship between the authors and the different versions of the text, see Curtis (1983).⁸

3.1.2.2 La vie de Saint Eustace

La vie de Saint Eustace tells the legend of the Roman general Placidus, who takes the name Eustace after his baptism in the the new Christian religion. The story centers on the many hardships endured by Eustace after his conversion, as his faith is put on test by God. Losing home, property, position and family, Eustace never renounces his Christian faith. When finally reunited with his long-lost wife and sons and promised full restitution by emperor Hadrian, Eustace and his family refuse to revert to the old gods and joyfully choose to die a gruesome death as martyrs at the hands of the emperor.

The legend was widely transmitted in the Middle Ages, reaching a zenith of popularity in France, where many versions were composed in verse and prose. The text of our corpus represents the oldest surviving prose adaptation and is a translation from a Latin original which also survives. The edition is that of Murray (1929), which is based on the manuscript 2464 of the *Bibliothèque nationale de Paris*, dated to the first half of the 13th century. According to the editor, the ms. does not present clear regional features and can be loosely attributed to the central region of France. Below the OF text, the editor publishes the Latin original, based on the oldest known manuscript, the ms. 5577 of the *Bibliothèque nationale*. The translation is faithful to the original in terms of content, but generally free in terms of language.

The text was extracted from the corpus *Base de Français Médiéval 2016* (Barbance-Guillot et al. 2017) and annotated manually according to the principles laid out in section 2.7. The text comprises 888 clauses, main and embedded included.

⁸A prologue contained in many manuscripts identifies a first author, Luce of 'Castle Gat' in the region of Salisbury in England. Neither his name nor that of this castle are documented elsewhere, and modern scholarship has found reason to suspect an invented identity here. His claim to have translated the story from a Latin original into French is also called into question. Other prologues and several epilogues refer to a second author, Hélié de Boron, whose alleged kinship with Robert de Boron is equally debatable (Curtis 1983).

3.2 The corpus data: Linear V2

In the rest of this chapter, I will present the corpus data and submit these to a detailed analysis with a particular focus on the structural position of the finite verb. We start out with quantitative, surface-oriented evidence and gradually progress towards the more fine-grained quantitative and qualitative evidence that ultimately provides the more reliable diagnostic of underlying syntactic structure.

I have chosen to present the two texts together, rather than devoting a section to each. Apart from purely practical considerations, the reason for this choice is twofold. First, the two prose texts are composed roughly at the same period, namely the first half of the 13th century. Secondly, they show a very high degree of consistency in virtually all aspects of grammar and in particular with respect to word order. We may therefore consider them jointly as synchronic evidence of the state of Old French syntax in the first half of the thirteenth century. The quantitative evidence for each text is of course kept apart and presented in separate tables or columns, while the qualitative evidence is presented with a mind to show the same phenomenon from both texts. The minor differences which exist will be addressed as we go along.

3.2.1 The linear distribution of the finite verb

We start out by considering some quantitative data, starting with the distribution of the finite verb in terms of linear order. This information is contained in tables (3.1) and (3.2). Notice that the two texts are very similar in virtually every aspect; four different linear positions of the verb are attested in both texts, and their relative distribution is also extremely similar. V4 orders are virtually non-existent. It is also worth noticing that there is no clearly discernable effect of the predicate class variable on the linear distribution of the verb. In particular, linear V2 is almost exactly as frequent with transitive verbs as with unaccusative verbs.

Table 3.1: Linear order of the finite verb in main clauses in *Tristan*

	Transitive	Unaccusative	Copula	Functional	Total
V1	44 (7.68%)	30 (12.20%)	12 (10.08%)	1 (1.67%)	87 (8.72%)
V2	431 (75.22%)	192 (78.05%)	96 (80.67%)	46 (76.67%)	765 (76.65%)
V3	96 (16.75%)	23 (9.35%)	11 (9.24%)	12 (20.00%)	142 (14.23%)
V4	2 (0.35%)	1 (0.41%)	– (0.00%)	1 (1.67%)	4 (0.40%)
Total	573 (100.00%)	246 (100.00%)	119 (100.00%)	60 (100.00%)	998 (100.00%)

Average number of constituents $\approx 3,46$

Null-subjects: $326/998 = 32.67\%$

Table 3.2: Linear order of the finite verb in main clauses in *Eustace*

	Transitive	Unaccusative	Copula	Functional	Total
V1	19 (7.01%)	9 (6.82%)	1 (1.89%)	3 (14.29%)	32 (6.71%)
V2	210 (77.49%)	104 (78.79%)	48 (90.57%)	14 (66.67%)	376 (78.83%)
V3	42 (15.50%)	17 (12.88%)	4 (7.55%)	4 (19.05%)	67 (14.05%)
V4	– (0.00%)	2 (1.52%)	– (0.00%)	– (0.00%)	2 (0.42%)
Total	271 (100.00%)	132 (100.00%)	53 (100.00%)	21 (100.00%)	477 (100.00%)

Average number of constituents $\approx 3,05$

Null-subjects: $154/477 = 32.29\%$

The tables also resoundingly show that the second position is the dominant one, reaching almost 80%. This is a very high figure, for instance higher than the 73.6% linear V2 in main clauses in the Old High German *Isidor* as reported in Lippert (1974), and considerably higher than the 68% reported from Walkden’s annotation of the Old Saxon *Heliand* (Walkden 2014), two Germanic languages which are considered verb-second languages. The amount of linear V2 in main clauses is therefore well within the output ranges expected for a verb-second grammar. In fact, this number, which was already evoked by Le Coultre (1875; see section 3.1.1.1) is surprisingly consistent across many investigations into Old French syntax (Roberts 1993; Vance 1997; Radwan 2011; Wolfe 2015b). We may therefore already draw a firm first conclusion:

Conclusion I:

The finite verb in late Old French regularly occupied the linear second position in main clauses.

There is nothing new to this observation, and it is clear that the numerical strength of linear V2 is not enough to conclude that we are dealing with structural verb-second syntax here. There is also a considerable amount of V3 orders, and even a non-negligible amount of verb-initial sentences. We will return to these word orders, but for the moment we leave them aside and consider the linear V2 orders in more detail, starting with an examination of the properties of the prefield.

3.3 The prefield in Old French

It is important to consider carefully how the prefield, in other words the position to the left of the finite verb, functions in Old French, as it is widely assumed in formal syntax that there exist substantive differences between V2 languages and non-V2 languages in this domain of the grammar. In non-V2 languages such as modern English or French, special constructions aside,⁹ the prefield in declarative clauses is an A(argument) position specialized

⁹Such special constructions include for instance ‘residual V2 structures’ in English or French. While these constructions are mainly restricted to wh-questions, they are also found in declarative clauses in the case of ‘Negative Inversion’ in English or (optionally) in conjunction with certain adverbs in French. If we accept that these are indeed remnants of V2 syntax (but see Kiparsky 1995), their presence in the grammar

for hosting the subject of the clause. In verb-second languages, as we saw in chapter 2, there is no such restriction and the prefield can host a range of different constituents.

The corpus supports the evidence adduced by much previous research in showing that Old French patterns like a verb-second language in this respect (Vanelli et al. 1985; Adams 1987a; Roberts 1993; Vance 1997; Benincà 2006; Wolfe 2015b). Thus, the initial constituent can not only be a subject, as in (95), but also a direct object (96), an oblique or prepositional object (97), a predicative complement (98), an adverb or a PP functioning as an adverbial (99) or even a non-finite verb (100).¹⁰ Notice that oblique pronouns, the pronominal adverbials *y/en* and the preverbal negative morpheme *ne/n'* are clitics on the verb and do not count for linearization purposes:

- (95) a. *[Bron] vint a Joseph et li dist...*
 Bron came to Joseph and him.CL said ...
 ‘Bron came to Joseph and told him...’ (Tristan, p.40 : 2.5)
- b. *[Eustaces] li respondi...*
 Eustace him.CL answered...
 ‘Eustace replied to him...’ (Eustace, p.13 : X.4-5)
- (96) a. *[Tel don] te fais je, biaux amis.*
 Such gift.ACC you.CL make I good friend.
 ‘Such a gift I give to you, my good friend.’ (Tristan, p.40 : 2.23)
- b. *Car [cest don] li dona Nostre Sires...*
 For this gift.ACC him.CL gave Our Lord
 ‘For Our Lord gave him this gift...’ (Eustace, p.45 : XXXIX.7-8)
- (97) a. *Rois, [de cest songe qui t’ est avenu] te dirai ge ce que je en cuit.*
 King, of this dream which you.CL is come you.CL say-FUT I that which I thereof.CL think
 ‘My King, I shall tell you what I think of this dream that came to you.’
 (Tristan, p.47 : 22.2-3)
- b. *Je aor le mien Seignor, Jhesu Crist: [a lui] faz je sacrefices e oroisons...*
 I worship the my Lord, Jhesus Christ: to him make I sacrifices and prayers...

of these languages does not invalidate the claim that the prefields of V2 and non-V2 languages function differently.

¹⁰This latter example is probably best analysed as fronting of the entire VP, as participles qua heads cannot occupy phrasal positions on standard assumptions. This does not mean that the first constituent of a V2 construction must necessarily be a maximal projection; an apparent counter-example is provided by ‘Long Head Movement’ in Breton, which has been analyzed as involving movement of a head to the left periphery in fulfilment of a V2 constraint (Borsley et al. 1996) and apparently in violation of the Head Movement Constraint (Travis 1984).

- ‘I worship my lord, Jesus Christ: to him I sacrifice and pray...’ (Eustace, p.39 : XXXV.7-8)
- (98) a. ... *et [compaignon d’armes] avoient il esté.*
 ... and companions of arms had they been.
 ‘... and they had been brothers in arms.’ (Tristan, p.52 : 33.12-13.)
- b. [*Granz e puissanz] est li dex as crestiens...*
 great and powerful is the god to-the christians
 ‘Great and powerful is the god of the Christians...’ (Eustace, p.44 : XXXVIII.12-13)
- (99) a. *Et [lors] ving je jusqu’ a ceste fontaine ...*
 and then came I all-the-way to this well
 ‘And then I came to this well ...’ (Tristan, p.52 : 35.12-13.)
- b. [*Aprés ce] repaira Eustaces a son ostel...*
 after this returned Eustace to his domicile
 ‘Afterwards Eustace went back home...’ (Eustace, p.13 : XI.1)
- (100) a. [*Honi] m’ a mes freres ...*
 dishonoured me.CL has my brother
 ‘My brother has brought dishonour upon me.’
 (Tristan, p.43 : 12.3)
- b. [*Oï] les avoit il sanz faille ...*
 Hear them.CL had he without failure ...
 ‘He had doubtlessly heard them ...’ (Tristan, p.65 : 68.2)

Notice also that Old French patterns like Scandinavian in allowing simple negation to appear in the prefield (101), where it assumes a tonic form *non* as opposed to the normal clitic negator *ne*.¹¹ Furthermore, it seems like discourse particles are also permitted in clause-initial position (102), suggesting that the Old French prefield is in some respects even more permissive than Germanic V2 languages in terms of the categorial status of its occupants. On the other hand, verbal particles are not encountered in the prefield in our corpus, but this could be an accidental gap due to the fact that these are generally very rare in Old French:

- (101) (Context: A strange man comes and sits down next to Tristan):
 ... *mes mot ne li dist, et [non] fist Sador a li.*
 but word NEG.CL him.CL said, and not did Sador to him

¹¹Just like in Scandinavian, negation in the prefield is closely associated with contrast; sometimes it bestows a contrastive reading on the subject, like in (101), while in other cases the contrast involved amounts to rejecting a previous discourse move, cf. Foulet (1930, pp.236–237)

- (i) *Ha! sire, fait la roïne, lessiez le moi, s’ il vos plest. Dame, fait il, [non] ferai*
 ha sir makes the queen let him.CL me if it you.CL pleases lady makes he not will-do
 ‘Ha! Sire, says the lady, please give it to me. My lady, he says, I will not ...’ (Tristan, p.54 : 39.4-5)

‘... but he did not speak a word to him, nor did Sador to him.’ (*Tristan* p.64, 67. 7-8)

- (102) *Coment puet il vivre? fait li rois; [ja] fu il gitez en la mer!*
how can he live makes the king PRT was he thrown en the sea
‘How can he be alive?’ the king asked, - ‘He was cast into the sea!’ (*Tristan* p.47, 22.9)

These examples serve to demonstrate that the prefield is not a position reserved for subjects, but rather an A’ position that does not impose any requirement on the category or grammatical function of the constituent it hosts. We may therefore conclude that the prefield behaves in a way similar to that of the Germanic V2 languages. It is important to emphasize, however, that the term ‘prefield’ is used here in a surface-oriented sense to refer to anything to the left of the finite verb, and does not say anything about the structural position of either the verb or the initial constituent. The evidence considered so far is therefore far from decisive in distinguishing between a V2 and a non-V2 grammar from a theoretical point of view, but it does allow us to draw another firm conclusion:

Conclusion II:

The prefield in late Old French was not reserved for the subject, but functioned as an A’ position hosting phrases with different categorial status and different grammatical functions.

This furthermore entails that subject-verb inversion is found in Old French declarative clauses in contexts where the modern language does not allow such inversion. None of the examples in (96–102) with the exception of (98b) are grammatical in modern French without changing the word order so that the subject precedes the verb. Notice also that argument fronting to the prefield, as in (96b), is not accompanied by clitic-doubling inside the clause (see also Roberts 1993:108, Vance 1997:234, Salvesen2013), as is virtually always the case in modern French (Rowlett 2007:178–180, De Cat 2009:98) These facts show that the Old French language behaved rather like the modern Germanic languages with respect to argument fronting, and that the language has subsequently undergone a significant change in some core property of syntax. To say that French has lost inversion is descriptively correct, but we would like to be able to say something more concrete about the structural underpinnings of these inversion structures, so as to better understand exactly what has been lost.

Having established that the prefield is qualitatively available to all types of constituents, we will next consider the actual quantitative distribution of different XPs in the prefield in linear V2 clauses. This is important, because we need to know if inversion in Old French was a rather marginal phenomenon, or if it is was in fact an option which was substantively used in the language.

3.3.1 The prefield in quantitative terms

We will now consider the distribution of different constituents in the prefield of linear V2 strings. This information is provided in table 3.3. Notice again the striking similarity between the two texts with respect to all constituents apart from non-finite verbs, which (presumably incidently) are lacking from *Eustace*. These numbers further underscore what the examples (95–102) above showed, namely that the prefield in OF was not in any sense a subject position, whether qualitatively or quantitatively speaking.

Table 3.3: *Tristan and Eustace*: XPs in the prefield of linear V2 strings in main clauses

Initial XP	Tristan	Eustace
DP subject	167 (21.83%)	90 (23.94%)
Pron. subject	205 (26.80%)	92 (24.47%)
Direct object	30 (3.92%)	18 (4.79%)
Predicative	11 (1.44%)	8 (2.13%)
Oblique object	9 (1.17%)	8 (2.13%)
Infinitive	6 (0.78%)	– (0.00%)
Participle	5 (0.65%)	– (0.00%)
Negation	2 (0.26%)	– (0.00%)
Adverbial	330 (43.14%)	160 (42.55%)
Subject-initial	372 (48.63%)	182 (48.40%)
Non-subject-initial	393 (51.37%)	194 (51.60%)
Total	765 (100.00%)	376 (100,00%)

It is also highly significant that the prefield in V2 strings is divided almost equally between subjects and non-subjects, with the latter even marginally outscoring the former. This is an important finding and must be assumed to be very salient from an acquisitional perspective. However, there is once again nothing atypical about it when we compare it to other findings in the literature, as table 3.4 below shows.¹² While these figures reveal that some variation is indeed observable in the historical corpus, presumably related to differences in style and sub-genre, the numbers are consistently high, and much higher than the 12.6% reported in (Kaiser 2002). It should be noted that Kaiser’s findings were based on *Les quatre livres des rois*, a late 12th century, partially rimed Anglo-Saxon translation of a Latin original. The text is known to differ in important respects from the historical corpus as a whole. Ingham points out that the text is unique among 12th century prose texts in allowing null-subjects in negative subordinate clauses (Ingham 2014:36), and a similar observation is made by Dupuis (1988). Zaring observes that *LQLR* patterns more like 12th century verse than 13th century prose with respect to embedded nominal inversion, making her speculate that the text might be ‘unusual prose’ in this respect (Zaring 2017:304).

As table 3.4 shows, the most frequent non-subject filler of the prefield by far is adverbial phrases, just like in modern Germanic V2 languages. Argument fronting also quite significant, reaching more than 5% in *Tristan* and 7% in *Eustace* if we combine direct objects with oblique arguments (all PP arguments of verbs, including indirect objects). This is comparable to the figures for the German and Icelandic (but not Old French) translations of *Les quatres livres des rois* in Kaiser (2002:141), and is also very much in line with the

¹²Sitaridou analyses several Old French texts, and the table from which the percentage in table 3.4 is collected represents the total (Sitaridou 2012:569) without differentiating between the different texts.

Table 3.4: Non-subject-initial linear V2 in main clauses in the literature on Old French

Text sample & Reference	Non-subject initial V2
<i>Villehardouin</i> (Rinke and Meisel 2009)	57,3%
<i>Sept Sages</i> (Rinke and Meisel 2009)	35,7%
<i>Clari, Sept Sages, Méd. liégeois</i> (Sitaridou 2012)	49,5%
<i>La Queste</i> (Wolfe 2015b)	53,7%

figures adduced for both modern and historical Germanic V2 languages in chapter 2 (Westman 1974; Fabricius-Hansen and Solfeld 1994; Bohnacker and Rosén 2008; Walkden and Booth to appear; see section 2.2). On the whole, we may safely conclude that the prefield in Old French shows striking similarities with that of modern Germanic V2 languages, both in being qualitatively accessible to a wide variety of constituents and in terms of the actual quantitative distribution of the elements found there.

In spite of this, it has been claimed that there is an important difference between the prefield in Old French and modern Germanic that reflects that the former language was not a true V2 language. This argument is developed by Rinke and Meisel (2009) (hereafter R&M). The core of R&M's claim is that Old French is a pro-drop language in the sense established by Rizzi (Rizzi 1982) and since elaborated by many others. The idea is that rich agreement morphology is pronominal in the sense that it can check the EPP-feature on T^0 (see also Barbosa 1995; Alexiadou and Anagnostopoulou 1998; Kato 1999). Due to economy considerations, the nominal subject stays low, '... in its post-verbal base position' (Rinke and Meisel 2009:97), a position they identify as Spec-VP. This allows Spec-IP to take on the role of an A-bar position which can host different kinds of constituents. Thus, the word order variation in the Old French prefield is a result of the interplay between syntax and information structure. More precisely, the initial constituent is topical, representing generally old or familiar information, while the postverbal material is focal, new information. The nominal subject only moves to Spec-TP (Spec-IP in our terminology) as a 'repair strategy [...] to escape a focus interpretation' (Rinke and Meisel 2009:109). Pronominal subjects are clitics on the verb, which only moves as high as T^0 .

R&M's claim that word order variation in the OF prefield is a result of the interplay between syntax and information structures is naturally correct when 'word order' is understood in a surface sense; for instance if SVO and OVS are taken as two different 'word orders'. However, it is also correct for the modern Germanic languages and presumably most languages in the world. The idea that preverbal subjects in OF are generally topics has also been voiced earlier in the literature, (Vance 1997; Marchello-Nizia 1999; Prévost 2001), and as a general rule of thumb, this observation also seems to be correct. The problem is again that, more often than not, this also applies to the initial constituent in Germanic V2 languages, which also has a strong tendency to be topical, a fact explicitly recognized by Rinke and Meisel (2009:111). However, they argue that the difference between verb-second inversion on the one hand and inversion in null-subject languages like the modern Romance languages or Old French on the other hand, is that the former is not restricted to topicalisation, while that is in fact the case in the latter. Thus, the initial constituent in German 'can be a topic, information focus, contrastive focus or an adverb that is neither the topic

nor the focus of the sentences' (Rinke and Meisel 2009:111).

It is not immediately clear how the informational structural properties of the prefield would be relevant to the V2 or non-V2 status of a language. As already stated, the word order variation in main clauses is an interplay between syntax and information structure in V2 languages as well, in the sense that the initial constituent will be a topic, focus, scene setter etc. based on the possibilities afforded by the discursive context. The only invariant feature is that the verb occupies C^0 , independently of the information structural value of the initial constituent. There is every reason to assume that there exist differences between languages regarding the pragmatic properties of the prefield which do not cut neatly along the divide between V2 and non-V2 languages, and such differences have indeed been reported in the literature. For instance, modern German is more liberal than the Scandinavian languages in allowing information focus in the prefield (Bohnacker and Rosén 2008; Bohnacker 2010), whereas in the V2 language Kashmiri, the prefield generally hosts foci rather than topics (Holmberg 2015); this does not affect their status as V2 languages.

Still, in order to dismiss the claim that inversion in OF is narrowly conditioned by a particular pragmatic partitioning of the clause, we will now consider the information structure of the prefield, restricting our attention for the moment to linear V2 strings.

3.3.2 The information structure of the prefield

Rinke and Meisel (2009) argue that the prefield in Old French is the structural position Spec-TP and that it is specialized for hosting topical information. At the moment, we are not in a position to evaluate the structural position of the prefield, so we focus on the second claim, that the prefield in OF is specialized for hosting topics. This claim does not stand up to scrutiny, as we do not have to search far to find evidence of the great variation afforded by the prefield in terms of information structure.

- (103) [*Cil Bron*] *avoit de sa moillier doze fiuz, mout biaux enfanz et mout saige et mout preu; et [mout] amoient de grant amor lor pere. Et [lor wise and very valiant and much loved of big love their father and their mere] n' avoient il pas ...*
mother NEG.CL had they not

'And this Bron had twelve sons by his wife, very beautiful and wise and valiant children; and they loved their father deeply with great love. They had lost their mother ...' (Tristan, p. 40 : 2.2-4)

Within the space of three sentences, the flexibility of the prefield offers first an aboutness topic (Frascarelli and Hinterhölzl 2007:1), mentioned in the preceding context (*Cil Bron*), followed by an initial adverb (*mout*) implying a scalar contrast amenable to an analysis as a focus (Vanelli 1998:82; Ledgeway 2008:450), and then even a new-information or *presentational* focus introducing a previously unknown actor into the discourse (*lor mere*). Perhaps one could argue that the latter example, although clearly introducing a new discourse-referent, is licensed because there is some kind of anchorage with a previously known referent through the use of the possessive anaphor 'their'. This might be true, but it is important to emphasize that without such anchorage, and without referring to entities that are assumed to be known to the listener from before, new information focus is very awkward in the prefield in at least some Germanic languages as well, as was demonstrated in section 2.2.1

in chapter 2. The reason seems to be that *brand-new* information in the sense of Prince (1981), that is information that is neither present in the preceding discourse nor assumed to be part of the common stock of knowledge between the speaker and hearer, is preferably realized in postverbal position. This seems to be the the case in Old French as well, (Vance 1997:57, Steiner 2014:171-172, Wolfe 2015b:89-90) where new or inactive discourse referents are much more likely to be introduced directly into the prefield if they have some kind of anchorage, witness (104):

- (104) [*Naburzadan, li frere Sador,*] *la resgarda par tantes foiz que ...*
 Naburzadan, the brother Sador her looked-at by so-many times that...
 ‘Naburzadan, Sador’s brother, looked at her so many times that ...’
 (Tristan, p.42 : 7.7)

In spite of this, the corpus still offers several examples of what might reasonably be characterised as new information focus in the prefield, as the following examples illustrate (see also (131) above):

- (105) *Sador aporta la demoisele en ce chastel que je vos di, et la mist*
 Sador brought the lady to this castle that I you.CL said and her.CL put
en une chambre. Et fu cele leanz bien trois jorz enz que ele manjast,
 in a room. And was that-one there well tree days before that she ate
car [tel paor] avoit eü de la mer ...
 for such fear had had of the sea
 ‘Sador brought the lady to the aforementioned castle and put her in a room. And she stayed there tree full days before she ate, for such was the fear instilled in her by the ocean.’ (Tristan, p.41 : 6.1-3)
- (106) *Et la dame, qui sa feme estoit, estoit apelee Madule, mout saige dame et*
 and the lady who his woman was was called Madule very wise lady and
mout cortoise; et de [haut linaige] n’ estoit ele mie estraitte.
 very polite and of high lineage NEG.CL was she not drawn
 ‘And the lady who was his wife was called Madule, a very wise and courteous lady; but she was not of noble birth.’ (Tristan, p. 49 : 26.4-6)
- (107) *...mout avoient esté bon ami entre lui et le roi Canor, et*
 ...much had been good friends between him and the king Canor and
[compaignon d’ armes] avoient il esté.
 companion of arms had they been
 ‘They had been good friends, him and king Canor, and they had been brothers in arms.’ (Tristan, p.52 : 33.12-13)
- (108) *...qant il vint en mi le flueve, qui estoit granz e lez, [uns lions]*
 ... when he came in middle the river which was great and strong a lion
issi del bois, qui ravi l’ enfant ...
 came-out of-the forest which seized the child
 ‘... when he came to the middle of the river, which was big and strong, a lion came out of the woods and seized the child ...’
 (Eustace, p.17 : XIV.11-12)

As for contrastive focal readings, these are also found, although they are admittedly not very numerous. They often involve adverbs of degree which are commonly analysed as implying a scalar contrast. One example was cited above in (104), another is given in (109), while (110) involves a clear and implicit contrastive focus.

- (109) [*Grant*] **fu** la joie et la feste qu' il firent au roi Pelias ...
 big was the joy and the party that they made to-the king Pelias
 'Great was the joy and great the party they made for king Pelias...' (Tristan, p. 53: 36.13)
- (110) [*Mieuz*] **est** que nos doignons a son cors aucune sepulture, que les bestes
 better is that that we give to his body some burial that the
 le manjassent.
 beasts him eat
 'It is better to give his body a burial than to let the beasts devour him.' (Tristan, p. 42: 11-12)

These examples show that inversion in Old French can in principle also reflect a focus-background division of the clause as well as the more current topic-comment structure, conclusions already reached by Steiner (2014) and Labelle and Hirschbuhler (2018). In addition to topics and foci, one also finds various kinds of adverbials (111–112) including some that may plausibly be considered scene-setters (113–114) (cf. section 2.4.4.1), as well as expletives (115–116) in the prefield.

- (111) [*Adonc*] **conterent** li chevalier a Eustace le comendement l' emperere ...
 then told the knights to Eustache the command the emperor
 'Then the knights told Eustache about the emperor's command ...' (Eustace, p.28 : XXIV. 9-10)
- (112) [*Ensi*] **vint** li enfes a sauveté la ou li rois l' avoit abandoné a
 thus came the child to safety there where the king it.CL had abandoned to
 destruction.
 destruction
 'Therefore the child was brought to safety where the king had left it to perish.'
 (Tristan, p.49 : 26. 1)
- (113) [*Lendemain*] **firent** il encore greignor feste e greignor joie ...
 the-day-after made they even greater party and greater joy
 'The next day they made an even greater and more joyful party ...' (Eustace, p.37 : XXXIII. 11-12)
- (114) [*Celi jor que la nef ariva en Cornoaïlle.*] **estoit** li rois montez en une
 that day that the ship arrived in Cornwall was the king ascended in one
 soe tor.
 his tower
 'That day when the ship arrived in Cornwall, the king had ascended one of his towers.' (Tristan, p.45 : 18. 7-8)

- (115) *Car [il] covient que tu soiez tentez ausi come fu Job e que tu*
 tor it behooves that you are.SUBJ tempted just like was Job and that you
veinques le deable par fine pacience.
 vanquish the devil by fine patience
 ‘For it is necessary that you be tempted like Job and that you defeat the Devil by
 noble endurance.’ (Eustace, p. 12: IX. 13-15)
- (116) *[Il] avint que li costiax chei delez Sador.*
 tt happened that the dagger fell beside Sador
 ‘The dagger happened to fall next to Sador.’ (Tristan, p.65 : 67. 29-30)

The pragmatic flexibility of the prefield in Old French is completely on a par with that found in the modern Germanic languages, a fact which enables us to draw another conclusion:

(117) **Conclusion III:**

The prefield in late Old French was not reserved for topics, nor is it possible to make any strong qualitative generalization regarding the informational structural partitioning of the clause in linear V2 strings.

3.4 Inversion

The previous section has shown that there are significant similarities between the prefield in Old French and modern Germanic with respect to both the qualitative dimension (3.3), the quantitative dimension (3.3.1), and information structure (3.3.2).

However, we are still a long way from establishing a credible V2 hypothesis. In order to do that, we need to be able to make a significantly stronger claim than just showing that the prefield is not reserved for topics or any kind of particular information-structure; namely to demonstrate that the fronting of a non-subject constituent *automatically triggers inversion*. We are not in a position to show this yet, since we have only considered linear V2 strings so far. Needless to say, in a linear V2 string with a non-subject constituent in the prefield, the subject (if expressed) will always follow the verb. Recall from table 3.1 that the corpus featured around 15% linear V3. The analysis of these strings will therefore be crucial to the overall understanding of the grammar. We will defer the discussion until section (3.7), and first consider inversion in more general terms.

It was observed in section 3.3.1 that almost half of all linear V2 clauses in both texts feature a non-subject constituent in initial position (cf. table 3.3). It must be emphasized that these percentages do not in fact show the amount of inversion in main clauses, for two different reasons. First, these percentages were calculated only from linear V2 strings. Second, the fact that over half of the sentences feature a non-subject constituent in initial position does not entail that all of these sentences exhibit surface inversion of the subject and the finite verb. The reason for this is that Old French, like its ancestor Latin and its Romance sister languages, was a *null-subject language*.¹³ As a consequence, *non-subject-initial V2 ≠ inverted V2*, since many of the V2 strings lack an overt subject. In chapter 2, a V2 language

¹³By this I simply mean that Old French regularly allowed referential pronouns to be phonetically unexpressed. As already noted, there exists a parametrically defined notion of ‘pro-drop language’, going back to the work of Rizzi (1982) and subsequently elaborated by others. Such a distinction can no longer be made in a non-parametric approach such as the present investigation.

was defined as a language which derives inversion structures by V-to-C movement (plus restrictions on the prefield). We must therefore look closer at the issue of inversion, since inversion strings must be considered completely fundamental to the acquisition of a V2 grammar. In principle, the subject of a clause can be either preverbal, postverbal/inverted, or null. The tables 3.5 and 3.6 provide information about how these options pattern in main clauses in the corpus, both in general and distributed over different predicate classes.

Once again, the texts are very similar, in particular with respect to the overall distribution in the ‘Total’ column. As for the possible interaction between predicate class and subject position, there is some variation, but no clear pattern of interest emerges, as most tendencies are particular to only one of the texts. There is slight preference for null subjects with unaccusative verbs, although this tendency is only statistically relevant in *Tristan* (p-value 0.0002, d.f. 1, Chi-square 13.87).

Table 3.5: The position of the subject (S) distributed over different predicate classes in main clauses in *Tristan*

	Transitive	Unaccusative	Copula	Functional	Total
Preverbal S	314 (54.80%)	103 (41.70%)	54 (45.38%)	24 (40.00%)	495 (49.55%)
Postverbal S	94 (16.40%)	40 (16.19%)	26 (21.85%)	21 (35.00%)	181 (18.12%)
Null S	165 (28.80%)	104 (42.11%)	39 (32.77%)	15 (25.00%)	323 (32.33%)
Total	573 (100.00%)	247 (100.00%)	232 (100.00%)	119 (100.00%)	999 (100.00%)

Table 3.6: The position of the subject (S) distributed over different predicate classes in main clauses in *Eustace*

	Transitive	Unaccusative	Copula	Functional	Total
Preverbal S	135 (49.82%)	62 (46.97%)	29 (54.72%)	10 (47.62%)	236 (49.48%)
Postverbal S	49 (18.08%)	21 (15.91%)	15 (28.30%)	3 (14.29%)	88 (18.45%)
Null S	87 (32.10%)	49 (37.12%)	9 (16.98%)	8 (38.10%)	153 (32.08%)
Total	271 (100.00%)	132 (100.00%)	53 (100.00%)	21 (100.00%)	477 (100.00%)

There are two very important lessons to learn from these figures. The first is simply that inversion is robustly attested, reaching more than 18% in both texts. This almost amounts to one fifth of all main clauses and must be considered a salient acquisitional cue, halfway between the 13.6% found in Westergaard’s study of the acquisition of V2 in Norwegian (Westergaard 2009:67) and Yang’s (2003) corresponding 23% for Dutch. Although it seems naive to put too much faith in any kind of magic number that triggers the acquisition of grammatical properties, we may conclude that, far from being a marginal phenomenon,

inversion was frequently employed in Old French. It should also be noted that the inversion strings include both nominal and pronominal subjects. The pronominal inversion string CVSpX, which is considered so crucial for the acquisition of a V2 grammar in Sitaridou (2012) – for reasons we will return to in section 3.5.2 – is reasonably well represented in both texts with 6.91% in *Tristan* and 8.79% in *Eustace*.

The second important observation is that inversion does not show any statistical sensitivity to the predicate class variable. This is crucial, since it suggests that inversion does not only arise in particular configurations where the subject is exceptionally allowed to surface in a position to the right of the verb. This is the case in many modern Romance languages, where postverbal subjects are much more available with unaccusative predicates than with transitive verbs. In Rinke and Meisel (2009), it was argued that this is the case for Old French too, and that the amount of inversion is a product of different lexical choices made in different texts, more specifically the occurrence of different types of verbs. R&M claim that ‘inversion is much more common and natural with some verbs than with others’ (Rinke and Meisel 2009:115). According to the authors, the predicates that provide propitious contexts for inversion are intransitive verbs of motion, the copula *être* and some other unaccusatives like *mourir*, *commencer*, *finer*.

The figures in tables 3.5 and 3.6 do not support this view. While there is a slightly stronger tendency for inversion with the copula, inversion is just as frequent with transitive verbs as with unaccusative verbs (in fact slightly more frequent). We may therefore draw another important conclusion:

(118) **Conclusion IV:**

The evidence does not support the hypothesis that inversion in late Old French is sensitive to the type of predicate employed.¹⁴

3.4.1 Foulet’s generalization and the position of the null-subject

We have already seen that inversion is a rather robust phenomenon in our corpus, reaching magnitudes of around 18% in both texts. However, there is reason to assume that the figures in tables 3.5 and 3.6 conceal an ever stronger inversion pattern than what can be directly read off the row ‘postverbal S’. The reason for this is that approximately one third of all strings lack an overt subject. Although inversion has been defined as a surface term in this thesis, making it all but senseless to talk about ‘covert inversion’, it is clear that the unexpressed subject must be assigned a structural position by the child acquiring the language. There exists a hypothesis going back at least to Foulet (1930, first edition 1919), which I will refer to as *Foulet’s generalization* and which states that null subjects in (later) Old French were generally only possible in postverbal position. This hypothesis has subsequently gained much currency in both the traditional and modern research literature.

Within the generative paradigm, a very influential and largely accepted analysis of this traditional claim was developed by Adams (Adams 1987b,a). Simplifying somewhat, Adams’ theory is as follows. The null subject is the empty category *pro*, and this category must be licensed by the head I^0 , which identifies the position and the content of *pro*. I^0 is only able to identify *pro* under government, and this structural configuration is only achieved in inversion structures when I^0 moves to C^0 . The content of *pro* is identified by coindexation with the features of C^0 . In subject-initial main clauses, the subject in SpecCP is not governed by the

¹⁴However, the position of the subject might potentially be sensitive to the type of predicate, cf. section (3.5.2.1).

verb, and in embedded clauses the verb does not raise to C^0 , and hence *pro* is not possible (see also Roberts 1993:110, Vance 1997:204).

Although some exceptions can be found, the generalization that null-subjects are structurally postverbal has been accepted by those researchers who see Old French as a verb second language (Vance 1993, 1997; Roberts 1993; Hirschbühler and Junker 1988; Hirschbühler 1990; Hulk and van Kemenade 1995; Vanelli et al. 1985; Salvesen 2013; Wolfe 2015b). On the other hand, it has been explicitly rejected by many researchers who also reject the V-to-C analysis of Old French (Kaiser 2002; Rinke and Meisel 2009; Zimmermann 2009). Rinke and Meisel (2009) argue that clauses without an overt subject cannot be used as an argument in favour of V2, as a pronominal null-subject in their view could occur both both pre-verbally and post-verbally, even adding that ‘null-subjects are of course more likely to occupy a preverbal position because they usually constitute the topic of the sentence’ (2009:97). In section (3.7.6, I will demonstrate that this position is empirically untenable, but since it is not a trivial thing to build up an empirical argument for the structural position of a null-subject, it is necessary to make a rather long detour. For the moment we must therefore leave this issue aside.

3.5 The acquisition of Old French phrase structure

The data considered so far has already permitted us to draw some clear conclusions about the major word order patterns and the nature of the prefield. These conclusions are clearly compatible with the hypothesis that Old French featured a verb-second syntax and even provide some suggestive evidence in favour of that hypothesis. However, the type of evidence considered so far has been mostly quantitative in nature and is therefore ultimately inconclusive with respect to the syntactic structure of the language. In order to establish this structure, we must submit the data to detailed qualitative analysis with a view to uncovering the phrase-structural organization of the clause that produces these surface word order patterns. This line of inquiry seeks to understand what kind of structure the child acquiring the language must assign to the input strings in order to make sense of the data. I repeat for convenience the *String-to-Structure-Assignment-Principle* from Chapter 2:

String-to-Structure-Assignment-Principle (SSAP):

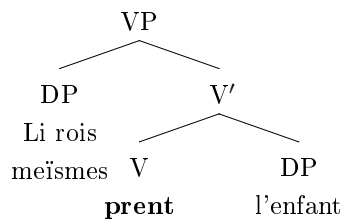
Children assign the minimal structure that is consistent with the global string input in a maximally economic way.

Taking this principle as our point of departure, we will now consider some more qualitative evidence in order to find out exactly how much structure the Old French main clause conceals. The next sections are therefore presented as a kind of ‘acquisitional tour’ of the Old French clause; needless to say, the actual acquisition process must be much more complex. The language acquirers do not get the evidence presented in such orderly, step-by-step fashion, but must presumably rather work with multiple competing analyses at a time based on chaotic, perhaps partially contradictory (due to production mistakes etc.) or at least sub-optimal input. Still, the actual logic of the process must be something along these lines; try to make do with what you already have, revise only if necessary.

3.5.1 Subject-initial strings: SVX and SpVX

We start out by taking into consideration the subject-initial strings which constitute slightly less than 50% in both of the texts. Consider the minimal example in (119). This clause, consisting of a transitive verb and its internal and external argument, can be expressed in accordance with the conventional X-bar schema within the confines of a single maximal projection, the VP:

- (119) [*Li rois meïsmes*] **prent** l' *enfant*
 The king-NOM himself takes the.CL child
 'The king himself takes the child'
 (Tristan, p.48 : 23.11)



If all clauses were like (119), then, the child would get away with the structure in (119), and Old French clausal syntax would be nothing more than a VP. In other words, I will not assume that the verb must raise to tense-related projections for independent reasons to merge with inflectional morphology. Many languages, like Modern English or Norwegian (in embedded clauses) do not need to raise their lexical verbs out of the VP to combine with inflection. Rather than assuming affix-lowering or covert movement of the verb, a simple way of accounting for this is to assume that hierarchical position and inflectional form are logically independent. Explicitly stated, this amounts to adopting some kind of *Lexical Integrity Principle*. I will not be concerned with the nature of the syntax-morphology interface.

Of course, no language can make do with such a minimal structure. All kinds of adverbial phrases must be accommodated, thereby stretching the clause considerably. Furthermore, Old French had already developed many of the modern periphrastic constructions involving auxiliaries related to the grammatical expression of aspect (120) and voice (121), thereby lexicalising further head positions in the clause:

- (120) [*Vos*] *m'* **avez osté** *de la greignor prison...*
 You me.CL have removed from the worst prison...
 'You have freed me from the worst prison... ' (Tristan, p.59 : 52.7)

- (121) [*Li comendemenz l'empeor*] **fu** *fez*
 the order-NOM the-emperor.OBL was made
 'The order of the emperor was carried out' (Eustance, p.40 : XXXVI.1-2)

Compared to (119), the last two examples provide a wealth of new information which calls for a dramatic revision of the clausal structure. The child is forced to expand the clause and make room for a higher verbal projection. At the same time, it is clear that the subject position cannot be Spec-VP, but rather the specifier of this higher projection. Notice

also that (120) provides evidence that oblique pronouns pattern differently than nominal objects, gravitating to a leftwards position in the clause as a clitic, although it is not clear from this example if the host is the first constituent (enclisis) or rather the finite verb in second position (proclisis).

Based on the evidence from (120 and (121) alone, the child might hypothesize that all kinds of auxiliaries head the same projection, which we might accordingly call AuxP. The immediately preverbal subjects in both cases would indeed seem to indicate this. The AuxP-hypothesis then breaks down when faced with strings like (122), since there are not two, but three verbal projections simultaneously lexicalised in this clause. The child is forced to further extend the verbal projection and thereby the clause, and (122) gives evidence that aspectual auxiliaries precede the passive auxiliary which again precedes the lexical verb:

- (122) [*J*] *'ai esté noriz en cele vilete ...*
 I have been nourished en this village ...
 'I was brought up in this village ...' (Eustace, pp.32-33 : XXVIII.31-32)

Furthermore, when we add adverbs (123) and negation (124) to the mix, their position provides positive evidence for a layer of projections between the finite verb and the lower, non-finite verbal projections:¹⁵

- (123) [*Je*] *ai bien veu les aumones que tu fez chascun jor as povres...*
 I have well seen the alms that you make every day to-the poor...
 'I have certainly seen the alms you give to the poor every day... ' (Eustace, p.5 : III.27-28)

- (124) ... [*il*] *n' avoit pas loiaument ovré envers le roi Canor.*
 ... he NEG.CL had not loyally acted against the king Canor.
 '... he had not acted loyally towards king Canor.' (Tristan, p.60 : 55.8-9)

Furthermore, adopting the cartographic logic established by Cinque (1999) and assuming that the position of these adverbs is fixed (by children or UG), rather than for instance adjoined freely at different junctions of the clause, different combinations of adverbs and negation allow children to construct the details of the IP-area in a piecemeal fashion. Thus, (125) shows that negation precedes *bien*, while (126) shows that certain temporal adverbials like *encore* precede negation. While these facts pertain to the IP-domain, they will become directly relevant later for the understanding of higher clausal syntax as well:

- (125) [*Il*] *n' ot mie bien sa parole finée, qant une voix li vint del ciel...*
 He NEG. had not well his word finished when a voice him.CL came from-the sky

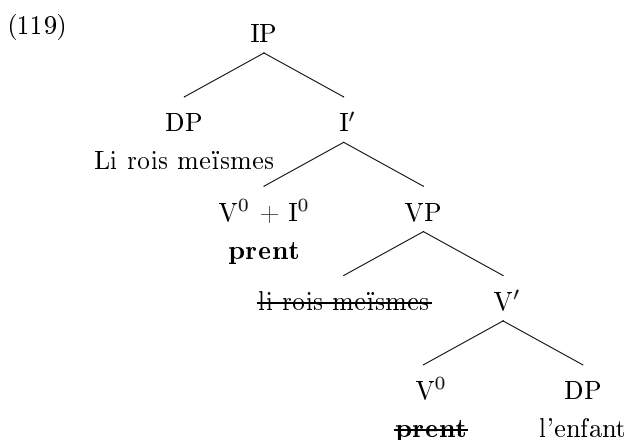
¹⁵Strictly speaking, *pas* (and *mie*, *point*, etc.) is not a negator, but a Negative Polarity Item (NPI) used to reinforce negation, the expression of which is only dependent on the clitic negator *ne* in Old French. Already at this stage of the language, the reinforcing function of *pas* is much bleached. See Inham (2014) for discussion and references.

‘He had barely finished speaking, when a voice came to him from the heavens...’
(Eustace, XX, 12-13.)¹⁶

(126) [*La roïne Chelinde*] *ne* **savoit** *encores pas que...*
the queen Chelinde NEG. knew still not that

‘Queen Chelinde did yet not know that...’ (Tristan, p.53 : 36.14-15)

By the same cartographic logic, the postverbal position of temporal adverbs and negation in (126) also reveal something else, namely that finite lexical verbs also raise out of the VP to head a higher projection, which can now also be identified as IP on the basis of similar distribution and the fact that modals and finite lexical verbs are in complementary distribution. By the SSAP given above, which states that children take into account the *global input*, the minimal clause in (119) can no longer be assigned a simple VP parse (119,) but must be modified:¹⁷



There will also have been evidence for at least one additional verbal projection, since modal verbs can precede and dominate aspectual auxiliaries. In our corpus, subject-initial strings only contain modal auxiliaries which dominate simple lexical verbs, but if we exceptionally allow ourself to minimally modify a CVX-string into a SpVX-string (127) by changing a preverbal expletive *si* into a pronominal subject,¹⁸ it is clear that the verbal projection must be expanded further:

¹⁶This example also shows a case of raising of the direct object to a position above the participle, an instance of a more general tendency in OF to employ short movement in the lower clausal area. The same might perhaps apply to the VP-adverb 'loiaument' in the previous example. These displacement superficially resemble (and may well be) scrambling of the kind found in the West Germanic languages, but it is not clear that they are driven by information-structure; in fact, I am rather inclined to consider this some kind of 'formal' or 'stylistic' optional movement. See also Salvesen (2013:141–142), and for a similar phenomenon in Old Italian, Poletto 2006.

¹⁷Notice that this is completely different from claiming that the child immediately parses the clauses into an IP due to something innate in UG. Head-movement is often assumed to involve left-adjunction (Roberts 2011), but I leave this extra structure out of the tree for simplicity.

¹⁸While it is generally not advisable to create examples and pass grammaticality judgements on dead languages, this example is completely uncontroversial. It is also clear that such strings would have been available in the input, although the corpus fails to provide them. The original string reads:

(i) ...si n' i **poïst** pas tant avoir demoré sans morir...

- (127) *[Il] n' i poist pas tant avoir demoré sans morir*
 he NEG. there.CL can not so-long have lingered without die
 'He cannot have stayed there so long without dying.' (Unattested, slightly modified
 from *Tristan*, p.51 : 32.3)

By the cartographic Principle of Transitivity, the child will now be able to deduce the whole verbal auxiliary sequence. Passives precedes the lexical verb (122), Aspect precedes Passive (123), and Modals precedes Aspect (127). The leftmost verb always carries the inflectional morphology¹⁹ and selects the morphological form of the next verb in the sequence, and we will therefore identify it with IP. Regarding the subject, we can also observe that SVX and SpVX strings provide evidence for a single subject position, Spec-IP. This hypothesis must be revised in the next section, when we consider the non-subject-initial strings.

This more or less exhausts the evidence it is possible to cull from subject-initial strings. Naturally, there is a wealth of adverbial positions in the IP-area, but these need not concern us beyond what has already been said. In cases of a contrastively stressed preverbal subject or (in the rare cases) of an indefinite focal, preverbal subject (128) we might ask if these activate left-peripheral positions associated with the corresponding functional projections uncovered by cartographic research.

- (128) *Lors li dist la dame: Beau sire, ou sont nostre enfant? Dame,*
 then him.CL said the lady beautiful sir where are your children? Lady,
dist il, [bestes sauvages] les ont devorez.
 said he, wild beasts them.CL have devoured.
 'Then the lady said to him : 'Good sir, where are your children?' 'My lady', he said,
 wild beasts have eaten them.' ' (Eustace, pp.35-36 : XXXI.11-12)

The approach adopted here dictates a negative answer to this question. The children might get a prosodic cue that allows them to relate such subjects with a slightly different reading than that of regular subjects (which we might consider to carry an aboutness topic reading by default), but SVX-strings as such provide no evidence for assigning them to a different syntactic position.²⁰

3.5.2 Non-subject-initial V2 strings:

The conclusions we arrived at in the previous section were not very controversial. These subject-initial strings gave no evidence of V-to-C movement and therefore no support to the hypothesis that Old French was a V2 language, although they were of course clearly compatible with such a hypothesis. In fact, judging by the evidence considered in the previous section alone, we would have to conclude that French of the thirteenth century had already developed the SVO-syntax of modern French, barring some variation in the lower part of the clause, due to seemingly optional local left-displacement of VP-material.

¹⁹This does not necessarily apply to non-subject-initial strings, where a participle or an infinitive may precede the verb, as was shown in (100) above. There is clear evidence that this is not a head position, however, so the child will analyse such cases as phrasal movement, presumably of the VP.

²⁰But as we shall see in the next section, inversion strings will provide evidence to children that the corresponding contrastive or focal readings occupy a (i.e. one)left peripheral position. Thus, the *global input* might suggest that SVX-strings should be assigned more than one syntactic structure.

We have already seen in section 3.3 that this is far from the case, and the situation becomes more complex when we now move to considering linear V2 strings with a non-subject constituent in initial position. These are even marginally more frequent than subject-initial strings in both texts, and include a considerable amount of overt inversion structures. These inversion structures provide incontrovertible evidence that Old French was not like modern French. The question is how to capture this difference analytically, and there is no general agreement on the answer to that question. One possible answer is that Old French was a V2 language which featured consistent movement of the verb into the C-domain. If V-to-C movement is consistent and always takes place, this means that all of the strings to be considered in this section are inversion strings from a structural point of view, in the sense that the verb has moved above the position of the subject. This would apply also to CVX strings, where there is no overt subject, in accordance with Foulet's generalization.

Since the S(p)VX strings we saw in section 3.5.1 provided strong evidence for a high subject position, presumably in Spec-IP, the conclusion that inversion strings involve V-to-C movement might even seem inescapable at first. However, the matter is somewhat more complicated, since the *global input* from inversion strings clearly shows that there is more than one surface position for the subject in Old French. The key evidence comes from cases of 'non-contiguous inversion' in the terminology of Vance (1997), in other words inversion strings where the finite verb and the postverbal subject are separated by one or more constituents. We will now consider these strings.

3.5.2.1 The position of the subject in inversion strings

There is more than one position available to the subject in inversion strings in Old French. First, there is a low position for subjects after non-finite lexical verbs. This predominantly occurs with unaccusative verbs (129), but sometimes also with passive transitive verbs (130). In the following examples, the subject is underlined.

(129) [*Tant*] *ont alé celi jor li marinier*...

So-far have gone that day the sailors...

'The sailors went so far that day. . . ' (Tristan, p.55 : 44.1)

(130) [*par son preeschement*] *fu torneé grant partie de la gent de cele terre a la loi*

by his preaching was turned great part of the people of that land to the law
crestiene.

christian.

'Through his preaching, a great part of the people of that land was converted to the Christian faith.' (Tristan, p.40 : 1.5-6)

These examples feature inversion strings where the subject follows both the finite auxiliary and non-finite verbs (see Roberts 1993:56, Lemieux and Dupuis 1995:92-93, Vance 1997:75-80, Salvesen and Bech 2014:212-214), what I will call R-inversion. However, example (130) shows that the subject is not necessarily string-final, and a natural interpretation of such cases would be that these subjects remain inside the VP, possibly even in complement position.²¹ ²² It is not uncommon that subjects of unaccusative verbs may occupy lower

²¹Vance suggests that they are rather in Spec-VP and that their surface position after participles and infinitives is rather to be derived through moving the latter to a projection above the VP (1997:82).

²²Some string-final subjects seemingly involve extraposition due to phonological weight or to provide following relative clauses with an immediate antecedent.

position in the clause than the subjects of transitive, or in more general terms, agentive verbs. For instance, subjects are allowed to follow non-finite forms of unaccusative verbs in Modern English and the Scandinavian languages too, provided an expletive occupy the prefield, cf. ME: ‘There has arrived a letter for you’/Norwegian: ‘Det har kommet et brev til deg.’²³

Although not as frequently as with unaccusatives, this low subject position can also be observed with transitive verbs, as witnessed by the following examples:

- (131) [*Beles miracles et bels vertuz*] *a fait* *ma dame Venus* ...
 beautiful miracles and beautifuls deeds has done my lady Venus
 ‘Great miracles and great deeds my lady Venus has performed!’ (Tristan, p.65 : 68.10)

- (132) ... [*tant*] *te devroit haïr Sador tes freres!*
 ... such you.CL must hate Sador your brother
 ‘... how much your brother Sador must hate you!’ (Tristan, p. 43: 11.7)

Perhaps the postverbal position of these subjects is linked to their focal nature. Interestingly, both these examples, which are the only ones in the corpus where the subject of a transitive verb unequivocally follows the non-finite main verb, are somehow *doubly focal* in the sense that both the subject and the initial constituent represent new information. This clearly applies to (131), since the passage from which it is taken contains no mention of either miracles or Venus in the preceding discourse. If we accept the frequently proposed, but certainly disputable, analysis whereby adverbials like *tant* are foci since they represent a scalar contrast (Vanelli 1998:82; Ledgeway 2008:450), this analysis extends to (132) as well, since *Sador* must here be considered inactive in the preceding discourse. It might be the case that these constructions carry focus in both a left-peripheral position and the string-final focus position associated with Heavy Inversion. What is clear, is that there is an emphatic, almost mirative reading at hand.

All of these inversion constructions, whether they subsume more than one single syntactic structure or not, involve DPs and show some surface similarity with the modern Romance inversion structures, since they surface to the right of the entire verbal complex. However, this is clearly not the normal position of the subject in our corpus, as illustrated by the following sentences, which include both transitive (133) as unaccusative (134) verbs. Here, the subject surfaces between the finite auxiliary and the participle:

- (133) [*Si*] *avoit ja li rois esleü ces qui le champ devoient garder.*
 SI had already the king selected those who the field should guard.

-
- (i) *Celi jor meïsmes... vindrent au chastel [noveles ou li rois Canor estoit].*
 that day self... came to-the castle news where the king Canor was
 ‘That very day, news came to the castle about the whereabouts of king Canor.’

The structural analysis of these ‘Heavy Inversion’ constructions varies somewhat in the literature (Déprez 1988; Valois and Dupuis 1992; Vance 1997), and we will not be further concerned with them, as the matter at stake is rather the position of the subject than that of the verb.

²³Notice however that the subject must be indefinite in English and Norwegian (*There has arrived the letter/Det har kommet brevet), indicating that there are additional pragmatic constraints in the latter languages which seem absent from Old French. In fact, Vance even claims the subjects of these constructions generally represent old information in OF (Vance 1997:77), but this can at best be a tendency, as (130) is a focal (although anchored) subject.

‘Thus the king had already picked out the ones who should stand ground.’ (Tristan, p. 60: 55.4-5)

- (134) [*Celi jor que la nef ariva en Cornouaille*], *estoit li rois* montez en une
That day that the ship arrived in Cornwall was the king ascended in one
soe tor.
his towers.

‘On that day when the ship arrived in Cornwall, the king had ascended to one of his towers.’ (Tristan, p.45 : 18.7-8)

- (135) [*Sanz grant senefiance*] *ne porroit pas ceste chose estre avenue*.
without big significance NEG could not this thing be happened.

‘For this thing could not have happened without some greater meaning.’ (Tristan, p. 47 : 21.8-9)

In other words, these examples involve G-inversion. With some minor, language-specific exceptions,²⁴this subject position is ungrammatical across the board with DP subjects in the modern Romance languages. I will henceforth assume that inversion strings with a single transitive verb instantiate the same syntactic structure, such that (136) is assigned the same structure as (133–135).

- (136) *A cest fet aperçut bien [Eustaces]...*
At that party noticed well Eustace...

‘At that banquet Eustace noticed...’ (Eustace, p. 14: XI.9)

What are the reasons that have led some researchers to reject a V-to-C analysis of these sentences, which appear to be *string-identical* to the inversion constructions of Germanic V2 languages?

The crucial observation, made for Old French within the generative framework by Deprez (1988) and since developed by others (Vance 1997), revolves around the position of the subject with regards to certain adverbs and negation. Recall from section 3.5.1 that we established a cartographic mini-sequence *pas* > *encores* > *bien* situated between the IP projection and the lower verbal projections. It can clearly be seen from the above examples that the subject position in main clause inversions is situated below the negation *pas* (135) and even below the lower adverb position *bien* (136), which might be taken to demarcate the edge of the VP (Cinque 1999). It is therefore assumed now by many researchers that the position of the subject may be no higher than Spec-VP (Lemieux and Dupuis 1995; Vance 1997; Rinke and Meisel 2009; Salvesen and Bech 2014) an idea which receives support from the independently established *VP-Internal Subject Hypothesis*, the hypothesis in transformational generative syntax that subjects of transitive verbs originate in Spec-VP (Zagona 1982).²⁵

²⁴European Portuguese optionally allows the subject to intervene between the finite verb and the participle in some cases (Ambar 1992:80).

²⁵There is another logical option concerning the position of these adverbs, provided we consider the possibility of adjunction. Concretely, if one considers the verb and the initial XP to reside in C⁰ and Spec-CP respectively and pronominal subjects to cliticize to C⁰, one could argue that the nominal subject is indeed in Spec-IP, if adverbs and negation are adjoined to IP. However, it has been demonstrated convincingly by Vance (1989) that IP-adjunction is banned in Old French with the exception of a few adverbs – in fact the NPIs *onques* and *ja*, see section 3.7.1.3 – since adverbs never precede the subject in Spec-IP in embedded clauses.

It should be clear why this is relevant to the position of the verb as well. Since children according to the SSAP take into consideration the global input, they might use the evidence provided by these inversion strings to analyze Spec-VP as the basic subject position. Furthermore, since children only construct the *minimal structure* consistent with the input, there is no need to push the verb into the C-layer to obtain surface inversion; it will quite suffice to raise the verb to I⁰. This is indeed analysis developed by those researchers who reject the V2-hypothesis for Old French (Kaiser 2002; Rinke and Meisel 2009).²⁶

The V-to-I analysis has many consequences, but before we explore them in more detail, it is important to clarify something. While it is true that the evidence clearly shows that the position of the subject is at least not always Spec-IP in Old French main clause inversion, the same applies with equal force to several of the modern Germanic languages. We have already seen this in chapter II, but I repeat for convenience the following, completely unmarked inversion structure from modern Norwegian:

- (15) *Brevet har dessverre sannsynligvis ikke [faren min] mottatt.*
 letter-the has unfortunately probably not father-the mine received.
 ‘My father has unfortunately probably not received the letter.’

The regular subject position for DP subjects in inversion structures is not Spec-IP in all Germanic V2 languages either. In other words, the relatively low position of the DP subject is in itself no convincing argument against V-to-C, as children acquiring modern Scandinavian V2 would get equally well away with a V-to-I parse for the sole sake of accounting for main clause inversion. Once again, it is the *global input* which forces a V-to-C parse in modern Germanic, and we shall see that exactly the same applies to Old French.

Based on similar facts from modern Germanic, Sitaridou (2012) argues that subject-verb inversion with nominal subjects is no guarantee for V-to-C movement, and that the relevant strings are those involving pronominal subjects, as these generally must be adjacent to the finite verb and cannot be separated from the latter by intervening material. It was already mentioned earlier that these are very well represented in our corpus and amount to around 7–9% of the total input. They also include several examples of the string ‘Adv-Aux-SVO’, which according to Kaiser (2002), building on Fodor (1998), constitute unambiguous evidence for the positive parameter setting of V2. I take it that ‘Adv’ is supposed to mean *adverbial*, not *adverb*, as it is hard to see what should be so special with an adverb as the first constituent. There are examples of initial adverbs as well, of course:

- (137) *[Et por ce] vouldroit il avoir doné la moitié de son reaume...*
 and for this would he have given the half of his kingdom
 ‘And therefore he would have given half of his kingdom...’ (Tristan, p.61 : 58.6-7)

- (138) *[par ceste chose] porroit il avoir Chelynde.*
 by this thing could he have Chelynde.
 ‘By this trick he could have Chelynde.’ (Tristan, p. 42: 9.3)

²⁶Whether rejecting V-to-C in fact automatically amounts to rejecting verb-second status depends on the definition of verb-second. For the researchers under discussion in this section, V2 = V-to-C (plus restrictions on prefield), so the answer is clear. Other researchers have also rejected the V-to-C analysis in favour of a V-to-I analysis without questioning the appropriateness of the label ‘V2’ for OF (Lemieux and Dupuis 1995).

- (139) *[miels] voudroie je morir a honor que vivre a honte.*
 better would \bar{I} die in honour than live on shame.
 ‘I would rather die with honour than live in shame.’ (Tristan, p.61 : 58.17)
- (140) *car [la] le porront il trover.*
 for there him.CL can they find.
 ‘For there they can find him.’ (Tristan, p.57 : 48.20-21)
- (141) *[Maintes foiz] le t’ avoie je desfendu.*
 many times it.CL you.CL have \bar{I} forbidden.
 ‘Many times I have forbidden you this.’ (Tristan, p.65 : 67.14)
- (142) *... car [ce] avoit ele requis a Nostre Seigneur, ...*
 ...for this had she asked of our lord ...
 ‘For she had requested this from Our Lord ...’ (Eustace, p.21 : XVIII.5-6)

Furthermore, pronominal subjects are always string-adjacent to the finite verb in linear V2 inversions, meaning that the string CVCSpX is unattested and that the pronominal subjects consistently precede all IP-adverbs (143–144) and negation (145). The fact that pronominal subjects without exception precede all IP adverbs and appear adjacent to the verb could be interpreted as evidence that they occupy a position at least as high as Spec-IP. If this is the case, then we are clearly dealing with V-to-C movement in these inversion structures:

- (143) *[de totes choses] veil je bien ovrer a vostre volenté.*
 from all things will \bar{I} well work at your will.
 ‘In everything I will act according to your will.’ (Tristan, p.40 : 3.4-5)
- (144) *[A ces enseignes] poons nos bien conoistre que il est hors de pooir.*
 On these signs can we well know that he is man of power.
 ‘We can tell from these signs that he is a man of power.’ (Tristan, p.55 : 43.6-7)
- (145) *[De ceste mort] ne le puis je pas oster.*
 From this death NEG him can \bar{I} not remove.
 ‘I cannot save him from this death.’ (Tristan, p.66 : 71.14)

But again, the situation is more complicated, since it is possible to argue that subject pronouns cliticize to the verb in inversion structures. If this analysis is correct, subject pronouns in OF evince a dual nature, since they are clearly not clitics when preverbal.²⁷ Furthermore, this analysis does not permit us to distinguish between a V-to-I and a V-to-C parse, since pronominal subjects would cliticize to the verb whether the latter is in I⁰ or C⁰, resulting in the same surface structure. This is also the argument employed by the researchers who reject the V2 hypothesis for Old French (Rinke and Meisel 2009; Kaiser

²⁷In OF it is for instance possible to separate the preverbal subject from the verb:

- (i) Veritez est que quant Joseph d’Abarematie se fu partiz de Sarraz ensi com vos et maint autre le sevent, je, qui estoie chevaliers del reame de Sarraz et chevaliers le roi Mordrain ... si ne **menai** si bone vie ...
- (ii) Et je que li **ferai** a cesti point?

and Zimmermann 2011). We are still not able to distinguish empirically between the two competing analyses, and in the absence of such evidence, the principle of the SSAP dictates the minimal parse, in other words V-to-I.

Although all the evidence reviewed so far has revealed remarkable similarities between Old French and the modern Germanic V2 languages, it is still possible to defend the more economical V-to-I parse. At this point, there is no more information to collect from linear V2 strings. We will therefore leave these strings aside for the moment and consider the other word order patterns which are attested in main clauses. First, we will submit V1 strings to scrutiny in section 3.6, while V3 strings will be the topic of section 3.7.

3.6 Verb-initial clauses (VX, VSX)

Verb-initial orders were attested in both texts of the corpus, reaching 8.72% and 6.71% of all main clauses in *Tristan* and *Eustace*, respectively. These figures cannot strictly speaking be called marginal. Given the theoretical assumption that verb-second grammars contain some kind of rule that prohibits the prefield from being left radically empty, often formalised through an EPP-feature that triggers the merger or movement of an XP to the relevant specifier position, we need to explain how these word orders arise in Old French.

A closer scrutiny of the data quickly reveals that appearances are somewhat deceptive in this case. In fact, not a single main clause in the corpus starts with the finite verb as the first *word* of the clause.²⁸ Rather, these V1 clauses are in their vast majority introduced by the item *et* – ‘and’ – while a handful of cases are introduced by the item *ne*, at the surface identical to the normal proclitic negative marker.²⁹ I deliberately use the vague term ‘item’, as we shall see that there is some debate over the exact categorial status of these expressions.

Both of these two constructions come in two guises; the vastly most productive pattern is the subjectless string *et/ne-VX*, while the other option is provided by the string *et/ne-VSX* with a postverbal nominal subject. Postverbal pronominal subjects in V1 strings are not attested in our corpus, meaning there is no occurrence of the string *et/ne-VSpX*. This has been observed before (Vance 1993), although the reason for this lack of pronominal subjects remains elusive.

The total absence of main clauses with the verb in absolute first position is also very much as expected in light of the research literature on the evolution of Old French word

²⁸There is one possible exception:

- (i) *Et en chascune bone vile, fust cité ou chastel, avoit adonc un perron*
 And in every good village BE.SUBJ. city or castle had then a platform.
 “And in every good town, be it city or castle, there was at that time a platform.”

It is clear that this is not really a true declarative clause or even a true main clause at all, but rather a special construction, intimately tied up with the subjunctive mood of the verb and a particular irrealis force which seems to straddle the border between a hortatory main clause and an adverbial conditional clause (*protasis*). Corresponding constructions feature V1 in several modern Romance and Germanic languages, including English, as shown by the translation.

²⁹Recall from chapter II, section 2.7, that only instances where *et* is considered to conjoin complete clauses have been included, while instances where *et* shares material with its preceding conjunct have been removed à priori and do not feature in the data provided in this thesis, as they do not strictly speaking represent full clauses. It should be noted that many cases are highly ambiguous, and that apart from a consideration of context, the sole guiding principle is often the punctuation provided by the editor. While there are multiple sources for potential misinterpretation here that might affect the quantitative data, the nature of the construction as such cannot be questioned; V1 clauses initiated by *et* are an authentic feature of Old French prose texts.

order. While earlier Old French texts not infrequently show the verb in absolute clause-initial position, Skårup observed that the initial position was no longer left radically empty by the turn of the thirteenth century (Skårup 1975:291). This is significant and seems to indicate that some change had made itself felt regarding the nature of the prefield. The conclusion lies near at hand that the prefield could no longer be left empty, but rather had to host some phonologically overt material. Interpreted this way, the nature of the V1 clauses could be linked to a V2 syntax.³⁰

But in what respect can we say that these clauses are structurally V2? The opinions differ in the literature about how to analyse these instances of *et-V* and *ne-V* clauses. Zimmermann and Kaiser (2010) provide a useful overview of this debate, the contents of which we will briefly recapitulate.

3.6.1 *Et-V* clauses

Generally speaking, two different options have been pursued by researchers when analysing these kinds of structures. Either *et* is considered to represent a coordinating conjunction in all cases, or it is considered to be a conjunction sometimes, and a kind of adverb in other cases. If the first option is pursued, the conclusion at first sight seems to be that we are truly dealing with verb-initial sequences. The other option, which was adopted by Foulet (1923) and accepted by others after him (Franzen 1939; Nissen 1943; Skårup 1975), makes it possible to argue that examples like the following feature the adverb *et* in first position, a full phrasal constituent triggering subject-verb inversion:³¹

- (146) *Sador aporta la demoisele en ce chastel que je vos di, et la mist*
 Sador brought the lady to this castle that I you.CL said and her.CL put
en une chambre. [Et] fu cele leanz bien trois jorz enz que ele
 in a room. And was that-one there well three days before that she
manjast...
 ate.SUBJ

‘Sador brought the lady to the castle that I spoke of, and put her in a chamber. And she was there for well three days before she took to eating...’ (Tristan, p.41 : 6.1-2)

- (147) *Lors la corut embracier e basier e acoler, [e] mercierent ambedui*
 Then her.CL ran embrace and kiss and hug and thanked both.NOM
mult le Sauveor del monde
 much the Saviour of-the world
 ‘Then he ran over and embraced her, hugging and kissing her. And both gave thanks to the Saviour of the world...’ (Eustace, p.35 : XXXI. 6-7)

³⁰These remarks are only valid to the extent that there is a continuity between the ‘real’ V1 clauses of the earlier Old French period (and if these changes truly reflect linguistic evolution; cf. section 6.3.1) and the *et-V1* clauses of the thirteenth century. This has been disputed, however (Vance 1993:300, see footnote).

³¹There is a long-standing debate in the literature about the possible influence of the particle/adverb *si* on *et*. I will have nothing in particular to say about this (for an overview, again see Zimmermann and Kaiser 2010), beyond the fact that they behave quite differently, as *si* regularly triggers inversion with nominal and pronominal subjects alike. The same point is raised by Vance, who even argues that in cases where *et* and *si* alternate in the same distributional contexts, it is rather the latter than is influenced by the former (Vance 1993:298). In our corpus, *si* has consequently been annotated as a full constituent in all cases and therefore enter the data as V2 strings (*si-V...*) or V3 strings (XP-*si-V*), never as V1 strings. For an analysis of *si* as a phrase, see Adams 1987a; Salvi 2004; Benincà 2006, and as head, Ferraresi and Goldbach 2003. See also Ledgeway 2008 for an analysis of *si* in Old Neapolitan.

On this analysis, such clauses in fact conform to the expected V2 pattern. There are two considerable complications for this analysis. First, if *et* is in fact adverb in these cases, the absence of postverbal pronominal subjects is completely unexpected, as adverbs generally allow inversion with all kinds of subjects (Vance 1993:291). Secondly, this analysis leads to inconsistencies, since clauses introduced by *et* are often followed by another constituent and then the verb, giving *et*-SVX strings (148a–148b) and *et*-CV(S)X strings (148c–148d) as well:

- (148) a. *Li mariniers, si come Dex le vost, morut, e [la dame] fu en*
 The sailor.NOM such as God it.CL wanted died and the day was in
sa delivre poeste.
 her free power.
 ‘The sailor died, at the will of God, and the lady was free again.’ (Eustace, p.21 : XVIII.7-8)
- b. *Et [cil] s’ agenoille devant lui, et reçoit le don, et l’*
 And this-one REFL.CL kneels before him and receives the gift and
en mercie mout durement.
 him.CL of.it-CL thanks much heavily.
 ‘And he kneels before him, receiving the gift and thanking him very heartily.’
 (Tristan, p.40 : 2.23-24)
 32
- c. *Sador remest avec ses freres qui estoient prudome et bon*
 Sador remained with his brothers who were prudhommes and good
chevalier durement. Et [mout] l’ amonestoient sovent qu’ il
 knights truly. And much him.CL counseled often that he
se mariast...
 REFL.CL married.SUBJ
 ‘And Sador remained with his brothers, who were *prudhommes* and good knights indeed. And they often strongly counseled him that he should marry...’ (Tristan, p.41 : 3.12-14)
- d. *E [ja] fu li cers mult esloigniez de tote la compaignie...*
 And already was the deer much removed of all the compaigny ...
 ‘And already the stag was far ahead of all the riders ...’ (Eustace, p.4 : II.19-20)

If we want to maintain a consistent analysis, we are forced to consider examples like (148c–148d) cases of linear V3, then. This is an unappealing and counterintuitive solution, and also very much against the otherwise firm tendencies of the language to put the verb in second position. Some researchers therefore prefer to consider *et/e* a true coordinating conjunction in these cases and therefore irrelevant to the computation of the clause itself. As pointed out by Zimmermann and Kaiser (2010), in the case of such a split approach to

³²Example (148b) illustrates well the difficulties in deciding the range of coordination of *et*. This example was annotated as a single clause, meaning that the two conjuncts *et reçoit le don* and *et l’en mercie mout durement* are considered sub-clausal structures, sharing the subject *cil* with the first conjunct. In principle, however, nothing excludes the possibility that this was intended as three or (more likely) two different clauses.

the categorial status of *et*, the choice of analysis is very much dictated by theory-internal considerations, in particular whether Old French is considered to be a verb-second language or not.³³ If there is indeed a categorial difference between *et* in (148a–148b) on the one hand and (148c–148d) on the other, one would like to see some independent evidence for it. Otherwise, the argumentation becomes circular.

In terms of semantics, little seems to speak for the split hypothesis. There is no obvious sense in which the *et* of *et-V* clauses is semantically richer than that of *et-SVX/CV(S)X* clauses; we cannot for instance translate it as ‘and also’. On the other hand, there is a piece of evidence that stems from the placement of pronominal and adverbial clitics. According to the descriptive generalization for Old Romance known as the Tobler-Mussafia Law (Mussafia 1898), pronominal and adverbial clitics are banned from appearing in absolute clause-initial position. In contexts where the verb itself occupies the initial position, such as in imperatives and polar questions, these weak elements therefore appear as enclitics on the verb, rather than in their normal preverbal position. In the following examples, the clitics are underlined.

- (149) *Pretez moi vin e viande...*
 Lend.IMP me wine and meat
 ‘Give me some wine and some meat...’ (*Eustace*, p.25 : XXII.9)

It turns out that clauses which begin with *et* or *ne* show the normal proclitic distribution; in other words, they behave as if they are not in clause-initial position:

- (150) *Et estoit li chastiax mout forz et mout bons, et l’ apeloient*
 And was the castle.NOM very strong and very good and it.CL called
cil del païs Lacoine.
 those of-the land Lacoine.
 ‘And the castle was very strong and good, and the people of the land called it Lacoine.’ (Tristan, p.53 : 36.9-10)
- (151) *Il a en une roche de mer un home de si loigtiegne terre com est*
 There has in a cliff of ocean a man.ACC of such distant land as is
Galilee, et i a cil hons demoré ja a grant tens passé ...
 Galilee and there has this man lingered already at big time passed
 ‘In a rock on the ocean there is a man from a land as far away as Galilee, and that man has already stayed there for a long time ...’ (Tristan, p.57 : 48.9-11)

Zimmermann and Kaiser discuss similar facts, but also point out that the Tobler-Mussafia Law had ceased to be rigorously observed by the thirteenth century, as it is possible to find examples from this period where the clitics do appear in absolute initial

³³To be fair, such theory-driven analyses are not uncommon in modern Germanic either, as we saw in chapter II. Many apparent cases of multiple frontings were considered to be a single, complex constituent in order to uphold the linear V2 rule, and exceptional cases of V1 are also often analysed as involving some null-element in initial position, even in cases where it is not clear exactly what that element would be if overt. Of course, the V2 status of the modern Germanic languages is generally accepted, and this makes an important difference. In Old French, using the supposed V2 rule as an infallible constituency test is arguably putting the cart in front of the horse.

position (Zimmermann and Kaiser 2010:274–275). For a similar observation regarding interrogatives, see Labelle and Hirschbühler 2005). Furthermore, they refer to Bergh (1952), according to whom clitics could attach to other conjunctions such as *ou* – ‘or’ – as well. The behaviour of the clitics is therefore no proof of the adverbial status of *et*, nor for the idea that the verb is somehow not in initial position in these cases.

Still, while the evidence is not decisive, it is suggestive. Counterexamples from other texts from the same period notwithstanding, our corpus does in fact obey the basic Tobler-Mussafia Law quite strictly. Beside the disputable cases of initial *et* or *ne*, there are for instance no cases where imperative verbs in initial position feature proclitics.³⁴ On the other hand, once an XP is fronted before the imperative verb, proclisis immediately obtains:

- (152) *Beau sire, tot ensi come li troi enfant de Babiloine furent esprové en*
 Good sir completely such as the three children of Babylon were tested in
la forneise e se proverent si bien qu’ onques ne te
 the oven and REF.CL proved so well that ever NEG.CL you.CL
renoièrent, [ensi] nos esprueve en ceste esprueve . . .
 renounced such us.CL test-IMP in this trial . . .
 ‘Good Lord, just like the children of Babylon were tested in the oven and proved
 themselves by never renouncing you, test us the same way in this trial . . .’
 (Eustace, p.41 : XXXVII.7-10)

Given that the basic mechanism of the Tobler-Mussafia Law seems to be intact in our corpus, the evidence provided by examples like (150–151) seems relevant. However, I would not interpret it as evidence that *et* is an adverb in these cases. I will assume that we are dealing with a coordinating conjunction in all cases, and that these clauses are all true V1 clauses from a linear and presumably also syntactic perspective, with the important proviso that *et* clearly plays a role in licensing them, given the preponderant evidence that main clauses with the verb in absolute initial position were shunned by language users. I would therefore like to suggest another possibility, namely that *et*-V1 clauses constitute a principled and cohesive construction living alongside the dominant V2 construction, without entering into competition with the latter on a functional level. This analysis therefore seeks to tie *et*-V1 clauses in Old French in with cases of V1 in some Germanic languages.

V1 clauses are quite common in historical Germanic (Eythórsson 1995; Hinterhölzl and Petrova 2010). Recall from chapter II, section 2.5.4 that they are still frequent in modern spoken German. Önnorfors (1997) describes at least five different kinds of V1 in use in modern German: *narrative*, *enumerative*, *deontic*, *causal* and *exclamative* V1. Pragmatically speaking, all of these serve quite different functions, a fact which should caution against assuming that V1 is only possible in (very) narrowly defined contexts in Modern German. We may therefore raise the question if it might be possible to consider the Old French V1 clauses in a similar way.

Admittedly, the similarities between the V1 clauses of our corpus and those discussed by Önnorfors are limited. The only alternative candidate would seem to be the ‘narrative V1’ type which is used to focus on the course of events, foregrounding what happened rather than providing argumentation or reflection, giving the effect of a vivid narrative with a rapid succession of events. While this description might fit in some cases, for instance if (148b) really contains three clauses, this hypothesis completely breaks down in other cases, witness

³⁴The corpus features only a few polar questions, and none of them contain clitics.

(150–151) above, as well as the examples in (153)–(154); in these passages, *et-V* clauses are not used to describe a rapid succession of events at all. In fact, they could be argued to have almost the opposite function, providing description and background, without ‘zooming in’ on the events so described. They feature the copula and static predicates rather than the dynamic predicates characteristic of narrative V1 in Germanic.

- (153) *Et neporquant tant repera li rois a la dame qu’ ele ot un enfant de*
 And nonetheless so-much returned the king to the lady that she had a child of
lui; [et] fu cil enfes masles, [et] fu apelez Cycoriades.
 him; and was this child male and was called Cycoriades.
 ‘And nonetheless the kind frequented the lady so often that she got pregnant and
 had a child by him; and the child was male and was called Cycoriades.’ (Tristan,
 p.51 : 33.2-4)
- (154) *Lonc tens li plot iluec a sejourner, [e] requist as peisanz de cele*
 Long time him.CL pleased there to sejourner and required to-the peasant of that
vile tant qu’ il le mistrent a lor chans garder, [e] fu lor
 village so-much that they him.CL put to their field defend, and was their
messiers ilueques jusqu’ a quinze anz.
 MESSIER there until to fifteen years.
 ‘For a long time he found rest there, and he made such service to the peasants that
 they put him in charge of supervising their fields, and for fifteen years he was their
 field-master.’ (Eustace, p.21 : XVII.5-8)

One could perhaps argue that cases like these still feature a succession of events, only at a higher level of narrative, since they feature verbs which are marked for perfective aspect and are chronologically ordered: ‘had a child’ - ‘the child was a boy’ - ‘the child was called Cycoriades’. These remarks do not extend to (150–151) above, however. More seriously, the construction seems to violate the most fundamental constraint on the various V1 structures discussed by Önnarfors, namely that of not having a Topic-Comment division. Several of the examples cited, for instance (151) and (153), do indeed seem to feature a postverbal topic. All in all, it is hard to pin down very exact characteristics that define all uses of this construction. What is clear, on the other hand, is that this construction shows clear clustering effects, as many *et-V* clauses follow directly after each other or with short intervals at several junctions in the text. This indicates that it performs some kind of stylistic function, a tool available to the narrator.

The formal analysis of the construction is also quite a challenge. Since we have rejected the idea that *et* is a constituent triggering inversion, there seems to be no obvious way to integrate *et-V* clauses with a V2 syntax or to make the former a subgroup of the latter. In particular, it is very hard to motivate the idea that *et-V* clauses feature a ‘null element’ of any kind in Spec-CP. They do not represent cases of topic-drop, and the lack of a temporal succession between many of the clauses conjoined by *et* makes it implausible to postulate some kind of ‘loco-temporal’ expletive in Spec-CP; in fact, many cases would become semantically incoherent if one were to add an adverb like ‘lors’ – ‘then’ – before the verb. The only plausible candidate would be a null formal ‘Platzhalter’ like *si*, but this hypothesis would again leave the absence of pronominal subjects completely unexplained, as overt *si* shows no similar ban of postverbal pronominal subjects. *Et-V* and *si-V* must therefore presumably be considered two different constructions. One would have to search for even more abstract entities such as some kind of operator; but what possible operator

could that be? The 'declarative operator'? 'The narrative operator?' It is unclear what such an operator might contribute, particularly given the quite varied discourse properties attested by the *et-V* clause.³⁵ It is true that *et-V* clauses are somehow dependent on the preceding clause in some intuitive sense, but then again this applies to every sentence in a paragraph or coherent subsection of a text or discourse. The discursive bonds that tie clauses together are supra-syntactic entities and should not be represented in syntax at all, but rather belong to the domain of text grammar.

In the light of these considerations, I will conclude that these clauses represent genuine V1 clauses with no preverbal position at all. This is the same analysis that Ötnerfors offers for V1 constructions in German. I should also to add a suggestion – although very tentatively, since the matter must be given much more detailed consideration than what can be done here – that the lack of a specifier triggers a particular *clause-typing* effect. This effect is not that of imposing a particular interpretation, as one might expect from a silent operator, but quite on the contrary to mark the clause as underspecified with respect to (syntactically encoded) information structure. There is no prefield to establish any kind of information-structural partitioning of the clause of the kind exploited in normal declaratives (topic/comment, focus-background). This clause-typing strategy is therefore employed to indicate how the clause should *not be interpreted*, and in consequence, to give free reign to a truly (that is, *non-syntactic*) pragmatic, supra-sentential interpretation of the clause based on the discursive context. This would go some way towards explaining why there is not one particular interpretation available to *et-V1* clauses, and it might also explain why there seem to be interpretive differences between superficially similar constructions in German and Old French. Of course, the actual discursive properties that might be established in Old French deserve the same kind of careful analysis as the one offered for German by Ötnerfors. Furthermore, these essentially functionalist intuitions must be embedded in an explicit, formal analysis. This is an interesting area for future research.

A (somewhat unsatisfactory) corollary of this solution is that the Tobler-Mussafia effects evinced by the construction at this stage must be considered prosodic in nature in that *et*, while not syntactically a constituent, still prevents the clitics from being counted as clause-initial at PF. As for the lack of pronominal subjects, I have nothing to offer. I see no reason to assume that these clauses are 'truncated' in the sense of being mere IPs, as suggested by Vance (1993), since the absence of pronominal subjects is not really explained by assuming that the verb only raises to I⁰. Furthermore, it would seem to predict, contrary to fact, that the construction should be available in embedded clauses, at least if one adopts a maximally simple theory of embeddability where any clause structurally smaller than a CP is embeddable. Since these constructions freely allow nominal inversion, I see no reason to assume that the verb raises to a different position in these clauses than in normal main clauses; what that position is, however, remains to be established.

3.6.2 *Ne-V* clauses

The other type of V1 clause found in our corpus is the construction that starts with the negative element *ne*. This discussion will be very brief, for two reasons. First, most of what applies to *et-V* clauses also applies to *ne-V* clauses, hence there is no need to repeat

³⁵ Another solution would be to say that there is a specifier, but that it is radically empty rather than hosting any null-element, and that the Tobler-Mussafia Law is sensitive to the presence of this specifier. This again seems too construed, and would also violate the common Minimalist idea that clausal structure projects from lexical items.

every detail from the last section. Secondly, our corpus contains only a handful of cases of *ne-V*-clauses.

As for the first point, the similarities with the *et-V* clauses include several important features, most notably the complete absence of postverbal pronominal subjects when *ne* appears in initial position of the clause. Secondly, the syntactic behaviour of *ne* could be described as inconsistent in as far as it can also appear sandwiched between an initial XP and the verb in *XP-ne-V* strings. The latter is of course the normal position of the negator, where it behaves as a preverbal clitic and in which structures postverbal pronominal subjects are freely permitted. Furthermore, initial *ne* is itself capable of hosting clitics and satisfying the TML, as clitics appear to the right of *ne*, but to the left of the verb: *ne-CL-V* (Skårup 1975; Labelle and Hirschbühler 2005:392).³⁶ In some cases, *ne* and the following pronominal clitic contract altogether, such that for instance *ne le* is written *nel*. According to Ingham, the contracted form is even much more frequent than the uncontracted forms until the early 13th century (Ingham 2014:31, see fn. 8). Note however that contraction takes place also when the negator appears clause-internally in its normal proclitic position (155), a fact which seriously undermines any claim that contraction is evidence for constituent-status of *ne* in clause-initial position:

- (155) [ll] *nel* *vost mie lessier sanz gerredon* ...
 He NEG-him.CL wanted not leave without recompense ...
 ‘He did not want to leave him without help ...’ (Eustace, p.2 : I.29)

Again, the dual patterns raise the question if we are dealing with two different and homonomous elements, or just one single item with inconsistent syntactic behaviour. Ingham (2014) opts for the former option, arguing that the conflicting properties of the simple negator receive a more satisfying explanation if we take into account its diachrony. Pointing out that the sole negator found in the earliest Old French texts is *non* and that a graphically intermediate form *nen* is encountered in the historical corpus, Ingham goes on to argue that 13th century *ne* masks two different lexical items representing two diachronically overlapping stages of evolution in Jespersen cycle of negation, more specifically a strong and a weak preverbal negator in the terminology of Zeijlstra (2004). The first is a phonologically reduced form *ne(n)*, which is a *simple clitic* in the sense of Zwicky (1977), since it must appear string-adjacent to the verb at PF, but which is otherwise still an independent constituent in the syntax, occupying a NegP situated above TP. The other item is the special clitic *ne*, which is merged in the VP and moves with the verb to its final position. The latter form is unable to satisfy the constraints of V2 or the TML.

As for the simple clitic in Spec-NegP, it is able to satisfy the TML, acting as a host for the real oblique and pronominal clitics. Adopting a criterial approach to head movement, Ingham suggests the verb moves to the corresponding head position of NegP in fulfilment of the Neg Criterion (Haegeman 1995). This is the structural correlate of *ne*-initial clauses, which do not count as V2 constructions on Ingham’s definition, since the verb in these constructions fails to reach the left-periphery. Only when the negator is focussed and moves from its in-situ position in NegP to FocP, where it changes form to the tonic *non* (cf. example (101) in section 3.3), are we dealing with verb-second construction.

This hypothesis is interesting, and Ingham adduces some suggestive quantitative data in favour of his theory, but how does it account for the lack of pronominal subjects in *ne*-initial clauses? Ingham follows Vance (1997) in assuming that *VSp*-order only arises when

³⁶Due to these conflicting properties, Foulet described *ne* as a *demi-adverbe* – a ‘semi-adverb’ (Foulet 1930:323).

the TopP or FocP is filled by a discourse-linked constituent. In the case of *ne*-initial clauses, ‘since it [the negative element in Spec-NegP – *EKP*] did not move to FocusP, post-verbal subject pronouns with initial *ne* were not licensed, unlike with *non* (Ingham 2014:34).’

It is unclear to me why PF-realization of pronouns should somehow be parasitic on a left-peripheral XP. Since Ingham assumes that pronominal subjects reside in Spec-TP in normal inversion structures, there is no structural reason why pronouns should not appear in *ne*-initial clauses as well if NegP is situated above TP, as argued by Ingham. The author further seeks to establish a more general theory of the syntax of Old French negation by expanding his analysis to include fronted NPIs like *onques* or *ja* above the negator. In such cases, Ingham suggests, we are dealing with a case of Stylistic Fronting à la Mathieu (2009), and since this fronting operation is dependent on a subject gap, this would explain the absence of pronouns in Spec-TP. We will deal with fronted NPIs in section 3.7.1.3, but for the moment it must be observed that this explanation cannot hold for simple *ne-V*-clauses, since the negator in fact appears in situ and there has been no XP-fronting at all, just Head Movement of the verb to Neg⁰ for criterial reasons.

An alternative view is to suggest that *ne-V* clauses are really just a negative version of *et-V* clauses, which were argued above to be true V1 clauses that are somehow discursively dependent on the previous clause in a different way than normal declaratives. In fact, our corpus presents a strong piece of evidence in favour of this view, since almost all *ne*-initial clauses are introduced by another element *ne*, resulting in two different *ne*'s. The first element must clearly be the conjunction *ne*, the heir of Latin *nec* and the ancestor of modern French *ni*. This can be seen in the following examples (156)–(159), where both forms of *ne* are underlined:

- (156) ... *il gisoient el mi leu del feu come il feissent en un lit*
 they lay in-the middle place of-the fire as they did.IMPF.SUBJ in a bed
*de roses ne underlinene **paroit** a chevol ne a robe qu' il*
 of roses NE NEG appeared at hairs NE at clothe that they
eussent arsüre de feu ...
 had.IMPF.SUBJ. burning of fire ...
 ‘They lay in the middle of the fireplace as if in a bed of roses and it did not appear from hair nor clothes that they had any burn marks ...’ (Eustace, p.43 : XXXVII.37-40)

- (157) *Molt s' en merveilloient tuit que si sodainement estoit adirez,*
 much REFL.CL thereof.CL marvelled all that SI suddenly was strayed,
*ne rien n' avoit leissié del suen, ne nel **pooit** l' en*
 NE thing NEG.CL had left of-the his, NE NEG.him.CL could man find.
trover.

‘Everyone marvelled that he had vanished so abruptly; he had left none of his things and they could not find him.’ (Eustace, p.15 : XII. 12-15)

- (158) *Et sachiez que a celi tens rendoient totes les regions dou monde rentes*
 and know.IMP that at that time rendered all the regions in-the word rents
*et treü a Rome. Ne n' **avoit** en celi tens encores nul crestien en Gaule*
 and fealty to Rome. NE NEG was still a that time no christian in Gaul

...
...

‘And know that in those days all the regions of the world owed rent and fealty to Rome. And there was as yet no Christians in Gaul ...’ (Tristan, p.58 : 5-7))

- (159) *Beles miracles et bels vertuz a fait ma dame Venus, qui ensi*
beautiful miracles and beautifuls deeds has done my lady Venus, who thus
l’ a retenu, ne n’ a mie sofert qu’ il s’
him.CL has retained NE NEG has NEG suffered that he REFL.CL
esloignast dou leu ...
remove.IMPF.SUBJ from-the place

‘Great miracles and great deeds my lady Venus has performed, who held him back and did not suffer him to leave the place ...’ (Tristan, p.65 : 68. 10.)

While it cannot be straightforwardly concluded that the above examples are sub-clausal coordination structures, they clearly show that the *ne*-initial clauses are strongly dependent on the immediately preceding clause. Although the corpus does not contain sufficient examples to settle the matter, it is unclear if *ne*-initial clauses are really a phenomenon distinct from *et*-initial clauses at all, since the coordinating conjunction *ne* is only a variant of *et* used in a negative context. The very few cases where a clause opens with *ne* without a preceding conjunction can be considered cases of asyndetic sentence-coordination:

- (160) *Aprés ce repaira Eustaces a son ostel, si nonça a sa fame qant que*
After this returned Eustaces to his home SI announced to his wife such as
Nostre Sires li avoit dit ... Ne demora mie grantment après que tote
Our Lord him.CL had said ... NE lasted NEG greatly after that all
lor mesniee chaïrent en une grant enfermeté ...
their household fell in a great disease ...

“Afterwards Eustaces returned home and told his wife of all the things the Lord had told him ... And it did not last long until all of his household was taken ill with a horrible disease ...” (*Eustaces*, XI.1-8)

I therefore tentatively suggest that there is only one kind of *V1* clause in thirteenth century OF, which is a positive (*et-V*) or negative (*ne-V*) declarative clause which is discursively strongly dependent on its immediately preceding clause, and which formally lacks a specifier position.

3.7 Linear V3 strings; CSVX, CCVX, CCVSpX, SCVX

The verb-third strings are very relevant to the debate on Old French V2 (see Prévost 2001 for a discussion) and have been singled out by many researchers since (Kaiser 2002) as the prime evidence against V-to-C movement in Old French. With an occurrence of slightly above 14% of all main clauses in both texts of the corpus, they are quite robust from a quantitative point of view, and we must therefore assume that these strings will have been very salient in the input to the children acquiring the language. The crucial question is whether these are exceptions that can be captured by some kind of generalization, such as lexical triggers or specific constructions, or if they rather eschew all attempts at a systematic explanation and appear as a free and productive alternatives to linear verb-second.

3.7.1 Lexically triggered V3

We will start by considering possible lexical triggers of linear V3. It is well known in the literature on Old French that certain adverbs tended not to trigger subject-verb inversion. Foulet identified *neporquant/neporec* – ‘nonetheless –’, *onques* – ‘(n)ever’ – *sanz faille* – ‘undoubtedly, assuredly, certainly’ – and *certes* – ‘certainly’ – as typical examples of expressions failing to trigger inversion (Foulet 1930:311), and Vance (1997:61-66) added the adverbs *ja* – ‘already’ – and *jamaïs* – ‘never’ to Foulet’s list, as well as the interjection-like oaths *par foi* – ‘by faith’ – *par mon chief* – ‘by my head’ – and *por Dieu* – ‘for God’s sake’. Many of these expressions are encountered in our corpus, in particular in *Tristan*, and in this section we will undertake a discussion of their characteristics and how they can be integrated in the general clausal structure of late Old French. We will divide them into three different groups, which all have in common that they regularly involve linear V3, but which otherwise show significant internal differences. The first group to be considered consists of *neporquant* and *certes*, which we argue to be discourse adverbs related to the expression of illocutionary force. The second group consists of *sanz faille* and *sans doute*, which in spite of having some features in common with the previous group will be considered a type of parenthetical interjection. The third and final group consists of *onques* and *ja*, which are Negative Polarity Items. The latter are not only occasionally involved in linear V3, but show some additional properties which make their precise analysis very elusive. The following discussion will rely substantially on previous research, since the corpus does not contain sufficient examples of the relevant expressions to draw firm conclusions. The focus will be on the relevance of these expressions for the general syntax and for the issues which are of central concern to this thesis. While a possible analysis will be suggested in each case, the following discussion therefore cannot be close to exhaustive.

3.7.1.1 *Neporquant* and *certes*

Neporec does not appear in our corpus, and *neporquant* is only encountered in *Tristan*, where it is employed in total 6 times, always in initial position and always involving linear V3. It can be followed by subject-initial *CSVX* (161) or inverted *CVSX* (162) strings alike, suggesting it does not interact with the syntax of the following clause.

- (161) *Et neporquant [il] dit a soi meïsmes . . .*
and yet he says to himself self

‘And yet he says to himself . . .’ (*Tristan*, p.58 : 49.15)

- (162) *Et neporquant [totevoies] revint ele en sa memoire et comença a mangier*
and yet still returns she en her memory and starts to eat
et a esforcier soi
and to reinforce herself

‘And yet her memory still comes back to her and she starts to eat and to regain strength.’ (*Tristan*, p.42 : 6.3-4)

As for *certes*, it too only appears in *Tristan* and is used in total five times. It behaves in a similar way to *neporquant*, apart from the fact that it seems to preferentially appear in the context of direct discourse; in the corpus, the word appears each time as the first word of a clause in direct speech, and in four of these cases, it is followed directly by an intercalated

clause featuring inversion and a *verbum dicendi*. The clause that follows this intercalated clause can be either subject-initial (163) or an inversion structure [164]:

- (163) *Certes, fait li rois, [ele] n' est pas saige qui tel duel demoine*
 certainly makes the king she NEG.CL is not wise who such pain lead
 'Truly, says the king, she is not wise to show this grief.' (*Tristan*, 18, 20-21.)
- (164) *Certes, dit li preudons, [mout] vos a Diex bien aidie*
 certainly says the prud'homme much you.CL has God well helped
 'Truly, says the gentlemen, God has in truth aided you greatly.' (*Tristan*, p-50 : 30.2-3)"

The behaviour of these adverbials led Foulet to conclude that they were 'without influence' (Foulet 1930:311) on the rest of the clause and Skårup to place them in a *zone annexee*, preceding the *zone préverbale* (Skårup 1975). The great 19th century philologist Burguy even categorized *neporquant/neporec* as conjunctions (Burguy 1869a:385-386). This may seem at odds with a modern understanding of the distinction between adverbs and conjunctions, since the latter generally expresses a relation between clauses, without performing a semantic role within the clause itself. It is clearly felt that *neporquant* contributes semantically. One might also object that it is possible to find clauses introduced by *neporquant* that are not conjoined at all with other clauses, but this argument is not necessarily decisive, as the same applies to *et* and even *quar*. Given the diachronic tendency for adverbs to grammaticalise into conjunctions (Ramat and Mauri 2011), we cannot exclude that *neporquant* is a kind of borderline case. In either case, the correct generalization is that the finite verb invariably follows the word that follows *neporquant* or *certes*.

Rather than adopting Burguy's suggestion that *neporquant* is a conjunction, which would anyway not extend to *certes*, I suggest that these are phrases, and more specifically clause-external adverbs generated in the left periphery of the clause to express a speaker-oriented, discourse-related semantics that takes scope over the entire clause. This hypothesis finds support in the fact that neither of these expressions appear clause-internally. Both seem to be closely related to the encoding of the speech act, although in completely opposite directions, as *neporquant* provides a concessive tone, whereas *certes* rather enforces and insists on the veracity of the following claim. This idea is further corroborated by the observation that they only appear in main clauses, a finding which mirrors that of Vance (1997, p.62). Both *neporquant* and *certes* show clear similarities with certain adverbs from modern Germanic; *neporquant* is comparable to 'nevertheless' in English (as pointed out by Wolfe (2015:94), *trotzdem* in German or *likefullt* in Norwegian. These can also be left-dislocated, and furthermore, in the V2 languages the following clause may be either subject-initial (165) or inverted (166):

- (165) [*Trotzdem*], [*ich*] **habe** meine Zweifel.
 nonetheless I have my doubts
 'Nonetheless/Still, I have my doubts.' (German)
- (166) [*Likefullt*], [*så mye tid*] **har** jeg ikke.
 nonetheless so much time have I not
 'Nonetheless/Still, I don't have that much time.' (Norwegian)

Certes also finds clear parallels in modern Germanic in interjection-like or speech-act oriented adverbs like *klar* or (more formal) *wahrhaftig* in German or *sannelig* in Norwegian, which may also be dislocated and precede both subject-initial (168) and inverted (167) clauses.

- (167) [*Klar*], [*das*] **habe** *ich* *auch gemacht*.
of-course, have I also done
‘Of course, I did so, too.’ (German)
- (168) [*Sannelig*], [*han*] **kaster** *ikke bort tiden!*
in-truth he throws not away time
‘He really doesn’t waste any time!’ (Norwegian)

Furthermore, the very same type of adverbials regularly fail to trigger inversion in historical stages of Germanic as well; according to Cichosz (2017), the adverbs *witodlice*, *soðlice*, *efne* – ≈ ‘truly, indeed’ – almost consistently fail to trigger inversion in Old English, ‘due to their extracausal status’ (Cichosz 2017:317).

- (169) *Soðlice* [*Dauid se witega*] **spræc** *to Drihtne*
truly David the prophet spoke to Lord
‘Truly, David the prophet spoke to the Lord’ (Old English, taken from Cichosz 2017:317.)

It should be clear from that the non-inverting character of *neporquant* and *certes* is hardly a convincing argument against the V2 hypothesis for Old French. The variable word order that may follow *neporquant* or *certes* should not be interpreted to mean that these adverbials as such are involved in free and productive word order variation. Rather, the adverbials are external and invisible to the computation of the word order of the clause they initiate, such that the following clause can be subject-initial or inverted according to the preference of the speaker and the pragmatics of the context, much like equivalent expressions in modern Germanic³⁷. Concretely, then, I suggest that they are base-generated in the left-periphery, possibly a very high position (although evidence for this position can only be obtained through word order facts based on Transitivity, cf. the SSAP) although a more concrete proposal must wait until we have achieved a clearer picture of the syntax of Old French.

3.7.1.2 *Por Dieu, sans faille/sans doute*

In *Tristan*, we find some examples of the interjection *por Dieu*, which is invariably used in direct discourse and is either clause-initial or clause-final. It is strongly associated with a verb in the imperative mood and is therefore not relevant at all to our discussion here.³⁸

More interesting are the adverbials *sanz faille*, *sans doute*. These show some semantic affinity with *certes* to the extent that they seem to reinforce the speech act, an observation which is echoed in Ingham, (2005:105) but they have a much more flexible syntactic

³⁷Admittedly, the modern Germanic adverbs discussed here can also appear clause-internally and also quite regularly trigger inversion. In this sense, *neporquant* and *certes* behave somewhat differently in that they consistently feature the extracausal pattern, which is just one of several options in modern Germanic.

³⁸Salvesen (2013:147) reports examples in declarative clauses and considers them scene setting elements. As interjections, one might also suggest that they are completely external to the clause, and even to the left periphery of the clause.

distribution. In clause-internal, that is *postverbal*, position they are not directly relevant to our concerns, but they can also appear in two other positions which both involve linear V3. First, they can appear clause-initially as in (170). In such cases, they scope over the entire sentence and perform much the same function as *certes*:

- (170) *Sans faille, [ce] faisoit l' eve qui estoit roide et forz a merveilles.*
 without failure this did the water which was rigid and strong to marvels.
 'Doubtlessly, this was caused by the current which was rapid and marvelously strong.'
 (Tristan, p.55 : 42.5-6)

Interestingly, these adverbs can also crop up in another position which also involves linear V3, namely directly after the first constituent of the clause. This is the same distributional pattern as so-called 'V3 adverbs' in modern Scandinavian V2 languages (cf. section 2.5.1). This occurs both with subject-initial clauses (171) and with inversion structures (172):

- (171) *[Cil de Cornoaïlle] sanz doute avoient mise lor dame en une tor en prison.*
 those of Cornwall without doubt had put their lady in a tower en prison.
 'The men of Cornwall had doubtlessly put their lady in prison in a tower.' (Tristan, p. 57 : 47.2-3)
- (172) *[Lor chambellan] sanz faille avoient il trové mort a la rive.*
 their chamberlain without failure had they found dead at the bank.
 'Truly, they had found their chamberlain dead at the shore.' (Tristan, p. 57: 47.6)

These examples cannot be dismissed by generating them outside of the clause proper in the left periphery, since they are not even in initial position, but are rather wedged in between the first constituent and the finite verb. We might reasonably ask if they constitute empirical evidence against V-to-C movement for the child acquiring the language. If *sans doute/sans faille* are taken to be sentential adverbs modifying the following extended verbal projection, this analysis is not entirely implausible. However, there are several reasons not to adopt this solution. First, it would be highly surprising that only a very limited class of lexically determined adverbs should be able to perform this clause-modifying function. Secondly, given their status as interjections, it does not seem unreasonable to consider them parentheticals which are added for emphasis. Strictly speaking, then, we are not really dealing with true V3 at all in these cases.³⁹

³⁹There is also another possibility, which is suggested by a peculiar feature of these expressions in the corpus, namely the fact that they only appear after initial constituents which are full DPs, not pronouns. This apparent interaction with the categorial status of the preceding constituent would be surprising if they scope rightwards over the verbal projection. One might therefore speculate that *sans doute/sans faille*, apart from their capacity to perform sentential-wide speech act modification as in (170), can attach to the XP on their immediate left for constituency scope, presumably by right-adjunction. Note also that *sans faille*, but not *sans doute*, has in fact kept this left-attachment property in modern French:

- (i) *Elle peut compter sur le soutien sanz faille de sa mère.*
 She can count on the support without failure of her mother.
 'She can count on the unyielding support of her mother.'

This use of *sans faille* is clearly more restricted in modern French than what was the case in OF, since the modern language only seems to allow it with deverbal DPs, suggesting the expression currently straddles the border between adjective and adverb, whereas it is clearly more interjection-like in the medieval stage.

This hypothesis receives some support from examples like (173), where *sanz faille* appears in an unexpected position directly after a past participle, and where the context excludes an interpretation with the following PP:

- (173) *Un soer quant je me gisoie en mon lit, vint avanz une voiz qui*
 an evening when I me.CL was-lying in my bed came forward a voice that
me dist: 'Creature vil et orde, lesse ton pechié! Je fui espeoentez mout
 me.CL said creature vile and impure leave your sin I was scared very
durement et convertiz sanz faille par cele parole, que je conui bien que
 thoroughly and converted without failure by that word that I knew well that
verité me disoit la voiz. . . .
 truth me.CL said the voice
 'One evening when I was lying in my bed a voice came to me, saying: 'Vile and
 impure creature, leave your sins behind!' I was very scared and altered, no doubt,
 by those words, for I knew well that the voice was telling me the truth. '
 (Tristan, p.50 : 29.9-12)

While the syntax of these expressions is interesting and merits further attention, there is little reason to attach much importance, in the context of the V2 hypothesis, to the fact that they can separate the subject and the verb, since they quite generally show a very flexible distribution (see also Ingham (2005:106) for an example where *sans faille* turns up in another clause-internal and somewhat odd position).

3.7.1.3 *onques, ja*

The final group of adverbs to be considered is *onques* and *ja*. These are by far the most frequent of the classes of adverbs involved in linear non-V2 order, and also by far the most complicated from a theoretical point of view. The following discussion draws heavily on the conclusions reached by Ingham, who has conducted the most detailed investigations of the syntax of Old French negation and Negative Polarity Items (Ingham 2005, 2007, 2013, 2014).

The first thing to notice about these NPIs in Old French is that they are 'symmetric', in the sense that they may both precede and follow the negator itself, yielding both the strings *ne... onques/ja* and *onques/ja...ne*. In this section, only the latter case will be discussed, as this is the constellation that involves linear V3. These fronted NPIs are either:

(a) followed directly by a nominal constituent (subject or non-subject) and then the verb, resulting in linear V3.

or

(b) followed directly by the verb and then no overt subject, resulting in the subject-less linear V2 string CVX.

While this always applies to *onques* as an inherent NPI, the situation is more complex for the adverb *ja* (derived from Latin *iam* – 'now'), which is strongly polysemous and can appear in affirmative and negative clauses alike. When used as a fronted NPI in negative declaratives, *ja* assumes exactly the same properties as *onques*. In affirmative clauses, it

Still, it does not seem unlikely that there is some continuity involved here. Since there are not enough examples in our corpus to test if this generalisation really holds, I will not pursue this option further here.

can function as a kind of underspecified temporal adverbial that interacts with the tense and aspect of the verb *ja* in complex ways to produce subtle semantic effects, as discussed by Buridant (Buridant 2000:523-524). It can also sometimes take on the role of a discourse particle. In both of these cases it can appear in initial position and trigger regular inversion with postverbal DPs and pronominal subjects alike; the adverb is illustrated in (174) and the particle in (175):

- (174) *e [ja] fu li cers mult esloigniez de tote la compaignie ...*
 and already was the deer much removed of all the compaigny ...
 ‘And already the stag was far ahead of all the riders ...’ (Eustace, p.4 : II.19-20)
- (175) *-Coment puet il vivre?- fait li rois; -[ja] fu il gitez en la mer!*
 how can he live did the king PRT was he thrown in the sea
 ‘How can he be alive? said the king, he was thrown into the ocean!’ (Tristan, p.47 : 22.8)

The subtle contrast between (a) and (b) described above has had the effect of leading both philologists and linguists astray. Foulet only gave examples of the pattern (a) above, concluding that *onques* had no influence on the syntax of the clause, in the same way as *neporquant*, *certes*, *sanz faille* (Foulet 1930:311). Vance noticed examples of the latter pattern (b) as well, concluding that *onques* belongs to an ‘unstable’ group of elements which sometimes trigger inversion, sometimes not (Vance 1997:62-62). This conclusion seems altogether natural as first, given the contrast in (a-b) above, and indeed even required, since Vance generally adopts the analysis of Adams (1987), according to which null-subjects are only licensed in postverbal position (*Foulet’s generalization*).

However, it was discovered by Price that *onques* displays a further particularity, namely the fact that it *hardly ever* features a postverbal pronominal subject (Price 1966, 1973). This claim receives support from our corpus, since no example with *onques* or *ja* in initial position of a negative clause features a postverbal pronominal subject. Given the high frequency of the string *CVSpX* in our corpus, also with initial adverbs of various kinds, this is completely unexpected. Furthermore, Ingham (2005) presents data from two early thirteenth century prose romances, claiming initial *onques* and NPI *ja* do not appear with preverbal pronominal subjects either, a claim which also holds for our corpus.

This hypothesis is considerably strengthened in a more recent and quantitatively more robust study (Ingham 2013), in which four new prose romances from the same period were analysed, yielding the same conclusion: pronominal subjects are completely banned from appearing when *onques* or NPI *ja* is in initial position of the clause. Ingham reports a single example of the string *onques-V-Sp*, raising the question if the example is native or the result of some transmission error (Ingham 2013:275-276). The string *ja-V-Sp* appeared four times in negative clauses, but as Ingham discusses, it is unclear that these cases are NPIs, since they may equally well be interpreted as cases of the discourse adverb *ja* appearing in a negative context (Ingham 2013:269-270, cf. 175) above. This makes all the difference, since it is clear that is not the negative polarity of the clause itself, but the appearance of an NPI in initial position that triggers this unusual syntactic behaviour, or to be even more precise: *the appearance of an initial adjunct NPI in initial position*, since Ingham argues that there is evidence that argument NPIs behaved differently, triggering regular V2 inversion structures with pronominal subjects when in initial position:

- (176) *[Nul si bon seignor] ne poriez vos servir*
 Any so good lord NEG.CL could you serve

glt ‘You could not serve any lord as good as this one’ (Le Haut Livre du Graal 616, from Ingham 2013:271.)

The findings of Ingham are quantitatively too robust to leave any doubt that we are dealing with a real phenomenon of Old French syntax here, and I can only add that our corpus behaves in exactly the same way. Furthermore, Ingham shows that evidence from a selection of 12th century verse texts clearly demonstrates that the same phenomenon is equally robust in verse. Suggesting that the syntax of verse might be conservative, he speculates that *onques* and *ja* conserve syntactic patterns of a diachronically older stage of the language. It is relevant in this respect that *onques* and *ja* are among the very few Old French adverbs to derive directly from Latin etyma (*onques* < *unquam* and *ja* < *iam*), the others generally being of later Romance creation (Herman 1963).

In fact, Ingham has argued elsewhere that there might be continuity from Late Latin in the word order distribution of these adverbs. On the basis of a comparison between selected texts from Classical Latin and the Late Latin period, Ingham (2007) demonstrates that, in the latter but not the former, the placement of *iam* and *(n)unquam* with respect to the verb is sensitive to the polarity of the clause. In negative declaratives in later Latin, *iam* and *(n)unquam* were significantly more likely to be fronted to a preverbal, and generally clause-initial, position. Ingham’s suggestion is that this pattern with *onques* et *ja* in preverbal, clause-initial position was already firmly established before verb-second became productive in Old French, and since this pattern in fact corresponded to the emerging V2 constraint (in terms of linear order), it was somehow left as it was. This included the ban on overt pronominal subjects, which were still not grammatical in the Late Latin period (Ingham 2007:72-73).

If this suggestion is on the right track, some properties of Late Latin syntax, notably the ban on overt pronominal subjects, ‘hitchhiked’ into Old French on the back of these NPIs. This is an interesting suggestion, but it still does not tell us how children acquiring the language would analyse such structures in the thirteenth century. A diachronic explanation cannot replace a synchronic explanation, although it can potentially account for some idiosyncrasies in the input to the child. We have already seen what these idiosyncrasies involve, now the challenge is to find out what the children made of them, and in particular to what extent they were integrated into the rest of the grammar.

Ingham (2013) has also provided a concrete answer to this question, couched in a Minimalist, phase-based approach to NPI licensing. Drawing on the insights of Martins (2000), he suggests that NPIs like *onques*, *ja*, *nul* lacked inherent negative features, but rather carried an uninterpretable, non-assertive polarity feature (Ingham 2013:272) that needed checking. The relevant head able to check this feature on the NPIs is the negator *ne*. Since there is no asymmetry in OF with respect to the linear relationship between the NPI and the negator, meaning *ne...onques/ja* and *onques/ja...ne* are equally fine, Ingham suggests that the negator is able to check the feature *[-pol]* in both Head-Complement and Spec-Head relations, adding that ‘the latter includes not only the clause subject but also adjunct constituents left-adjoined to TP’ (Ingham 2013:273). This latter addition is a rather idiosyncratic interpretation of what falls under a Spec-Head configuration, but it is needed in Ingham’s analysis to account for the grammaticality of linear V3 cases with nominal subjects, in other words cases where *onques/ja* are followed by another constituent before the verb and hence cannot be in specifier position of the verbal projection.

Ingham goes on to suggest that the different behaviour of argument NPIs like *nul* and adjunct NPIs like *unques/ja* can be captured by a phase-based derivation. Since arguments,

whether subjects or objects, are merged in the VP-complex, they can be checked by the negator *ne* before the verb (and the negator) raise to T. This must entail that *ne* itself is already added in the VP, then.⁴⁰ On Ingham's account, this explains why argument NPIs may participate in cyclic movement to higher, left-peripheral projections, since they may also carry criterial features like *Wh-* or *Topic*. In other words, argument NPIs may participate in the fulfilment of the verb-second constraint, moving to a specifier position in the left periphery above the verb in C after the checking of the *[-pol]* feature in the low vP phase.

The situation is different for adjunct NPIs like *onques* and *ja*, since these as T-related adverbs are merged after the vP phase. In addition, since they cannot be topicalised, they will only reach as high as a position adjoined to TP, where they will be checked by the higher phase head C. In such structures, CP will be non-overt. In Ingham's view, this explains why we never find any pronominal subjects in these constructions:

‘In Old French Spron was either in Spec CP or was a clitic on C. Since the C system was not engaged in *ja/onques*-initial clauses, the absence of Spron follows automatically.’ (Ingham 2013:275)

I believe this solution raises just as many questions as it seeks to answer. If there is no null-subject because there is no CP-projection, then where and what is the subject of the clause? There must quite simply be a null-subject in the clause, otherwise the verb would not be able to release/check all its Theta-roles. Ingham does not explicitly address this problem, but if we for the sake of argument accept that clauses initiated by *onques/ja* are somehow truncated and do not involve a CP, perhaps we could suggest that subject clitics are only PF-readable at the CP-level. In that case, the pronominal subjects that only cliticize to the verb in a lower position, such as I⁰, must remain silent. This would seem to get the facts right, but the solution is admittedly very ad hoc, lacking any theoretical or independent empirical justification. Notice also that Ingham's truncation hypothesis rests fundamentally on the assumption of a general verb-second grammar in OF, since whoever accepts this hypothesis has thereby seemingly accepted that *CVSpX* strings are the output of a verb-second grammar with the verb in C⁰. Let us therefore raise the question of there is any reason to assume that these NPI-initial clauses are TPs/IPs rather than CPs. However, this question cannot be answered yet, since we are still very much in the process of establishing the correct structural description of main clauses, and both V-to-I and V-to-C parses are still candidates. We must therefore rather ask if there is any reason to assume that *onques/ja*-initial clauses feature movement to a lower projection than other main clauses.

The String-Structure-Assignment-Principle (SSAP) states that children try to account for the global input in a maximally economic way; we must therefore assume that they will have tried to integrate *onques/ja*-initial clauses as far as possible into the general pattern of their emerging I-grammars. As already stated, I do not believe that the absence of pronominal subjects is strong evidence in favour of a truncated structure. Leaving aside the subjectless CVX-strings, we therefore turn to option (a), where nominal subjects occur. These strings involve linear V3. Now, if these strings had been consistently of the

⁴⁰Some details of Ingham's analysis are not entirely clear to me, for instance the exact position of *ne* inside the VP, such that it allows the checking of the *[-pol]* feature on the NPI subject or object. Ingham adds in a footnote that he leaves open ‘the question of whether a NegP was projected in Old French’ (p.274). Be that as it may, *ne* must still occupy a head position in the VP that gives rise to the required Head-Complement or (expanded) Head-Specifier configuration. From Ingham's bracketed notations (p.273), the intended head seems to be *v*⁰, which would also follow from the general idea that *v*⁰ is a phase head.

non-inverted type CSVX, this would indeed have been very strong evidence in favour of Ingham's claim. However, the fact remains that *onques/ja*-initial clauses regularly feature overt inversion strings with nominal subjects, witness the following example:

- (177) *Onques* [a paroles que il deüssient de li leanz] ne **respondi** Sador
 ever at words that they said of him there NEG.CL answered Sador
 'Sador never answered to the words they spoke of him.' (Tristan, p.64 : 66.9-10)

Ingham mentions these inversion strings, but concludes that they may be analysed with the subject in Spec-VP (Ingham 2013:270). This is in line with previous observations already reviewed in section 3.5.2 that nominal subjects occupy a low position in Old French inversion strings (Lemieux and Dupuis 1995; Vance 1997; Rinke and Meisel 2009; Salvesen and Bech 2014). Since Ingham assumes a V-to-C analysis of pronominal inversion structures, this means that Ingham posits two different positions for the finite verb in main clauses.

I will assume instead that children try getting away with a single position, and that the absence of pronominal subjects is not analysed as the result of any particular behaviour of the verb, but rather as an idiosyncratic property of the construction itself, some kind of blocking effect that rules out the co-occurrence of initial *onques/ja* and pronominal subjects. It may well be correct, as Ingham suggests, that this pattern is a relic from Late Latin. If it is correct that that Latin/early Romance predecessors of *onques* and *ja* at some point became obligatorily fronted to a sentence-initial position in negative clauses, we might speculate that this obligatoriness, being at odds with the general productivity of the prefield to host different fronted elements, somehow made the whole construction some kind of semi-idiomatic 'island'. At this given stage of evolution, pronominal subjects had to remain unpronounced, since the language was still a Consistent Null-Subject language in the terminology of Roberts and Holmberg (2010). We may even suppose that the finite verb in Latin carried interpretable inflectional morphology, making the overt expression of a weak pronominal subject not only redundant, but incoherent, as the external theta role of the verb had already been checked by the verbal morphology.

This means that for every successive generation of children acquiring the language, subjects were not to be heard in the PLD in the context of this island, even when they started appearing elsewhere, possibly as a consequence of the gradual weakening of the inflectional morphology. Now, it would be senseless to say that the language had conserved an 'island' of the old grammar with its interpretable inflectional morphology, since the verb would carry no stronger inflectional morphology in these clauses than in other contexts. Plausibly, children might have reinterpreted this ban on pronominal subjects as a property of the construction itself, adding a corresponding blocking rule to the lexical entry. Of course, this solution entails that the *onques/ja*-initial clauses are 'in the lexicon', that is, that they must be considered constructions. While it would admittedly be preferable to derive every single clause from the application of basic compositional operations of a uniform kind, languages do exhibit idiosyncratic properties that are not easily integrated into the core grammar, what is often referred to as a the *periphery*. The NPI-initial clauses would be a prime example of the latter, a case of 'historical residues [...] which we can hardly expect to – and indeed would not want to – incorporate within a principled theory of UG', to borrow a frequently cited passage from Chomsky (1981:8–9).

Before rounding off this discussion about *onques/ja*, it is worth pointing out that the lack of pronominal subjects is not the only particular feature of these items, or rather this construction, as we have suggested (since *onques/ja* do not exhibit any of these particular

features when used clause-internally, in the *ne...onques/ja*-pattern). As Ingham (2013:270–271) observes, *onques/ja*-initial clauses also tend to involve multiple constituents in front of the verb to a much higher degree than what is normally observed in Old French. Thus, alongside the general V3 pattern *onques/ja*-XP-verb, there are also some cases that involve linear V4. This claim also finds some support in our corpus, as a couple of the extremely few cases of linear V4 in main clauses do in fact involve initial *ja(mes)*:

- (178) ... *quar jamés [en ton ostel] [deus si proudomes] n' entrerunt.*
 for ever en your house two so prudhommes NEG.CL entered
 '...for never did two so valient *prudhommes* enter your home.' (Tristan, p. 62: 59.13)

In this and other similar cases reported by Ingham, it is tempting to speculate that *onques/ja* are able to attach to other constituents syntactically. It is worth mentioning that *jamais* seems to retain this property even in modern French. Apart from fixed expressions like *jamais de ma vie*, we also find cases where the host constituent is a locative PP instead of a temporal, as in (179):

- (179) *Jamais en France [un tel hommage] n' a été rendu à un chanteur.*
 never in France a such tribute NEG has been rendered to a singer
 'Never in France has such a tribute been rendered to a singer.'
 (From *Liberation* online, published 09.12.2017)⁴¹

As we already saw in chapter 2, section 2.5.3, parallel expressions are found in the modern Germanic languages as well (cf. the English translation of (179)), where 'never'-adverbs like *nie(mals)* in German or *aldri* in Mainland Scandinavian, although not NPIs, behave like focus adverbs/particles in attaching to other constituents :

- (76) [*Nie zuvor in Deutschland*] *hat sich jemand für eine Fernsehserie so*
 never before in Germany has REFL-CL some for a television-series so
kopfüber in die Vergangenheit gestürzt...
 headlong in the past plunged.
 'Never before in Germany has anyone dived so headlong into the past because of a television series.'
 (German, from *die Welt* online, 10.10.2017)⁴²

It is therefore not entirely implausible that some apparent cases of V4 involve complex constituents and should rather be counted as V3, and by extension, that some cases of V3 should be counted as V2. However, this analysis may hardly be extended to account for all cases (cf. (177) above; see also Ingham 2013:270-271). There is no way escaping that there is more than one constituent in front of the verb in (180):

- (180) *Voire, fet li autres, onques [mielz] [nus hom] ne resembra autre.*
 true makes the other ever better no man NEG resembled other
 'True, says the other, never did any man more resemble another man.' (Eustace, p.26 : XXIII.4-5)

⁴¹<http://www.liberation-champagne.fr/50145/article/2017-12-09/bien-elu-mais-mal-aime>

⁴²<https://www.welt.de/kultur/plus169444371>

I will therefore conclude that *onques/ja*-initial clauses are not the output of a V2 grammar, but a case apart, possibly a left-over from a diachronically earlier system, that was internalised as an idiosyncratic construction with special properties, notably a ban on pronominal subject and more lax conditions on the prefield. On the other hand, the inversion facts, which are otherwise identical to the general nominal inversion patterns of the language, suggest that we lose more than we gain by parsing these clauses into ‘truncated’ structures. This solution does not anyway provide a natural explanation of the absence of pronominal subjects, while clearly complicating the grammar to be acquired by postulating two different positions for the finite verb.

Summary In this section, we have seen that many cases of linear V3 involve a very limited class of adverbs or adverb-like expressions. Within this class, it is possible to discern at least three different groups that differ minimally from each other with respect to syntactic distribution and other properties such as their capacity to co-occur with pronominal subjects. While it is not possible to provide a definite syntactic analysis of these constructions without first establishing the general syntactic structure of main clauses – an analysis which is still pending – I have suggested a categorial status and partial analysis for each group, and concluded that neither the illocutionary force adverbs *neporquant* or *certes*, nor the interjection-like parenthetical adverbs *sans faille/sans doute* provide evidence against V-to-C movement. As for the fronted Negative Polarity Items *onques* and *ja*, the situation is more complicated, and we have tentatively accepted (parts of) the analysis proposed by Ingham that these items are remnants of a former non-V2 negative syntax. I have argued that the latter can plausibly be considered *lexical* in the sense that it is not possible to deduce general, productive rules from their syntactic behaviour; in consequence, they do not arise through any violation of the general rule of the language to put the verb in second position.

3.7.2 Left-dislocation (LD) structures

Another source of linear V3 in the corpus is provided by constructions featuring left-dislocated (LD) phrases which occur clause-initially, followed by another constituent and then the verb in linear third position. Recall from chapter 2 that such constructions are possible in all the Germanic V2 languages on the condition that the initial constituent be linked to a resumptive element inside the clause (cf. section 2.3.1). This is often formalised by co-indexing the LD constituent and the resumptive. In the following, ‘LD’ is used generically to designate any kind of left-dislocated phrase; in section (3.7.2.1), the precise type of LD found in the corpus is briefly discussed.

While such constructions are not very numerous in our corpus, they behave just like the corresponding constructions in Germanic. They show a certain tendency to appear in contexts where the initial DP is ‘heavy’, either through substantive modification or when interrupted by a parenthetical clause. In this respect, the discourse factors governing these constructions (see Marchello-Nizia 1998 for some discussion) seem to be different from the very frequent left-dislocation structures found in modern spoken French (De Cat 2009). In the following examples, clauses which are considered parenthetical are enclosed in parentheses :

- (181) [*Et la demoisele_i*], (*qui joene estoit et novelement mariee*), (*quant ele sot*
and the damsel who young was and recently married when she knew
son mari pres de li, et celi ocis qui illec l’ avoit amenee),
her husband close to her, and that-one killed who there her.CL had brought,

[*ele_i*] *s'en* ***fui*** *pour sauver son cors.*
 she REFL.CL therefrom went to save her body.

‘And when the lady, who was young and recently married, recognized her husband, and saw that the man who had brought her there had been killed, fled from there to save her life.’ (Tristan, p.65: 67.20-22)

- (182) [*Naburzadan_i*], (*quant il vit que ses freres n' estoit mie venuz*), [*il_i*] *en*
 Naburzadan, when he saw that his brother NEG was not come, he thereof
fu *auques liez* ...
 was quite happy
 ‘When Naburzadan saw that his brother had not come, he was very happy ...’
 (Tristan, p.42 : 9.1-2)

- (183) *Et [li rois_i], (qui par la men la tenoit,) (maintenant qu' il trova l'*
 and the king who by the hand her.CL held now that he found the
anel), [il_i] li osta et le mist en son doit.
 ring he her.CL removed and it.CL put on his finger.
 ‘And not that the king, who was holding her hand, found the ring, he took it off her
 and put it on his own finger.’ (Tristan, p.54 : 39.4-5)

- (184) *Et [li chevalier_i, qui regardent le roi,] (quant il s'aperçoevent qu' il n'*
 and the knights who watch the king when they perceived that he NEG
est mie morz), [il_i] dient: Comment! Encores est il vis!
 is not dead they say how still is he alive
 ‘And when the knights, who were watching the king, realized that he was not dead,
 they said: What? He is still alive!’ (Tristan, p.56: 44.18-20)

We have already seen that non-subject arguments that are fronted to the prefield in linear V2 strings do not provoke clitic-doubling (section 3.3), unlike the situation in Modern French (Rowlett 2007:178–180, De Cat 2009) or indeed modern Romance in general, where such constructions generally go by the name *clitic left-dislocation* (CLLD). It is therefore highly interesting that LD arguments do in fact trigger such doubling (185). If the left-dislocated phrase corresponds to the object rather than the subject, the resumptive does not occupy the prefield (185), but this is of course just a natural consequence of the clitic status of object pronouns:

- (185) [*Tote la terre que li barbarin avoient saisie_i*], [*il*] *la_i delivra* ...
 and the land that the barbarians had seized he it.CL freed
 “All the lands that the barbarians had seized, he freed them.” (Eustace, p.30: XXVII.
 3-4)

The examples in (181)–(185) reveal a clear similarity between Old French and the modern Germanic V2 languages in the syntax of left-dislocation. In this sense, the V3 strings featuring LDs are not only compatible with a V2 grammar, but in fact provide quite suggestive evidence in favour of it.

3.7.2.1 Hanging Topics (HT) or Contrastive Left Dislocations (CLD)?

It is natural at this point to ask what kind of left-dislocated constituent we are dealing with in the cases above, or whether there is only one kind. In modern Germanic, a distinction is

made between two different kinds of LD constituents (cf. section 2.4.4.1). Hanging Topic Left Dislocations (HTLDs) are generally taken to be base-generated in the left-periphery. An HTLD is an NP/DP that carries an aboutness topic interpretation and generally does not display connectivity effects into the clause (Riemsdijk 1997; Frey 2004a; Grewendorf 2008, 2009). Its most salient feature is that it generally carries default nominative case, regardless of whether the resumptive acts as the subject of the ensuing clause or not. For this reason, it is sometimes also called *nominativus pendens*, since there might be a mismatch in case features between the HTLD and the resumptive. It is often assumed to be unembeddable (Grewendorf 2009:69).

The other and more frequent LD construction is known as Contrastive Left Dislocation (CLD). The term goes back to Thráinsson (1979) and is misleading, since CLDs do not need to be contrastive (Frey 2004a), and arguably not even topics (Repp and Drenhaus 2011). CLDs are categorically more unrestricted than HTLDs and display certain connectivity effects into the core clause, such as coherent case morphology, reconstruction effects for binding and sensitivity to weak islands, while lacking others, the most notable being their ability to escape the linear V2 requirement that generally holds for moved elements (see Ott 2014 for a discussion). Because of these somewhat conflicting properties, there is no agreement about whether to analyse CLDs as a movement dependency⁴³ (Grohmann 2003) or a base-generated element (Zaenen 1997; Frey 2004a); see also Ott (2014) for a third alternative, where the CLD is considered the result of ellipsis of an entire main clause. CLDs are embeddable at least in the Scandinavian (Thráinsson 2007:359), and possibly all, Germanic languages (see Bayer 2001:24 for an example from German).

Unfortunately, the evidence does not really allow us to diagnose the nature of the LDs in our corpus in a satisfactory manner. It is simply impossible to test for most connectivity effects, since this would require manipulation of the material and grammaticality judgements on phenomena like reconstruction or cross-over effects. On the other hand, it is possible to check if the case morphology is coherent or not. All of the LDs found in our corpus do in fact match the resumptive in case morphology. This is not strong evidence against their status as HTLDs, however, since there is only one case where the dislocated element is not the subject of the core clause (185), and in this particular case, the LD is a feminine noun which is morphologically identical in the nominative/*cas sujet* and the *cas régime* in Old French. On the other hand, one might perhaps expect Hanging Topics to be by default accusative/*cas régime* in Old French, rather than nominative, since there is already a clear tendency to overproduce the accusative at this stage (Foulet 1930:35-36). The fact that none of the LDs in our corpus are *accusativi pendentes*, in spite of the fact that many of them are disrupted by much parenthetical material, could tentatively be interpreted to mean that at least some of them are CLDs. However, it is also possible to turn this argument around; if one does not accept the idea that accusative/*cas régime* should be considered a default, perhaps one could argue that the fact that most LDs in (181–(185) are interrupted by much parenthetical material rather suggests that they are strongly disintegrated from the rest of the clause, and that this in turn favours an analysis as HTLDs. The matter cannot be settled here, but the possibility that some of these LDs are in fact hanging topics cannot be discarded.⁴⁴ The issue of their embeddability must wait until chapter 4.

⁴³See Alexiadou (2006) for overview and discussion.

⁴⁴This would be in line with Salvesen's (2013) view that both types of LD were available in Old French. However, the only example of a HT provided by Salvesen is an initial direct address ('vocative') in nominative case which is picked up by a resumptive in accusative (2013:146). Such examples are not entirely decisive, not only since the morphological vocative had vanished out of the language in favour of the nominative, but

3.7.3 The initial subordinate clause

We now turn our attention to the most frequent source of linear V3 by far, namely initial subordinate clauses. To the best of my knowledge, this deviation from the general linear V2 pattern is not discussed at all in the traditional literature on Old French. In the modern research literature, it has been pointed out repeatedly (Roberts 1993; Vance 1997; Donaldson 2012; Salvesen 2013; Wolfe 2015b and many more), and has sometimes been used as an argument against the V2 status of Old French (Kaiser 2002; Elsig 2009).

The corpus contains a wide variety of different types of embedded clause in initial position of main clauses. By far the most frequent of these are initial temporal adverbial clauses, in particular the ones introduced by *quant* – ‘when’ –, which on their own constitute the bulk of the relevant examples. Other elements that introduce temporal clauses are *si tost come* – ‘as soon as’ –, *ancois que/ainz que* – ‘before’ –, *apres ce que* – ‘after’ –, *en ce que* – ‘while’ –, *que que* – ‘(all the) while’ – *Ensi com* – ‘while thus’ – *maintenant que* – ‘now that’ –, as well as temporal clauses introduced by *la ou* – ‘as, while’. In addition to these, we also find causal adverbial clauses introduced by *por ce que/puis que* – ‘since, because’ –, conditional clauses introduced by *se* – ‘if’ – and comparative clauses like *(tout) aussi/ensi come* – ‘(just) like’.⁴⁵

As for the way the initial subordinates pattern with respect to their matrix clauses, it is possible to distinguish between three different groups (186)–(188), one of which (186) can be further divided into two sub-groups. This gives the following four options:

- (186) The initial subordinate clause is followed directly by the matrix clause without the use of a resumptive, giving linear V3. The matrix clause can be either:
 - a. Option I : subject-initial (string CS(p)VX)
 - or:
 - b. Option II: inverted (string CCV(S(p))X)
- (187) Option III: the initial subordinate clause is followed by a resumptive in the matrix clause – generally *si* or an adverb like *adonc* – and then the verb, giving the V3 string C-resumptive-V.... Postverbal subjects can be nominal, pronominal or null.
- (188) Option IV: the initial subordinate clause is followed directly by the finite verb, giving the V2 string CV.... Postverbal subjects can be nominal, pronominal or null.

Notice that while option (186a) differs from the other options in being the only pattern not involving inversion in the matrix clause, option (188) is the only one in fact triggering inversion and thereby involving linear V2. Because of this subtle and potentially confusing variation, I will for clarity borrow the terminology of the field model (Drach 1963/1937) in the following discussion and use the term ‘prefield’ to mean the position *directly to the left of the verb*. Thus, option IV (188) is the only configuration where the initial subordinate clause is in the prefield. In the other cases, the initial clause is external to the prefield, what Skårup (1975) termed *la zone d’annexe* – ‘the appendix zone’.

While all of these options are attested in our corpus, their distribution is very uneven. This is a domain of clausal grammar where the differences between Old French and the

also since a (true) vocative could not possibly be resumed by a case-matching resumptive anyway. Also, it is not clear in what sense such cases can be interpreted as aboutness topics, if this is to be a defining property of HTLDs.

⁴⁵There is some variation between the texts here: while for instance *que que* is exclusively used in *Eustace*, *maintenant que* is on the other hand only encountered in *Tristan*.

modern Germanic languages are very salient. While the latter generally make use of options (187) and (188), Old French has a strong predilection for option (186a). This pattern, where the verb comes in linear third position without a resumptive in the prefield, is generally ungrammatical in the Germanic V2 languages. However, some cases of this pattern can even be found in Modern Germanic, in particular in the case of so-called initial biscuit conditionals (Krifka 2017; Csipak 2018; see section 2.5.2.1 in chapter 2). This is potentially revealing, since it indicates that there is a connection between the logico-semantic relationship between the conditional and the following proposition on the one hand and their syntactic integration on the other hand.

Furthermore, option (187) is heavily used in spoken verb-second varieties (Faarlund et al. 1997; Eide 2011), and there is reason to assume that the presence of the resumptive is to some extent sensitive to the weight and complexity of the initial clause. In a sense, then, resumptives seem to signal a certain reluctance on the part of the speaker to use the ‘pure’ verb-second option (188). Taken together, this evidence could be interpreted as suggesting that embedded clauses as complex constituents lead a somewhat uneasy life in the prefield of V2 grammars. Diachronically, there is evidence that subordinate clauses are among the last elements to incorporate into the prefield and to yield to the general V2 pattern. In Old High German, initial adverbial clauses often give rise to linear third position of the verb, both with and without a resumptive element in the prefield (Axel 2007:229), a situation which carried on into Middle High German as well (Demske 2012).

In Old French, initial embedded clauses were presumably never fully integrated into the prefield.⁴⁶ The tables (3.7) and 3.8 show how the different clauses pattern with regard to the options in (186).⁴⁷ Although there is some variation, the overall message from these tables is that embedded clauses are highly unwelcome in the prefield. In *Eustace* they are not attested at all, whereas in *Tristan*, 6.35% of initial subordinates involve V2 constructions. Option III, which is the resumption strategy, is much more prevalent in *Eustace* than in *Tristan*.

3.7.3.1 Option I : non-inversion

The dominant pattern after initial subordinate clauses is non-inversion of the following matrix clause. I will return to the actual syntactic analysis shortly in section 3.7.4, so for the moment this pattern is simply illustrated in (189)–(193) for various kinds of subordinate clauses:

- (189) [*Quant il li ot tot conté*], [*sa feme*] *s’ escria e li dist*
 when he her.CL had all told his wife REFL.CL cried and him.CL said
 ‘When he had told her everything, his wife cried out and said to him . . .’ (Eustace, p.8 : VI. 3-4)

⁴⁶ According to Donaldson (2012), there was a period from the late 12th to the early 13th century where initial subordinates showed some tendency towards integration with the matrix clause. In Vance et al. 2009 it is even suggested that the situation in the early 13th century is already due to a weakening of the V2 rule, implying that the non-inverting pattern I might be an innovation, ‘the first wave of loss of V2’.

⁴⁷ The term ‘embedder’ in these tables is used to encompass both heads and phrases. Most of these are phrases (in Spec-CP, presumably), although the status of *quand* is not entirely clear, and *se* is most likely a head, C⁰/Fin⁰.

Table 3.7: *Tristan*: Subordinate clauses preceding their matrix clause: patterns of syntactic integration

Embedder	I (<i>CSVX</i>)	II (<i>CCV...</i>)	III (<i>C-res-V...</i>)	IV (<i>CV...</i>)	Total
<i>Quand</i>	75 (91.46%)	2 (2.44%)	5 (6.10%)	-	82 (100.00%)
<i>Apres que</i>	-	-	-	1 (100.00%)	1 (100.00%)
<i>Maintenant que</i>	4 (80.00%)	-	-	1 (20.00%)	5 (100.00%)
<i>En ce que</i>	3 (100.00%)	-	-	-	3 (100.00%)
<i>La ou</i>	7 (100.00%)	-	-	-	7 (100.00%)
<i>(Tout) aussi/ensi come</i>	2 (50.00%)	-	-	2 (50.00%)	4 (100.00%)
<i>Por ce que/puis que</i>	3 (50.00%)	-	-	3 (50.00%)	6 (100.00%)
<i>Se</i>	12 (66.67%)	5 (27.78%)	-	1 (5.56%)	18 (100.00%)
Total	106 (84.13%)	7 (5.56%)	5 (3.97%)	8 (6.35%)	126 (100.00%)

Table 3.8: *Eustace*: Subordinate clauses preceding their matrix clause: patterns of syntactic integration

Embedder	I (<i>CSVX</i>)	II (<i>CCV...</i>)	III (<i>C-res-V...</i>)	IV (<i>CV...</i>)	Total
<i>Quand</i>	29 (80.56%)	-	7 (19.44%)	-	36 (100.00%)
<i>Encois/Ainz que</i>	2 (100.00%)	-	-	-	2 (100.00%)
<i>Que que</i>	5 (100.00%)	-	-	-	5 (100.00%)
<i>La ou</i>	1 (100.00%)	-	-	-	1 (100.00%)
<i>(Tout) aussi/ensi come</i>	-	2 (100.00%)	-	-	2 (100.00%)
<i>Por ce que</i>	1 (100.00%)	-	-	-	1 (100.00%)
<i>Se</i>	3 (60.00%)	1 (20.00%)	1 (20.00%)	-	5 (100.00%)
Total	41 (78.85%)	3 (5.77%)	8 (15.32%)	-	52 (100.00%)

(190) [*Mes si tost come li lions fu pres des amis Nostre Seignor*], [*il*] *beissa*
 but as soon as the lion was close to-the friends Our Lord.OBL he lowered
le chief ...
 the head

‘As soon as the lion came close to Our Lord’s companions, it lowered its head ...’
 (Eustace, p.40 : XXXVI. 3-5)

- (191) [*...e ançois qu' il a moi repairast*], [*uns lions*] *sailli* del bois ...
 and before that he to me returned a lion came-out of-the forest
 '... and before he could return to me, a lion came out of the woods...' (Eustace, p.32 : XXVIII. 28-29)
- (192) [*Et en ce que celi chevauchoit*] [...] [*il*] *avint* que aventure l'
 and in this that this-one rode it happened that adventure him.CL
aporta jusqu' a un rochoi ...
 brought all-the-way to a cliff
 'And while he was riding [...] fortune happened to bring him to a cliff...' (Tristan, p.41 : 4. 4-6)
- (193) [*Que qu' il parloit ensi*], [*il*] *ploroit e sospiroit*...
 what what he talked such he cried and sighed
 'And while he was talking, he was crying and sobbing ...' (Eustace, p.20 : XVII. 1-2)

3.7.3.2 Option II: a rare pattern, but why?

The tables 3.7 and 3.8 also reveal another interesting thing. Option II (186), which is the pattern where an initial subordinate clause is followed by an inverted main clause, is also very unpopular and systematically avoided except occasionally after conditional clauses introduced by *se*. This is strictly speaking completely unexpected. Given a theoretical framework where information structure is constrained uniquely by syntax, one might wonder why a left-peripheral embedded clause should interfere with the word order of the following matrix clause, as the former must be assumed to neither occupy nor move through the prefield of the latter. Recall that around half of all main clauses in both texts feature a non-subject constituent in initial position of linear V2 strings (see table 3.3 in section 3.3.1). However, an initial subordinate clause almost systematically blocks any kind of XP-fronting to the prefield in the following matrix clause. The rarity of pattern II suggests we are dealing with an independent discourse constraint here. Observe furthermore that this pattern is not really banned, it is just preferably avoided, as evidenced by the following example:

- (194) ... [*après ce que il m' a servi et honoré en sa terre, se je aucun*
 after this that he me.CL has served and honoured in his land if I some
gerredon ne l' en rendoie en la moie], [*a felonie*] le
 recompense NEG.CL him.CL of.it-CL render in the mine of felony it.CL
porroit l'en tenir.
 could man hold.
 'After he has served and honoured me in his land, if I do not return him the favour in my own, people might consider it dishonourable.' (Tristan, p.56 : 46.5-8)

This shows that pattern II is indeed possible, a fact which almost makes the paucity of examples even harder to explain, since a total ban might be amenable to some kind of principled syntactic explanation. As it stands, one might conclude from examples like (194) that this word order is completely fine at the level of *grammar*, but that it is marginal at the level of *usage*, but this is hardly an explanation, since the difference between real discourse constraints and *usage* is not sufficiently worked-out from a theoretical perspective. Here is

another interesting topic for future research, but not one that I will pursue any further here, since the matter is strictly speaking somewhat peripheral to our concerns.

3.7.3.3 Option III: *si* and the resumption strategy

Occasionally, a resumptive element appears in the prefield immediately after the initial subordinate clause. The most frequent of these elements by far is the element *si*, a strongly polysemous and much discussed item of Old French grammar (see Marchello-Nizia 1985:15-18 (fn.) for references). According to Marchello-Nizia, it is possible to distinguish between 18 different uses of the element *si* in Old French. One of these, the use of *si* as a kind of resumptive in the prefield, is well attested in our corpus and in Old French in general (Einhorn 1974:115-116, Benincà 1995:333, Vance 1997:64-65, Wolfe 2015b:98-100). Two examples are provided in (195)–(196). On the rare occasion, a temporal adverb like *adonc* – ‘then’ – is used instead (197):

- (195) [*Et quant ce vint encontre le suer*], [*si*] *comença a changier li tens...*
 and was it came towards the evening SI started to change the weather
 ‘And when the evening approached, the weather started changing...’ (Tristan, p.44 : 14.2-3)
- (196) [*e quant il ot assez sermoné de la loi as crestiens*], [*si*] *les baptisa el non del Pere...*
 and was he had enough preached of the law to-the christians SI them-CL
 baptised en-the name of-the father
 ‘... and when he had preached the Law of the Christians for a long time, he baptized them in the name of the father...’ (Eustace, p.9 : VII.9-10)
- (197) [*Quant li jorz fu venuz et il pot bien veoir entor li*], [*adonc*] *fu il un po plus aese...*
 when the day was come and he could well see around him then was he a
 bit more calm
 ‘When the day had come and he could see well around himself, he was a little more calm...’ (Tristan, p.49 : 27.7-8)

The analysis of *si* in this particular use is not straightforward. The most fundamental questions revolve around the categorial status and the syntactic position of *si*, two issues which are strongly interrelated. As a very light and discourse-oriented element, it is tempting to suggest that *si* is a particle with X⁰-status in the syntax. If so, it is possible to analyse *si* as a particle lexicalising a left-peripheral head position, for instance Fin⁰. This is the central claim of the analysis developed for Old French by Ferraresi and Goldbach (2002) and by Ledgeway (2008) for Old Neapolitan.

However, there is problem with this analysis for Old French, namely the fact that it falsely predicts the order *XP-si-subject*, which is unattested in the corpus and generally not possible according to the consensus view in the research literature, which is that *si* must appear left-adjacent to the verb (Adams 1987a, Lemieux and Dupuis 1995:96, Ferraresi and Goldbach 2002:11, Salvesen 2013:142). Non-inverted main clauses in Old French invariably feature the subject at least as high as Spec-IP, as was demonstrated in section (3.5.1). If *si* is a particle lexicalising a C-head, we would not expect inversion after the particle, in particular since the lexicalisation of C-heads through base-generation of particles is considered an alternative

and competing strategy to verb movement (Roberts 2004; Ledgeway 2008). Notice that it does not help to assume that *si* rather lexicalises a high left-peripheral head like Force⁰, thereby structurally allowing enough space for inversion, for instance by having the verb move to Fin⁰. The reason is that such an analysis leaves unexplained why this construction is allowed to escape the EPP-effects, or in more neutral terms, the ban on verb-initial clauses (recall that *et/ne*-V1 clauses must be considered a construction apart) which is manifestly a very strong principle of Old French syntax.

Admittedly, the examples (195)–(196) do not feature an overt subject, and furthermore, this tendency towards null subjects after *si* is strong in the historical corpus in general (Marchello-Nizia 1985:48, Vance 1997:53, Wolfe 2015b:98–99), so one might raise the question if we are really dealing with structural inversion here. According to *Foulet’s generalisation*, which states that null-subjects are only permitted in postverbal position (Vanelli et al. 1985, Adams 1987a), these cases must involve structural inversion, but not all researchers accept that conclusion (Kaiser 2002; Rinke and Meisel 2009; Zimmermann 2009). While it is true that *si* in this particular use has a very strong tendency to trigger non-expression of the subject and might therefore also be considered a marker of *topic continuity* (Marchello-Nizia 1985:165, Benincà 1995:333, Vance 1995:183–184),⁴⁸ one does not have to search too far to come across examples of inversion after clause-initial *si*; the examples in (198)–(199) not only feature pronominal inversion, but also show that *si* triggers proclisis in accordance with the Tobler-Mussaffia Law:

- (198) *[se je en deusse orendroit morir], [si] le feisse je por la volenté mon*
 if I of-it.CL must now die SI it.CL would-do I for the will my
seignor accomplir
 lord.OBL accomplish
 ‘(Even) if I were to die now from it, I would still do it in order to carry out the wishes of my lord’
 (La Queste, taken from (Vance et al. 2009). Glosses slightly adapted.)
- (199) *... [se l’en vous donoit tout l’ empire], [si] l’ auriez vous bien*
 if man you.CL gave all the empire SI it.CL would-have you well
deservi
 deserved
 ‘if one gave you the whole empire you would deserve it.’
 (Villehardouin, adapted from Vance et al. 2009)

On the whole, these examples provide direct evidence against the hypothesis that *si* is a C-particle, and I therefore suggest that it is a phrase which can be used as an expletive in clause-initial position and as a resumptive after initial subordinate clauses, in line with the similar suggestion for Old Italian by Poletto (2005). It cannot be resolved at this point whether the expletive occupies Spec-CP or Spec-IP.

Occasionally, the resumptive *si* is used after other constituents than initial subordinates. This particularly happens after adverbial expressions of time like *après/puis* – ‘afterwards’ – or the like, much like the situation in modern Scandinavian, where such resumptives are ubiquitous in spoken language (Ekerot 1988; Nordström 2010; Eide 2011). It also seems to be the case that this use of *si* is easily triggered by the presence of intervening, parenthetical

⁴⁸But, as pointed out by Wolfe, (2015:100) it is also possible to find cases of inversion after initial *si* featuring indefinite and hence focal subjects.

material (200)–(201), a fact which strengthens the analysis as a resumptive, although it must be emphasized that this does not apply to all cases (202):

- (200) *et [après], (quant vos savrez lor volenté), [si] en overrons a*
 and afterwards when you will-know their will, SI of.it-CL we-will-work at
vostre conseil.
 your deliberation
 ‘... and afterwards, when you have learned their will, then we will act according to your judgement.’ (Tristan, p. 40, 2.8)
- (201) (Context: two of king Pelias’ knights have found Tristan sleeping. They recognize him and contemplate killing him. . .)
Mes alon au roy Pelias [...] e li conton ceste nouvelle.
 but let-us-go to-the king Pelias and him.CL let-us-tell this news
[Puis], (s’ il velt), [si] l’ ociron.
 afterward if he want SI him.CL we-will-kill
 ‘But let’s go to king Pelias and tell him this news. Then afterwards, if he wishes, then we’ll kill him.’ (Tristan, p.64 : 64.22-23)
- (202) (Context: the king has a dream about a lion and a leopard. First, the lion eats the leopard. . .)
Et [puis] [si] s’ en venoit par le roi et se gitoit
 and then SI REFL.CL of.it-CL came towards the king and REFL.CL threw
desor lui et le devoroit erranment.
 over him and him.CL devored quickly
 ‘And then it turned on the king and threw itself over him and devored him quickly.’ (Tristan, p. 46 : 20.9-10)

It is also possible to come across *si* after an initial noun phrase. There are no examples in *Tristan*, but (203 is a case from *Eustace*:

- (203) (Context: the emperor is angered that Eustace and his family refuse to revert to the old gods, and orders them to be sent to the arena to be fed to the lions. . .)
[L’ areinne] [si] estoit une mult grant place en Rome. . .
 the arena SI was a very big place in Rome
 ‘The arena was a very big place in Rome. . .’ (Eustace, p.40 : XXXV.22-23)

One might ask if the role of *si* in such cases is really just an extension of the resumption strategy. Wolfe (2015:98) cites similar examples, interpreting them as Hanging Topics. This entails that the initial noun phrase is in left dislocation. As we saw in section 3.7.2, unambiguous cases of left-dislocation in Old French always involve a resumptive element in the core clauses, contrary to normal, inversion-triggering topicalisation. Wolfe therefore suggests that *si* is able to fulfil the same function as the subject pronoun in resuming the NP/DP (2015:98).

There is a feeling that *si* contributes a bit more semantically than just being an index linked to the initial constituent. Marchello-Nizia claims that *si* encodes strong assertion, a commitment to the truth value of the proposition; this position is also adopted by Lemieux and Dupuis, who posit a projection ΣP above IP which hosts the initial XP in the specifier

and *si* in the head (Lemieux and Dupuis 1995). It may therefore well be that cases like (203) should be kept apart from the resumption strategy after initial subordinate clauses.

If we want to maintain that cases like (203) are in dislocation, an alternative is to assume that there is a resumptive null-subject in the core clause. This hypothesis receives some support from a rare and very interesting V4 case like (204). Here, the first constituent is clearly left-dislocated, since it is followed by an embedded clause which cannot be described as a parenthetical, being a (biscuit) conditional of which the matrix clause expresses the consequence. Since embedded clauses clearly belong to a position to the left of the main clause proper, the fact that the initial constituent in turn precedes this clause indicates that it occupies a very high position; this is in other words a very likely candidate for the Hanging Topic position. In the matrix clause, there is initial *si* followed by overt pronominal inversion:

- (204) (Context: Eustace compares himself to Job from the Bible, who was also tested by the Lord and dispossessed of his property. Eustace concludes that Job was after all in a better position than himself. . . :)

[*Cil*], [*se il n' ot rainseaus*], [*si*] *ot il racine: ce ert sa fame*. . .
 that-one if he NEG had branches SI had he root that was his wife

‘For he, even if he didn’t have branches, he did have a root: his wife. . .’ (Eustace, p.20 : XVI, 4-5)

This provides more evidence that *si* is not a C-particle, but a phrase in a specifier position, but more importantly, it suggests that the role of *si* in such cases is not to be a resumptive for the Hanging Topic, since this role is fulfilled by the inverted subject pronoun. This leaves two possibilities: either *si* is used as a resumptive after the conditional clause (recall from tables 3.7 and 3.8 that resumptive *si* is particularly frequent after conditional *se*-clauses), or it is used to express strong assertion, an interpretation which is clearly very plausible here. In fact, the two explanations do not exclude each other, at least not in this particular case.

On the other hand, it does not seem plausible to extend the assertion analysis to the cases involving initial subordinate clauses, since it would be very odd indeed if the appearance of an initial subordinate in general triggers strong assertions more easily than elsewhere. I therefore conclude that *si* may fill the role of a simple resumptive after initial subordinate clauses and presumably also some other temporal adverbial expressions. On the other hand, on the strength of the evidence from (204), I believe cases like (203), repeated below, rather feature a left-dislocated phrase, possibly a Hanging Topic, which is resumed by an inverted null-pronoun. The role of *si* in such cases is to emphasize the truth value of the proposition, as suggested by Marchello-Nizia (1985) and Lemieux and Dupuis (1995):

- (203) [*L' areinne*] [*si*] *estoit une mult grant place en Rome*. . .
 the arena SI was a very big place in Rome

‘The arena was truly a very big place in Rome. . .’ (Eustace, p.40 : XXXV.22-23)

3.7.3.4 Option IV: inversion

The final option is inversion of the matrix clause. Although this pattern is relatively rare, it is possible to come across bona fide examples; (205–206) clearly seem to feature the initial subordinate clause in the prefield, as evidenced not only by the following inversion, but also

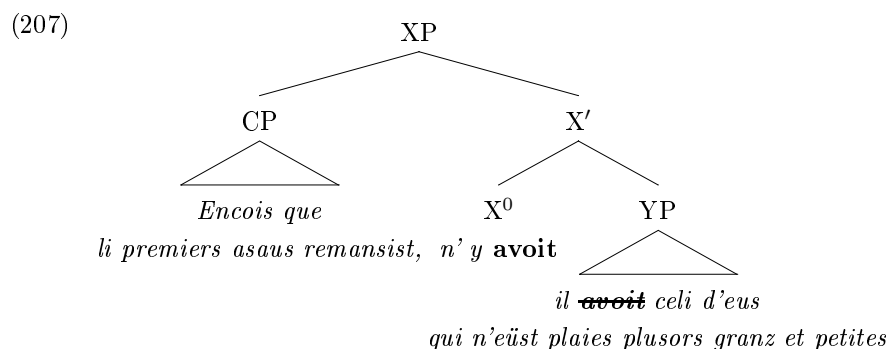
the TML-compliant proclisis of the pronominal and adverbial clitics in preverbal position. Furthermore, these examples feature inverted pronominal subjects, meaning the sequence cannot be a combination of an initial subordinate clause followed by a V1 clause:

- (205) [*Mes por ce que formé estoit a lor semblance, et creature,*] *le*
 But for this that formed was in their likeness, and creature, him.CL
secorront il por pitié de nature ...
 helped they for pity of nature ...
 “But since he was a living being that looked just like them, their compassion made them help him.” (*Tristan*, 31. 15-17)
- (206) [*... ençois que li premiers asaus remansist*], *n’ y avoit il*
 before that the first assault remained, NEG.CL there.CL had it
celi d’eus qui n’eüst plaies plusors granz et petites ...
 that-one of them who NEG.CL had wounds several great and small
 ‘... before the first assault was over, none of them was left without several wounds both small and big ...’ (*Tristan*, p.60 : 56. 6-7)

These examples show that subordinate clauses are not treated in uniform manner in 13th century Old French. While they are generally ‘invisible’ to the inversion mechanism of the language, they sometimes manage to trigger it. There is apparently some degree of optionality in the grammar here. It is impossible to know if this hesitation is the expression of the inversion grammar at its height or rather its first stage of decline. In other words, we cannot tell if early 13th century OF was briefly in the process of conquering the subordinate clause as well for the inversion grammar, a process which never materialised completely, or if the stage we witness here is already the first phase of the decline of inversion, which subsequently spread to other constituents.

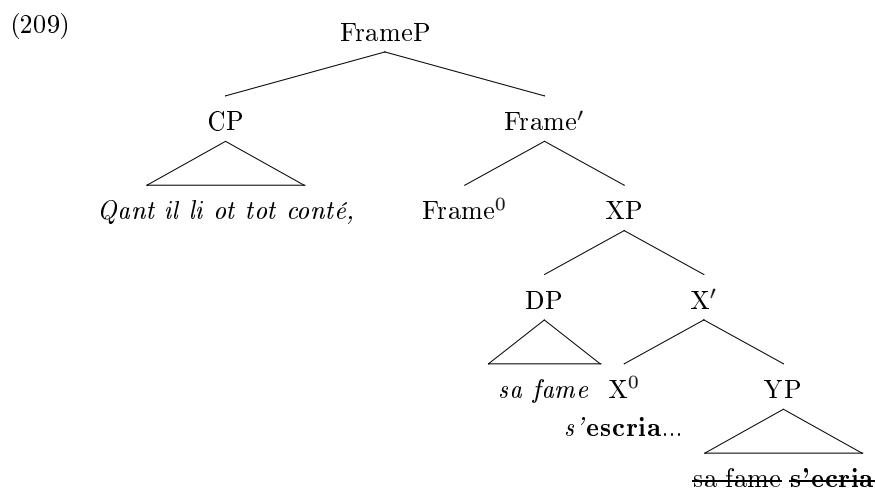
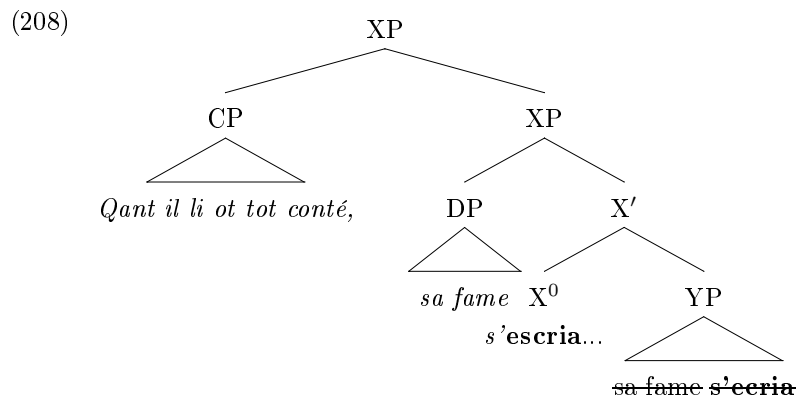
3.7.4 The formal analysis of initial subordinate clauses

We have just seen that initial subordinate clauses can enter into four different surface configurations relative to their matrix clause. As for their syntactic integration on a structural level, however, we may assume that there are only two different options. In some rare cases (option IV), the initial subordinate seems to be in the prefield and to trigger subject-verb inversion like any other constituent. The structure of a clause like (206) can therefore be represented as in the following tree, where XP and YP for the moment stand in for the more precise labels that we still need to establish for such inversion structures:



This represents the marked option where the initial subordinate triggers inversion and linear V2. In all cases of linear V3 on the other hand, whether they belong to type I, II or III, the initial subordinate occupies a higher position at the left edge of the clause. In the case of pattern III, a resumptive then occupies the prefield.

As for the dominant, non-inverted pattern I, there are two basic options available to represent these structures: either we adjoin the initial subordinate to the maximal projection of the core clause (208), or we generate it in the specifier position of some dedicated functional projection such as for instance FrameP (209):



Which of the phrase markers in (208–209) is the more adequate? In order to evaluate this, it is necessary to have some idea of what possible theoretical distinction they could be thought of as representing. Notice in this respect that (209) is the more informative structure, since it attempts to establish a correlation between the syntactic position of the initial subordinate clause and its information-structural reading in cartographic fashion, while no such claim is made in (208). Accordingly, it is also possible to interpret (209) as representing a more fully *productive* pattern of the grammar, if the projection FrameP is conceived of as the locus of all scene-setters. Interpreted this way, (209) clearly embodies the stronger and hence theoretically more interesting claim, since it predicts that initial subordinate clauses fulfill the role of scene-setters and that such elements are external to

the core clause, regularly giving rise to linear V3. It has indeed been suggested before that initial subordinate clauses function as scene-setters and accordingly occupy a high position in FrameP (Donaldson 2012; Salvesen 2013), where they can be first-merged after the clause has been constructed. Moreover, Wolfe has recently claimed that this is generally possible for scene-setting elements in Old French (Wolfe 2015b).

The hypothesis that there is a productive projection hosting Frame-setters high at the left edge of the clause in Old French is interesting, but we cannot evaluate it just yet. We must defer it, along with several other pending questions, until we have a clearer picture of the general syntax. Concretely, we need to know if it is indeed correct that scene-setters can generally occur in this high position in Old French, and we have not seen the data on this just yet. For this reason, it is hard to say whether the adjunction-analysis in (208) or the analysis with the dedicated frame projection in (209) is the more adequate. I will shortly return to this issue.

It is important to emphasize, however, that regardless of which of these analyses is chosen, the V3 strings featuring initial subordinate clauses should probably not be interpreted as evidence against V-to-C movement. The reason for this is simply that the initial subordinate clause precedes the main clause entirely and hence extends the phrase marker on top of it. Ignoring the aforementioned relative paucity of inversion structures after the initial subordinate clause, the word order facts of the matrix clause are thus left wholly unaffected. It would seem that the child acquiring the language has no choice but to accommodate the initial subordinate clause by generating structure on top of the main clause, regardless if the latter is derived by V-to-C movement or not. If these V3 strings have received much attention in Kaiser (2002) and subsequent literature (Elsig 2009, 2012) as 'incompatible with verb-second', it is therefore not because they provide evidence against V-to-C movement, but rather because they fall outside a certain narrow definition of verb-second that does not allow linear V3 without the use of resumptives in the prefield, a move which is justified by appealing to an alleged universal ban on CP-adjunction. According to such a definition, only patterns III and IV in (186) are compatible with a verb-second language. This is a valid move, but nothing more than a *definitional move*, and also one that will run into severe empirical problems if intended to capture the totality of cases in modern Germanic (cf. biscuit conditionals, section 2.5.2.1).

3.7.5 Remaining V3 patterns

In the preceding sections, I have reviewed various kinds of deviations from the linear V2 order in the corpus. In section 3.7.1, it was demonstrated that in Old French, a very limited group of adverbial expressions fails to trigger inversion and hence feature linear V3. It was argued that, with the possible exception of the NPIs *onques* and *ya*, none of these expressions provide evidence against V-to-C movement. In section 3.7.3, I suggested that the same applies to initial subordinate clauses. Furthermore, the behaviour of both groups is predictable and systematic – although not entirely without variation, particularly in the case of the subordinates – and should accordingly not be analysed as free and productive word order variation at the level of clausal syntax, but rather as individual constructions whose idiosyncratic syntactic behaviour must be acquired and stored in a piecemeal fashion. Because of this state of affairs, these expressions do not fall foul of the definition of a verb-second language employed in this thesis.

Let us now examine the quantitative contribution of these various groups to the overall

amount of linear V3 in main clauses in order to get a clearer picture both of their impact on the data on linea order reported in tables 3.1 and 3.2 as well as the amount of ‘residue’ in the form of V3 orders that do not fall out from any of these groups. This information is presented in table 3.9.

Table 3.9: Triggers of V3 in main clauses in *Tristan* and *Eustace*

Trigger	Tristan	Eustace
Neporquant	3 (2.11%)	- (0.00%)
Certes	3 (2.11%)	- (0.00%)
Sans faille/sans doute	5 (3.52%)	- (0.00%)
Onques	2 (1.41%)	2 (2.99%)
Left dislocation	4 (2.82%)	3 (4.48%)
XP- <i>si</i> -V	3 (2.11%)	3 (4.48%)
Initial subordinate	118 (83.10%)	52 (77.61%)
Total	138 (97.18%)	60 (89.55%)
Total V3, main clauses	142 (100.00%)	67 (100.00%)
Residual V3	4 (2.82%)	7 (10.45%)

It is clear that initial subordinate clauses make up the lion’s share of linear V3 in both texts. At the same time, there is a significant difference between the two texts with respect to the amount of residue; as for *Tristan*, the amount of residual V3 is extremely low, accounting for only 2.82 % of all V3 strings, while the corresponding number for *Eustace* is almost four times as high (10.45%). The expected frequencies are too low for a Chi-square test, but a Fisher’s exact test shows that the difference in amount of ‘residual V3’ is statistically significant (p-value 0.0402). Let us briefly examine why this might be the case.

In fact, almost all exceptional cases of V3 in both texts are of a similar kind. They feature an initial constituent which functions as an temporal adverbial, often a PP, followed by a non-inverted main clause. This would at first sight seem to support Wolfe’s recent claim that there is a high FrameP in the left-periphery that can host initial scene-setters by base-generation, and which is therefore external and invisible to the computation of inversion (Wolfe 2015b:93). Wolfe suggested that initial subordinate clauses could occupy this position, which is why they generally fail to trigger inversion, as well as other adverbial expressions of time and place which have the appropriate semantics and are able to scope over the entire clause:

- (210) *[A mie nuit], (sanz plus atendre e sanz le seu de lor mesniee),*
a mid night without more wait.INF and without the knowing of their house
[il] en alerent a l’ evesque des crestiens...
they of.it.CL went to the bishop of-the christians
‘In the middle of the night, without delaying and without the rest of the house knowing, they went to the bishop of the Christians...’ (Eustace, p.9 : VII.2-4)

However, a closer inspection reveals that there are very salient differences between the two texts of the corpus in this respect, both quantitatively and qualitatively speaking. In *Tristan*, initial subordinate clauses aside, there are only 4 cases out of a total of 998 main clauses which feature such constructions, in other words 0.40% of the total amount of main clauses. This does not give the impression of a productive projection which can generally host scene-setters. Furthermore, when we consider those cases, they reveal a very consistent pattern; non-inversion is triggered by the presence of much intervening material (211)–(213), which apparently has the effect of dislocating the first constituent. This explanation does not easily extend to (214), since the intervening material is so short, but this is an isolated case:

- (211) [*Après la passion Nostre Seignor Jesu Crist*], (*par cui mort et*
 after the passion our saviour.OBL Jesus Christ.OBL, by whose death and
par cui travail nos fumes osté de la prison tenebreuse et de la
 by whose toil we where removed from the prison tenebrous and from the
mort pardurable.) [*Joseph d'Abarematie*], (*qui avoit esté son deciple*
 death eternal, Joseph of Arimathea, who had been his
feal et leal.) **vint puis en la Grant Bretagne** ...
 disciple faithful and loyal, came afterwards in the Great Britain ...
 'After the passion of our Saviour Jesus Christ, by whose death and suffering we
 have been removed from our tenebrous prison and from death eternal, Joseph of
 Arimathea, who had been his faithful and loyal disciple, came thereafter to Great
 Britain ...' (Tristan, p.40: 1. 1-4)
- (212) [*Hui cest jor*], (*quant je cuida estre fors de ceste forest*.) [*je*] **trovai**
 today this day, when I thought be.INF out of this forest, I me.CL found
devant la roiche meïsmes ou mes chevaux morut.
 before the rock same where my horse died.
 'This very day, when I thought I was on my way out of this forest, I found myself in
 front of the very same rock where my horse died.' (Tristan, p. 52: 35. 12-12)
- (213) [*A l' endemain*], (*quant li jorz aparut biaux et clers*), [*il*] **comencent**
 at the day-after when the day appeared beautiful and clear they start
a regarder le roi ...
 to look-at the king
 'The next morning, when day had broken clear and beautiful, they started looking
 at the king ...' (Tristan, p.55 : 43.1-2)
- (214) [*et au cheoir que il fait*], [*il*] **vole tot de plain en l' eve**.
 and in-the fall that he makes, he flies all of plain in the water.
 'as he falls, he drops straight into the water.' (Tristan, p.53: 38. 8-9)

Apart from these cases, all similar expressions trigger inversion. Of course, Wolfe's claim is not to be interpreted to mean that scene-setting elements *must* fail to trigger inversion or that they obligatorily give rise to V3 orders; clearly it is possible for such expressions to trigger the inversion mechanism. However, the evidence from *Tristan* even calls into the doubt the hypothesis that it is possible at all to let scene-setters precede the main clause. Examples abound in all parts of the text of candidate scene-setters, adverbial expressions of time (216)–(220), place (221)–(222) or reason (223)–(224) that might plausibly qualify

as scene-setters on the definition provided by Wolfe (cf. section 2.4.4.1)⁴⁹ and which are altogether parallel to the expressions he adduces for other Old Romance languages, where such elements indeed regularly fail to trigger inversion. Yet they invariably trigger inversion in the text. The following is just a very small sample:

- (215) [*Cele nuit*] **demora** Sador delez la fontene...
 that night lingered Sador next-to the fountain
 ‘That night Sador stayed there next to the fountain...’ (Tristan, p.42 : 8.10-11)
- (216) [*Celi jor que la nef ariva en Cornouaille.*] **estoit** li rois montez en une
 that day that the ship arrived in Cornwall was the king ascended in one
soe tor.
 his tower
 ‘That day when the ship arrived in Cornwall, the king had ascended one of his towers.’ (Tristan, p.45 : 18.7-8)
- (217) [*Un suer*] **gisoit** li rois en son lit...
 an evening laid the king in his bed
 ‘An evening the king was lying in his bed...’ (Tristan, p. 46: 20.1)
- (218) [*A l’endemain*] **revint** li phylosophes devant le roi ...
 On the day-after returned the philosopher before the king ...
 ‘The day after the philosopher returned before the king ...’ (Tristan, p.47: 22.1)
- (219) [*A celi tens que je vos cont*] **estoit** li reaumes de Cornouaille et celi
 at that time that I you.CL tell was the kingdom of Cornwall and the-one
de Leonois en la subjection au roi de Gaule.
 of Leonois in the subjection-of-the king of Gaul
 ‘At the time that I am talking about, the kingdoms of Cornwall and Leonois were subjects to the king of Gaul.’ (Tristan)
- (220) [*En ceste partie*] **dit** li contes que...
 en this part says the story that
 ‘Here the story tells that...’ (Tristan, p.49: 27.1)
- (221) [*En Cornouaille*] **avoit** a celi tens un roi paien qui estoit apelez Canor ...
 In Cornwall was at that time a king heathen who was called Canor ...
 ‘In Cornwall there was at that time a heathen king called Canor ...’ (Tristan, p.45: 18.1-2))
- (222) [*Leianz en cele cité*] **demorerent** il trois semaines e plus ...
 There en that city remained they three weeks and more
 ‘There in that city they stayed on for three weeks and more...’ (Tristan, p.62: 61.1)

⁴⁹Recall Wolfe’s definition of a scene-setter:

‘The pragmatic characteristics of this group of elements is homogeneous. They have adverbial characteristics, scope over the entire clause and anchor the speech-act either temporally, spatially or aspectually’. (Wolfe 2015b:14)

- (223) [*Et por le sens dont il estoit*], *le* **tenoient** *il tuit a*
 and for the wisdom of-which he was him.CL held they all for
phylosophe...
 philosopher
 ‘And because of the wisdom he possessed, they all held him for a philosopher...’
 (Tristan, p.47 : 21.4-5)
- (224) [*Et por la biauté de li*] *l’* **apelierent** *il Apolo l’ Aventureus...*
 and for the beauty of him him.CL called they Apolo the fortunate
 ‘And because of his beauty they called him Apollo the Fortunate’...
 (Tristan, p. 49: 26.8)

What these examples serve to illustrate is that scene-setters, rather than residing in a dedicated functional projection that precedes and is exempt from the operation of the inversion mechanism, are subsumed under this latter construction, just like in modern Germanic V2 languages. As we have just seen, initial subordinate clauses on the other hand behave quite differently. Now, it would be odd to suggest that Old French has the peculiar property that only initial subordinate clauses qualify as scene-setters, since cartographic projections in the left-periphery are A’ projections defined by their information-structural properties, and since we cannot reasonably conclude that these kinds of adverbial expressions have different IS properties in different languages. The difference between Old French and the other Old Romance languages must accordingly be sought in the syntax, and the evidence strongly suggests that initial subordinate clauses pattern differently due to their *syntactic status as clauses*, rather than due to any inherent informational-structural properties they might carry.

However, the situation is quite different in *Eustace*. Although the inversion mechanism is quite consistently maintained in this text as well, there are 7 cases of exceptional V3, constituting 1.86% of the total amount of main clauses – more than four times as much as *Tristan*. An example of a non-triggering adverbial expression was cited above (210), some others are provided in (225)–(228). These examples must be characterised as violations of the inversion mechanism :

- (225) [*Aprés*] [*il*] **comenda** *a chascun qu’il l’alassent querre...*
 afterwards he ordered to each that they him.CL should-go seek.INF
 ‘Afterwards he ordered everyone that they should go out and seek for him...’ (Eustace, p.22: XIX, 8-9)
- (226) *e* [*tantost*] [*li feus*] **devint** *douz e soef ausi come rosee...*
 and immediately the fire became calm and mild as like dew
 ‘and the fire immediately became calm and mild as dew...’ (Eustace, p.43 : XXXVII. 34-35)
- (227) [*Emprés*], (*qant il ot son afere atorné*), [*il*] *s’* **esmut** *a aler en*
 Afterwards when he had his matter prepared he REFL.CL moved to go en
bataille...
 battle
 ‘Afterwards, when he had prepared his journey, he rushed to set out for battle...’
 (Eustace, p.30 : XXVII.1-2)

- (228) *[au departir], [il] les beisa e acola e comenda a Dieu.*
 at-the depart.INF he them.CL kissed and embraced and commended to God
 Upon departure, he kissed and embraced them and commended them to God.’ (Eustace, p.28 : XXIV, 13-14)

Interestingly, then, there is a subtle, but still quite noticeable difference between *Tristan* and *Eustace*; while the former text displays an almost exceptionless adherence to a linear V2 rule, outside of those particular cases that were reviewed above, the latter already reveals some signs of the weaknesses in the inversion grammar that are familiar from the literature and that would only increase in the following two centuries. This pattern is also familiar by now from several modern Germanic varieties. Urban vernaculars in Germany and the Scandinavian languages (see Walkden 2017 and references therein), Germanic heritage languages in America (Schmid 2002; Larsson and Johannessen 2015; Arnbjörnsdóttir et al. to appear; Westergaard and Lohndal to appear), as well as Flemish dialects near the French border (Haegeman and Greco 2016) all show the exact same option of using V3 after initial circumstantial adverbial expressions. In these cases, it makes sense to assume that there is a productive FrameP available at the edge of the clause. It is very questionable, however, if this projection was generally available/fully productive in earlier stages of Old French.

This brings us over to a very important point. Since the two texts display these subtle differences, it is crucial to emphasize how to interpret the corpus as a whole. *Tristan* shows a grammar which not only systematically rejects ‘scene-setting’ V3 constructions of the kind occasionally found in *Eustace*, but which also allows for inversion after initial subordinate clauses with non-marginal frequency (6.35%), something which is never encountered in *Eustace*. Although the texts are considered to be roughly contemporaneous, there can be no doubt that *Tristan* represents a more robust state of the inversion grammar in its diachronic evolution. Of course, this is not to be interpreted as saying that the differences between *Tristan* and *Eustace* are diachronic. They might stem from other sources, such as diatopic variation. It is also possible that the fact that *Eustace* is translated from Latin plays some role. None of this really matters for the current argument, which is that *Tristan* is an authentic witness which reveals something about the Old French inversion system *at some point in time and space*, since there is very little chance that a quantitatively so robust textual sample is either the result of accidental gaps, or alternatively, that the adherence to a V2 pattern is the expression of some written norm. It is therefore very tempting to conclude that cases like (225)–(228) were outright ungrammatical in spoken French at some point. However, since it is not possible to draw firm conclusions on the basis of negative evidence, I will rather limit myself to a weaker, but still quite strong claim about the Old French inversion grammar:

Conclusion V:

At some stage of its diachronic evolution, Old French featured a linear V2 constraint which could only be circumvented in certain, narrowly definable constructions.

Crucially, these narrowly definable constructions did presumably not originally include initial scene-setters. The adoption of a productive FrameP above the locus where inversion is computed is presumably not the right characterisation of the Old French inversion system, which in fact was even stronger and resembled modern Germanic more closely. The FrameP identified by Wolfe (2015) is historically real, but it is already a sign of weakness, foreshadowing the loss of the inversion grammar.

I must emphasize that this is no claim, as of yet, about the structural underpinnings of the inversion system. However, if one accepts the view that V2 languages can be both I-V2 languages and C-V2 languages, the conclusion that Old French featured one of these V2 systems seems unescapable, since there clearly is a linear constraint at work, and there is no imaginable parse that does not bring the verb at least as high as I^0 . However, the assumption (or rather definition) adopted in this thesis is that a V2 grammar always involves V-to-C movement, and whether there was V-to-C movement or not in Old French still cannot be resolved on the basis of the evidence reviewed. On the other hand, the evidence built up until now does allow us to approach another important question which we had to leave aside earlier, namely the issue of the structural position of null subjects.

3.7.6 Foulet’s generalisation and null subjects again

The descriptive observation that I have dubbed Foulet’s generalisation states that null subjects in Old French are only permitted in postverbal position, in other words, in inversion structures. This empirical observation was developed into an explicit formal hypothesis within the generative framework by Vanelli et al. (1985) and in particular Adams (1987; 1987), which was very briefly reviewed in section 3.4.1. This analysis makes many concrete assumptions about the structure of the Old French clause, in particular that it was a V2 language that consistently moved the verb to C^0 . Some researchers have rejected the analysis of French as a V2 language and in consequence, they also reject Adam’s analysis. Furthermore, the very descriptive generalization that null subjects are only permitted postverbally is also rejected (Kaiser 2002; Zimmermann 2009; Rinke and Meisel 2009). The most explicit statement comes from Rinke and Meisel, who not only claim that null subjects could just as well be realized in preverbal position, but even add that they were even more likely to do so because they ‘usually constitute the topic of the sentence.’ (Rinke and Meisel 2009:98). For this reason, Rinke and Meisel concluded that CVX strings cannot be used as evidence for verb-second.

We are now finally in a position to approach the question of the position of the null subject in Old French. In order to approach this problem empirically, we start by observing that Foulet’s generalisation and Adam’s theory of pro-drop involves only a one-way implicational relationship between inversion and null subjects. In other words, the fact that null subjects are only licensed in postverbal position does not entail that all postverbal pronominal subjects must be phonologically null. We have seen ample evidence for this, as the pronominal inversion string CVSpX is well attested in both texts of the corpus, reaching almost 7% in *Tristan* and almost 9% in *Eustace*. Clearly then, pronominal subjects *can be postverbal*.

This does of course not logically entail that all pronominal subjects, even the unexpressed ones, must also be postverbal. Furthermore, it is not self-evident, although this seems to be an implicit assumption in much of the literature, that the only thing that distinguishes overt pronominal subjects and null subjects is that the former are given PF realization while the latter are not. Although the expression or non-expression of the subject pronoun, apart from in initial position, seems to be an optional choice, there might be governing, or at least influencing, factors. Beyond pragmatic factors such as the avoidance of ambiguity when there are several possible referents, one might imagine information-structural factors, or even syntactic factors like cliticization. But all of this strictly pertains to the possible factors governing the alternation between overt and null subjects, not their position in the clause. If it is true that null subjects may also be preverbal, this means that the CVX string can potentially conceal either a SCVX string (if the null subject precedes the first

constituent), or CSVX (if the null subject intervenes between the initial constituent of the verb. In both cases we get a V3 string rather than a V2 string. But if this is the case, *it should be possible to find such strings*. After all, nobody has ever suggested that it should be impossible for the subject to be pronounced in these configurations, nor are the prospects for such a hypothesis very promising. In other words, we may approach the question of the position of the null subjects indirectly by considering the strings SCVX and CSVX counterevidence.

After a complete scrutiny of the data from main clauses in the corpus, the answer is very clear: these strings are virtually not found. As for the string SCVX, it is encountered twice in both texts. In *Tristan*, both cases involve the expression *sans faille* intervening between the subject and the verb. It was suggested in section 3.7.1.2 that this is a common parenthetical interjection, but the exact analysis matters less than the fact that this is a particular case, a conclusion which receives strong support from the very fact that no other instances of this strings are encountered. In *Eustace*, the string SCVX also occurs twice, and involves an initial DP followed by *si* and then the verb. In short, this is very strong evidence that the string SCVX is generally not possible in Old French outside of these familiar contexts.

The same applies to the string CSVX. Although this string is very frequently encountered, almost all instances feature an initial subordinate clause followed by a non-inverted main clause, or cases where initial *neporquant*, *certes* or *sans faille* precede a subject-initial clause. We have already seen the few exceptions which exist in section 3.7.5; these involve initial adverbial expressions which occasionally fail to trigger inversion, and are markedly more frequent in *Eustace* than in *Tristan*. In a text like *Tristan*, such cases reduce to 0.40% of all main clauses, a very robust finding which clearly shows that Old French, at the height of its inversion grammar, did not allow the kind of V3 structures with preverbal pronominal subjects that have been reported for Old English (van Kemenade 1987) or Old High German (Tomaselli 1995) in the Germanic diachrony. It is therefore clear that Foulet's generalisation held for this stage of the language. This permits a final conclusion regarding the syntax of main clause:

Conclusion VI:

In early 13th century Old French, as a very robust generalisation, null subjects in main clauses are possible only in postverbal position.

In consequence, the rate of structural inversion in main clauses can be considered to coincide quite accurately with the amount of non-subject initial, linear V2 strings, and is therefore around 50%.

3.7.6.1 Stylistic Fronting in main clauses?

I will now suggest that there is one exception to Foulet's Generalisation. In section 3.3, it was demonstrated that the prefield is in principle able to host a great variety of different constituents in Old French. In section 3.3.1, however, we saw that certain constituents are only rarely encountered in the prefield. Among the less frequent are non-finite verbs like infinitives and participles. This is not surprising, since VP-fronting is a quite marked construction that is not employed very frequently in modern Germanic either. However, on closer scrutiny, many of the cases where infinitives and participles occupy the prefield in the corpus, the resulting construction does not resemble VP-fronting. Consider the examples in (229)–(232). First, notice how they involve a bare infinitive rather than an obvious case

of VP-fronting. Secondly, the infinitives express new, unexpected information, very much against the general tendency of V2, which is reluctant to put new information focus in the prefield. Thirdly, none of the examples involve overt inversion, just a subjectless CVX string. And finally, all the examples involve an impersonal predicate.

- (229) *Et il responent: [A saillir] t' i covient. . .*
 and they answered to come-out you.CL there.CL behooves
 'They answered: you must come out. . .' (Tristan, p.44-45: 16. 7-8)
- (230) *Ha! fait li roys, [a trover] le covint.*
 ha does the king to find him.CL behooves
 'Ha! says the king, you have to find him.' (Tristan, p.64: 65. 15)
- (231) *[A morir] te covient après ton lecheor.*
 to die you.CL behooves after your adulterer
 'You must die after your adultery.' (Tristan, p.65: 67.24)
- (232) *[a dire] vos estoit que mes freres vos a mesfait. . .*
 to say you.CL is what my brother you.CL has mistreated
 'you must tell what wickedness my brother has done to you.' (Tristan, p. 43: 11.15-16)

It is highly unlikely that all of these different and unrelated properties coalesce accidentally in all of these cases. I therefore suggest that these examples are not instances of the normal Old French inversion mechanism, but rather a different construction. In the next chapter, it will be argued that this is a fronting operation which is related, although not identical, to Stylistic Fronting in Icelandic (Maling 1990). Since this construction is much more frequent in embedded than in main clauses, and since it will play an important role in understanding the syntax of embedded clauses, I will defer both the general description and the analysis of Stylistic Fronting until chapter 4. The only thing which is important to emphasize at this point is that Stylistic Fronting is generally taken to be dependent on a subject gap in the clause, such that it cannot take place in the presence of an overt subject. The fact that the examples in (229)–(232) all feature a non-referential null-subject must be considered highly relevant, since this strongly suggests that Stylistic Fronting is triggered by the lack of a preverbal subject, and furthermore, that non-referential subjects might be dropped in preverbal position, contrary to what is the case for other null subjects.

This is certainly not the only way to interpret these data, and I will return to this in more detail in chapter 4, since this phenomenon might be key to understanding some subtleties of Old French syntax. In rounding off, let me also raise the question if this construction sometimes takes place even with referential subjects. Example (233) features a coordination structure and is therefore quite ambiguous, since the locus of coordination is not clear (CP or IP). In this particular case, it matters less than the fact that the second conjunct lacks a subject. In this conjunct clause, a past participle *retenu* – 'retained' – is fronted in bizarre manner to the position in front of the finite verb, yielding a string which seems infelicitous from the perspective of V2. This might suggest that this is a case of SF.

- (233) *Childeïs, li filz Maroveux [...] ala par maintes foiz veoir Sador, e*
 Childeis the son Marovex.OBL went PRT any times see Sador and
[retenu] l' eüst a compaignon, s' il vousist.
 reainted him.CL he-had.SBJV to companion if he wanted.SBJV

‘Childeis, the son of Marovex, went very many times to see Sador, and would have made his him companion, if he had wanted.’ (Tristan, p. 62: 61. 2-4)

It should be emphasized that no such cases featuring referential subjects were found in independent, that is non-coordinated, main clauses.

3.8 Summary

In this chapter, I have conducted a detailed review of many different aspects, both quantitative and qualitative, of the syntax of main clauses. The evidence has made it possible to draw several important conclusions.

First of all, the prefield in Old French was not reserved for subjects, but functioned as an A' position hosting phrases with different categorial status and a wide variety of grammatical functions. Secondly, the prefield in Old French was not reserved for topics, nor is it possible to make any strict qualitative generalization regarding the informational structural partitioning of the clause in linear V2 strings, although new information focus is not preferred in preverbal position. Thirdly, the evidence does not support the view that inversion in late Old French is sensitive to the type of predicate employed; rather, the inversion mechanism seems to be a completely syntactic principle which is automatically triggered by the fronting of a non-subject constituent to the prefield. Furthermore, null subjects are generally only licensed in postverbal position in accordance with Foulet's generalisation.

Finally, the Old French grammar generally restricted the number of constituents in the prefield to exactly one. V1 clauses are generally not permitted except for clauses starting with *et* or *ne*, and whatever the proper analysis of these, they must be kept apart from the general syntax of declarative clauses. Exceptions from the linear V2 pattern can be found, the most important from a quantitative perspective being initial subordinate clauses, which generally fail to trigger inversion, yielding linear V3. Other exceptions feature a narrowly definable class of adverbial expressions and parenthetical interjections. The most peculiar of these are the NPI-items *onques* and *ja*, which exhibit special syntax when fronted to the first position of the clause, such as a ban on pronominal subjects and a tendency to allow several constituents to appear before the verb. It was suggested, partially in line with conclusions made by Ingham, that these constructions are the remnants of an older stage of the language, an idiosyncratic island that had to be acquired on a lexical basis.

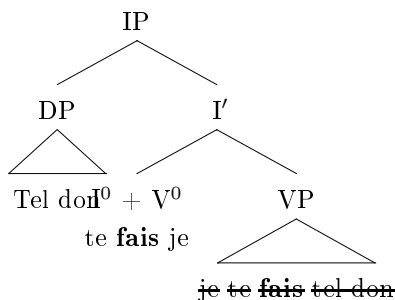
Outside of these particular domains, V3 is hardly found in *Tristan*, although *Eustace* occasionally features initial adverbial expression of time and place which may plausibly be considered to serve a scene-setting function. These constructions, it was suggested, should not be considered an integral part of the inversion system of the language in its original state, a state which is better preserved in *Tristan*, but rather as early signs of hesitation foreshadowing its future demise. It is also possible to conclude, on the basis of this evidence, that Old French, in descriptive terms, featured a linear V2 constraint which could only be circumvented in certain, narrowly definable constructions.

All of this provides very suggestive evidence in favour of considering Old French a V2 grammar derived by V-to-C movement. However, this latter conclusion cannot be drawn with complete certainty on the basis of the evidence from main clauses. Let me briefly recapitulate why this is the case.

3.8.1 A V-to-I model for Old French?

It is possible to argue that the finite verb only raises as high as I^0 , and that the position I have referred to as the prefield in this chapter is Spec-IP, an A' bar projection open to any kind of constituent. With these assumptions, it is possible to argue that inversion structures may be parsed into IPs rather than CPs. This line of argumentation is made possible by the crucial observation that, just like in the modern Scandinavian languages, nominal subjects are often not adjacent to the verb in inversion structures, being separated by the reinforcing negative adverb *pas* as well as other IP-adverbs. Therefore, the base position of the subject seems to be in Spec-vP. As for pronominal inversion (the string CVSpX), these subjects are always adjacent to the verb, invariably preceding the aforementioned adverbs. While this might be interpreted as a strong cue for V-to-C movement, it is still possible to maintain that postverbal pronominal subjects are clitics and that this explains why they are always adjacent to the verb. In other words, a clause like (96) might be given the following parse:

- (96) [*Tel don*] *te fais je, biaux amis.*
 Such gift.ACC you.CL make I good friend.
 ‘Such a gift I give to you, my good friend.’ (Tristan, p.40 : 2.23)



However, it is important to emphasize that in this model, I^0 must be equipped with an EPP-feature, since V1 clauses are generally not possible. Notice that no such EPP-feature is mentioned in the V-to-I models proposed by Kaiser (2002) or Rinke and Meisel (2009). However, without such a feature, a V-to-I model with the subject in Spec-VP would mean that Old French was in fact a VSO language. This conclusion, which presumably is unintended, (but see Ferraresi and Goldbach 2002), is incapable of dealing with the absence of true V1 orders in main clauses. If one adopts the EPP-feature, one might assume that it has the effect of attracting the subject from Spec-vP as the closest argument in the absence of true topicalisation or focalisation fronting, or alternatively, the highest adverb in the IP field. This would be in line with ‘formal movement’ approaches to V2 in Germanic (Fanselow 2002; Frey 2004b), although at the level of the IP rather than the CP. In addition, one might assume that a light adverbial like *si* or *lors* – ‘then’ – might be merged directly in Spec-IP as a Last Resort strategy. Finally, this model also has to avoid V3 orders somehow. It does not help to just adopt a ban on CP adjunction, one would in fact also need a ban on IP-adjunction (or something equivalent) to explain the general restriction to a single constituent in front of the verb. As already mentioned, this is of course already some kind of V2 model, but one that avoids V-to-C movement; essentially the model proposed by Lemieux and Dupuis (1995).

This model must be taken seriously, since a fundamental theoretical assumption in this thesis is that children only assign the minimal structure than is consistent with the global input. Other things being equal, a V-to-I parse is more economical and hence preferable to a V-to-C parse. However, we have not seen the global input yet. This is the topic of the next chapter, where we will consider the data from embedded clauses, and where we shall see that other things are in fact not equal at all.

Chapter 4

Old French: embedded clauses

In this chapter, I will consider the syntax and word order of embedded clauses in Old French. The syntax of embedded clauses is known to differ cross-linguistically from that of main clauses. In general terms, embedded clauses tend to be more constrained in terms of the word order variation permitted (Hooper and Thompson 1973; Cruschina 2010; cf. also the ‘Penthouse Principle’ of Ross 1973). This means that we might expect that the unmarked word order might appear more clearly than what is the case in main clauses.

Of particular relevance to the current investigation is the fact that the main-embedded asymmetry is a central characteristic of the verb second phenomenon, as was illustrated in chapter 2, where it was suggested that this asymmetry might in fact be common to all V2 languages. The reason this asymmetry arises is assumed to be the presence of a complementiser or subjunction in Fin^0 which blocks access to the left periphery, thereby bleeding V-to-C movement. If the Old French inversion system was indeed derived through V-to-C movement, we expect this to be clearly reflected in the quantitative and qualitative data. Nonetheless, we have to keep in mind that numerous exceptions are attested in the literature, including V-to-C in the complement clauses of viaduct verbs (see section 2.3.3) as well as in certain *peripheral* adverbial clauses (see section 2.3.5).

It has been claimed in the literature that children acquire the properties of their I-grammars (almost) exclusively from unembedded data (Lightfoot 1989, 1991), the so-called *degree-0 hypothesis*. This assumption is not adopted in this thesis. Rather, the guiding hypothesis here is that children are sensitive to the global input. This does not only mean that they take main and embedded data into equal consideration, but even that relevant cues for setting the syntax of main clauses may in fact be found in embedded clauses. In other words, embedded clauses may contain important information that helps narrow down the range of possible hypotheses regarding the structure of main clauses, the exact opposite direction of inference from what is assumed under the degree 0-hypothesis. The claim is of course not that children generally construct the grammar of main clauses based on embedded data. The hypothesis is that the global input is recruited to construct the grammar of both root and embedded clauses, and that the inferences may in principle run in both directions.

Structure The chapter is structured as follows. Section 4.1 discusses some major quantitative facts related to linear order and the prefield. Section 4.2 focuses on various quantitative and qualitative aspects of embedded inversion; considerable space is devoted to a discussion of the fronting phenomenon known as ‘Stylistic Fronting’ and its relevance to our

understanding of the general syntax of the language. In section 4.3, the issue of embedded V3 is addressed, with a particular focus on its consequences for the Force-V2 analysis (Wolfe 2015b). The final section 4.4 picks up some loose threads and suggests a concrete formal analysis of the Old French clause structure, based on the evidence from both main and embedded clauses.

4.1 Linear order and the prefield

The embedded clauses in the corpus were categorized into four different classes: complement clauses, adverbial clauses, interrogative clauses and relative clauses, and statistical information was extracted for each group individually and as a whole. We will proceed in similar fashion to what was done for main clauses in chapter 3, starting with some major, surface-oriented facts of a purely quantitative nature. In tables 4.1 and 4.2, the linear distribution of the finite verb in embedded clauses is presented. It was not found practical to include information about the different predicate classes, since there would be too many variables to present in one and the same table, but let it suffice to say that the predicate class has no interesting effect on linear word order in embedded clauses.

Table 4.1: Linear order of the finite verb in embedded clauses in *Tristan*

	Complement	Adverbial	Relative	Interrogative	Total
V1	– (0.00%)	6 (1.24%)	2 (0.64%)	1 (1.89%%)	9 (0.81%)
V2	250 (94.34%)	462 (95.65%)	199 (63.58%)	52 (98.11%)	963 (86.45%)
V3	13 (4.91%)	15 (3.11%)	112 (35.78%)	– (0.00%)	140 (12.57%)
V4	2 (0.75%)	– (0.00%)	– (0.00%)	– (0.00%)	2 (0.18%)
Total	265 (100.00%)	483 (100.00%)	313 (100.00%)	53 (100.00%)	1114 (100.00%)

Null-subjects (excluding relative and interrogative clauses): 52/748 = 6.95%

Table 4.2: Linear order of the finite verb in embedded clauses in *Eustace*

	Complement	Adverbial	Relative	Interrogative	Total
V1	– (0.00%)	2 (1.40%)	1 (0.63%)	1 (5.88%%)	4 (1.00%)
V2	80 (97.56%)	135 (94.41%)	122 (77.22%)	16 (94.12%)	353 (88.25%)
V3	2 (2.44%)	6 (4.20%)	35 (22.15%)	– (0.00%)	43 (10.75%)
Total	82 (100.00%)	143 (100.00%)	158 (100.00%)	17 (100.00%)	400 (100.00%)

Null-subjects (excluding relative and interrogative clauses): 9/225 = 4.00%

The first thing to notice is that one does not get the right impression by looking at the ‘Total’ column. There reason for this is that relative clauses feature a very particular

distribution in both texts that has a significant impact on the total.¹ When we disregard relative clauses, linear V2 is almost completely unchallenged in embedded clauses, as V3 orders are not common and V1 vanishingly rare. There is in other words a clear asymmetry with respect to main clauses that calls out for an explanation. Notice also that null-subjects are much rarer than in main clauses, where the corresponding figure for both texts was slightly above 32%. Although figures are low for both texts, there are much more null subjects in *Tristan* than in *Eustace*, a finding which is presumably not accidental, as we shall see.

Relative clauses evince particular word order properties with high proportions of linear V3 that set them apart. The reason for this is that relative clauses show a strong propensity for a particular construction which is illustrated in (234), and where an XP intervenes between the relative pronoun and the finite verb, causing linear V3 : ²

- (234) *Si m' eïst Diez! dist la dame [qui] [avec le chevalier] chevauchoit*
 if me.CL helps-SBJV God said the lady who with the knight rode
 'Good lord! exclaimed the Lady who rode with the knight.' (*Tristan* 24.5)

This fronting phenomenon has been attracted considerable attention in the literature on Old French (Dupuis 1989; Roberts 1993; Cardinaletti and Roberts 2002; Mathieu 2006b; Salvesen 2011; Labelle and Hirschbühler 2014) and has been equated by Mathieu (Mathieu 2006a,b, 2009, 2013) with the phenomenon of 'Stylistic Fronting' found in Icelandic (Maling 1990). I will return to it later in section 4.2.2, where I will argue that it provides very important insights into the syntax of Old French. In fact, as we will see, some embedded word order strings are ambiguous between Stylistic Fronting and verb-second.

However, as for the general syntax of relative and interrogative clauses, I will not be more concerned with it in this chapter. The reason is simply that these clause-types show

¹A word on how relative clauses and interrogative clauses were annotated is in order. For reasons which are explained in detail in the user manual that goes along with the data files in the TROLLing Repository (Klævik-Pettersen 2018), relative and interrogative clauses were annotated in an 'asymmetric' way for Old French. The relative pronoun or wh-phrase is never counted as a constituent when deciding linear order; however, the gap inside the clause that corresponds to the relative pronoun or wh-phrase *is* counted. This means that a subject relative clause like (i) is annotated as linear V2, since the gap of the relative pronoun is counted. In (ii), on the other hand, the gap is postverbal (since it corresponds to the direct object position) and is therefore not counted, meaning this clause is also counted as linear V2 :

- (i) *Li rois Canor avoit un frere qui [_] estoit apelez Peliadés*
 the king Canor had a brother who was called Peliades
 'King Canor had a brother who was called Peliades.' (*Tristan*)

- (ii) *ce sunt mi enfant que [j]' ai perduz _.*
 that are my children that I have lost
 'They are my children that I lost.' (*Eustace*)

The reason for this choice is naturally to be able to treat such clauses in uniform fashion and to say what is natural, that they are both normal, non-inverted V2 clauses. I consider this approach the best solution to the practical problems of annotating relative and interrogative clauses (see also Maling 1990 for a similar argument), but the approach is only justified as long as there is no doubt about the position of the clause internal gaps. For Latin, a different approach was adopted, since the position of the subject inside the clause is highly unclear.

²Notice that, although the relative pronoun is not counted, I use brackets around it for practical reasons to signal that the lacking subject is counted as a constituent. Since this fronting operation is generally dependent on a subject gap, there is reason to assume that the fronted XP occupies a position above IP and that it therefore precedes the subject gap. These facts will be discussed in some detail in section 4.2.2.

no signs of verb-second syntax in the corpus. I must clarify what is meant by this, since both clause types feature a majority of V2 orders, and this order is even almost exceptionless in interrogatives. The point is that there is *absolutely no inversion* in these clauses, just a staunch and unrelenting sequence *subject-verb*, with the exception of the aforementioned fronting construction in relative clauses. This yields a consistent linear V2 pattern, but there is no reason to expect V-to-C movement; in relative clauses like all other clauses, only inversion can be considered strong evidence for V-to-C movement. This is also much as expected when compared with the modern Germanic languages, where inversion is generally excluded in these contexts in all languages (except for some varieties of German, see section (2.3.5)). On the other hand, the modern Romance inversion structures are indeed possible in relative clauses, but they are not found in the corpus.

For these reasons, relative and adverbial clauses will be excluded from the rest of the discussion in this chapter, which focusses on the syntax of complement clauses and various kinds of adverbial clauses.

4.1.1 The prefield

The differences between main and embedded clauses become even clearer when we consider the V2 strings in more detail by looking at the type of constituents that appear in the prefield. We recall that subject-initial and non-subject initial strings were very evenly divided in main clauses. As table 4.3 illustrates, there is another sharp main-embedded asymmetry in this domain of the grammar, as subject-initial clauses now make up the overwhelming majority in both texts.³ This is a strong quantitative indication that the prefield functions differently in embedded clauses than in main clauses. The subject-initial pattern is slightly stronger in adverbial clauses than in complement clauses, but the difference is not significant (p-value 0.0662 if adverbial clauses from both texts are compared to complement clauses from both texts). *Tristan* also displays more non-subject initial clauses than *Eustace*; the difference between the texts is not significant for each clause type individually, but significant if complement and adverbial clauses are combined for each text (p-value 0.0178, d.f. 1, Chi-square 5.62).

³Notice also that there is a strong asymmetry as well with respect to the categorial profile of the subjects in the prefield. In main clauses (see table 3.3), pronominal subjects were only slightly more frequent than nominal subjects. In all embedded clauses, on the other hand, pronominal subjects are much more frequent. This is probably related to the general backgrounding function of embedded clauses, which do not introduce new discourse referents as easily as main clauses.

Table 4.3: *Tristan and Eustace*: The constituents in the prefield of V2 strings in embedded clauses

Initial XP	Complement clause		Adverbial clause	
	Tristan	Eustace	Tristan	Eustace
Nominal subject	68 (27.20%)	25 (31.25%)	98 (21.21%)	25 (18.52%)
Pronominal subject	151 (60.40%)	50 (62.50%)	325 (70.35%)	105 (77.78%)
Direct Object	3 (1.20%)	– (0.00%)	3 (0.65%)	2 (1.48%)
Oblique Object	– (0.00%)	1 (1.25%)	3 (0.65%)	1 (0.74%)
Predicate	– (0.00%)	2 (2.50%)	8 (1.73%)	– (0.00%)
Infinitive	2 (0.80%)	– (0.00%)	2 (0.43%)	– (0.00%)
Participle	– (0.00%)	– (0.00%)	2 (0.43%)	– (0.00%)
Adverbial	26 (10.40%)	2 (2.50%)	21 (4.55%)	2 (1.48%)
Subject-initial	219 (87.60%)	75 (93.75%)	423 (91.56%)	130 (96.30%)
Non-subject initial	31 (12.40%)	5 (6.25%)	39 (8.44%)	5 (3.70%)
Total	250 (100.00%)	80 (100.00%)	462 (100.00%)	135 (100.00%)

It was argued in chapter 3 that a V-to-I model, coupled with some additional assumptions, was reasonably well-equipped to account for the main clause data. However, this model faces considerable problems when confronted with the data in table 4.3. It is unclear why such a marked asymmetry between main and embedded clauses should exist in a grammar that produces inversion in main clauses by moving the verb only as high as I^0 . Since the entire IP is equally available in embedded clauses, the model does not predict this asymmetry. Granted, it is perhaps slightly too strong to claim that a V-to-I approach to inversion is outright incompatible with asymmetry between main and embedded clauses, since it is possible to argue that such asymmetries arise from differences in information structure in main and embedded clauses, and that these differences cannot simply be read off syntactic structure in cartographic fashion, but rather follow from independent principles. However, unless coupled with a concrete theory of what these independent principles might be and how they constrain the syntax, the null-hypothesis of the V-to-I parse is non-asymmetry between main and embedded clause.

Another problematic aspect of the V-to-I analysis is that it complicates the grammar and hence the acquisition process considerably. In particular, it makes it necessary to postulate far-reaching differences between main and embedded clause syntax beyond the accessibility or otherwise of the left periphery. In particular, Spec-IP, which was claimed to be an A' position in main clauses, capable of hosting all kinds of different phrases with different syntactic functions, all of a sudden starts behaving much like a position reserved for the subject of the clause. The idea that the same syntactic position can be an A' position in main clauses and an A(argument) position in embedded clauses is theoretically costly, but more importantly, it is not necessary. Spec-IP was shown to be a possible subject position

in main clauses as well, and if we maintain this insight, the syntax and word order of embedded clauses can to a large extent be reduced to a subset of main clauses; essentially the structure of subject-initial clauses that we established in section 3.5.1. There is no need for an independent and parallel acquisition of embedded syntax, a fact which must be considered a major advantage. Although I avoid the very strong assumption that children are ‘degree-0 learners’ in the sense of Lightfoot (1989; 1991), establishing their grammar (almost) exclusively from unembedded data, it seems reasonable to assume at the very least that children make no additional hypotheses about embedded clauses unless compelled to do so by the evidence. This is in fact included in the SSAP through the proviso that children account for the global evidence in a ‘maximally economic way.’ If one imagines, as we did in chapter 3, that the V-to-I and the V-to-C analyses compete in the internal grammar of the child as the appropriate representation of main clause inversion, the data from embedded clauses and its impact on the global evidence has the effect of shifting the balance in favour of the latter hypothesis. In fact, a V-to-C approach is not only compatible with the data in table 4.3, it concretely predicts that there will be asymmetries of exactly the kind observable in the corpus, since the C-layer is generally unavailable in embedded clauses due to the presence of the complementiser in C/Fin⁰. In this respect, a strong and interesting theoretical claim of the V-to-C analysis receives support.

It might be objected at this point that the alleged asymmetry between main and embedded clauses is not all that categorical and that the differences when it comes to the prefield are really more a matter of degree than a truly qualitative difference. After all, embedded clauses are not categorically SVX, as table 4.3 shows, but also feature other word orders, although to a much lesser degree than main clauses. This is correct, but it is important to emphasize that verb-second languages are not predicted to display a total and categorical ban on inversion in embedded clauses. On the contrary, embedded inversion is completely expected and is found in all of the modern Germanic V2 languages, albeit with some subtle variation among the different branches. It therefore behooves us to consider in more detail the contexts for embedded inversion in our corpus. Before doing so, however, it is important to consider the quantitative dimension of embedded inversion.

4.2 Embedded inversion

Recall that, in principle, a non-subject initial clause is not the same as an inverted clause, since the subject may be null. According to Foulet’s generalisation, null subjects are in fact only possible in inversion structures, suggesting the two notions are in fact equivalent in Old French. In chapter 3, it was argued that this generalisation holds for main clauses, and to the extent that there might be exceptions, these arise because the inversion grammar itself sometimes fails to be triggered. However, it cannot be taken for granted that embedded clauses behave in exactly the same way, so we must therefore examine the overt evidence for inversion. This information is presented in table 4.4. The predicate class variable is excluded for practical reasons, but let me again emphasize that inversion does not show any tendency at all to interact with the predicate class.

Table 4.4 demonstrates well how important it is, in principle, not to equate non-subject initial clauses with inversion from a surface perspective. When comparing the figures in the row ‘Postverbal S’ with the row ‘Non-subject initial’ in table 4.3, the difference is very clear, as the figures for inversion are considerably lower. This was the case in main clauses as well. Notice, however, that inversion is extremely rare in *Eustace*, where it is in fact not attested

at all in complement clauses. When we compare *Tristan* and *Eustace*, there is a quite salient difference with respect to the clause types, as inversion is appreciably less common in adverbial clauses in the former, while the opposite in fact holds for the latter. This is very interesting. If Foulet’s generalisation holds in full generality in embedded clauses as well, such that null subjects are only permitted in those cases where there is structural inversion, the asymmetry between the two texts would seem to be just accidental. There is reason to believe that this is not the case, as we will see later.

Table 4.4: *Tristan* and *Eustace*: preverbal, postverbal and null subjects (S) in embedded clauses

	Complement clauses		Adverbial clauses	
	Tristan	Eustace	Tristan	Eustace
Preverbal S	234 (88.30%)	77 (93.90%)	437 (90.48%)	135 (94.41%)
Postverbal S	16 (6.04%)	– (0.00%)	12 (2.48%)	4 (2.78%)
Null S	15 (5.66%)	5 (6.10%)	34 (7.04%)	4 (2.78%)
Total	265 (100.00%)	82 (100.00%)	483 (100.00%)	143 (100.00%)

At this point, we have reviewed enough quantitative evidence to establish with certainty the general syntax of embedded clauses. The data establish beyond reasonable doubt that Old French had already developed a basic SVO word order. An average of more than 90% subject-initial embedded clauses very strongly suggests that the prefield is an A position reserved for the subject, and does not lend support to the view that Old French displayed ‘Celtic’ tendencies (Ferraresi and Goldbach 2002:1). However, there are cases where an embedded clause features a non-subject constituent in the prefield, and furthermore, there is also a non-negligible amount of cases where there is overt subject-verb inversion. We must therefore consider some qualitative evidence in order to understand how to evaluate this variation.

4.2.1 Inversion in complement clauses

I will start by considering inversion in complement clauses, as this is the domain of embedded V2 in the modern Germanic languages that has been studied in most detail. In these languages, inversion in complement clauses is restricted by the matrix verb, with only some groups of predicates allowing embedded V2. Research has shown (Andersson 1975; Vikner 1995; Heycock 2006; Julien 2007; Salvesen and Walkden 2017) that these verbs generally overlap well with the class of verbs allowing root-phenomena in the important study of Hooper and Thompson (1973), in other words predicates of the classes A, B and E, which I will refer to collectively as ‘viaduct verbs’, following Walkden and Booth (to appear).

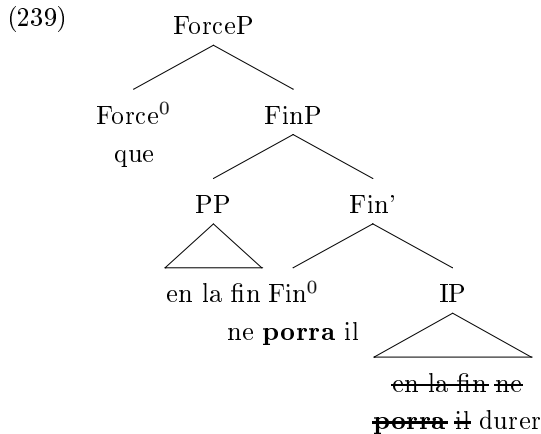
In testing if this pattern holds for the corpus as well, I will start out with the assumption that non-subject initial clauses are inversion structures, or to put it differently, that the string CVX is the product of inversion in the underlying syntax. I will return to this point

shortly and discuss to what extent this assumption really holds in embedded clauses as well.

As for *Tristan*, the observed cases of embedded, non-subject initial V2 almost completely overlap with the familiar group of predicates allowing embedded V2 in modern Germanic, containing in particular verbs of saying like *dire* – ‘say’ – (14 tokens) or *conter* – ‘tell’ – (1 token), verbs of thinking like *penser* – ‘think’ – (2 tokens), plus various ‘semi-factive verbs’ such as *savoir* – ‘know’ – (6 tokens), *voir* – ‘see’ – (2 tokens) and *conoistre* – ‘know’ – (1 token). The following examples, which are selected because they have an overt subject and are therefore particularly unambiguous, illustrate :

- (235) ... *dient que [ceste povreté] ne soefre il mie se non par neanté de cuer.*
 say-3PL that this poverty NEG.CL suffers he not if not by baseness of heart
 ‘They say that he does not suffer from such poverty if not for the wickedness of his heart.’ (Tristan, p.64 : 66. 8-9)
- (236) *il pensa que [par ceste chose] porroit il avoir Chelynde*
 he thought that by this thing could he have Chelynde
 ‘... he thought that through this chance he might have Chelynde.’ (Tristan, p.42: 9. 2-3)
- (237) *sachiez que [cist oraiges et ceste tempeste] [...] nos a Diex envoieé por le pechié d’ aucun de nos...*
 know that this thunder and this storm us.CL has God send for the
 sin of each of us
 ‘know that God has send this thunderstorm for the sins of each of us ...’ (Tristan, p.44: 15.2-4)
- (238) ... *il conoist bien que [en la fin] ne porra il durer.*
 he knows well that en the end NEG.CL can-FUT he last
 ‘He knows well that he cannot last until the end. (Tristan, p.61: 58. 6)

This must be considered very strong support for the view that embedded word order in Old French is constrained by very much the same syntactic principles as those operative in the modern Germanic V2 languages. In particular, inversion does not just occur randomly at the odd occasion, but rather appears to be highly systematic and predictable. The evidence from *Tristan* therefore mirrors with great precision the findings of Salvesen and Walkden (2017) in their investigation of embedded V2 in *La quête de Graal*, and provides strong support for the V-to-C hypothesis. The natural way to interpret these cases is to assume that viaduct verbs may select a high complementiser in Force⁰, thereby opening up the left periphery for XP-fronting and concomitant inversion. Adopting the assumption that the verb only moves as high as necessary to produce inversion in accordance with the SSAP, and ignoring other possible projections in the left periphery for the moment, a clause like (238) can therefore be represented as in (239):



4.2.2 Stylistic Fronting

In spite of the quite remarkable overlap between embedded V2 in modern Germanic and our corpus, there are some caveats which are important to bear in mind here. First, there are two examples of what one might call ‘unexpected V2’ in complement clauses in *Tristan*. The first is a case where the matrix verb is a copular predicate triggering a factive reading on the complement (240), in other words an instance of a class C predicate in Hooper and Thompson’s (1973) schema, a class which is generally hostile towards embedded root phenomena. The second is a negated verb of thinking (241), which is another staunch non-V2 context in modern Germanic languages:

- (240) ... *il s’ en vet grant aleüre, liez et joianz de ce que [ensi]*
 he REFL.CL of.it goes great speed, happy and joyful of this that such
li est venu de la roine.
 him.CL is happened of the queen
 ‘He departs in great haste, joyful and happy that it had turned out in this way for him with the queen.’⁴ (Tristan, p.54: 40. 4-5)

- (241) ... *car il ne cuidoient mie que [en tote Cornouaille] eüst un sol*
 for they NEG.CL think not that in all Cornwall had.SUBJ a single
chevalier qui encontre le roi Pelias osast porter armes.
 knight who against the king Pelias dared.SUBJ carry.INF weapons
 ‘...for they did not think that there was a single knight in Cornwall who dared to carry arms against king Pelias.’ (Tristan, p.59: 54. 12-14)

These examples at first seem quite unexpected, as the embedding predicates should not allow a main clause phenomenon like V2 inversion. However, appearances are probably

⁴It has been suggested to me that (240) can also be interpreted as a non-complement clause, either some sort of free relative or an adverbial clause of reason. However, the most natural interpretation to me is that of a complement clause, albeit possibly the complement of the preposition rather the adjective itself. This is also the opinion of Vance (Vance 1997:143). In Modern French, the complements of several verbs and copular expressions vary between taking the complementiser directly or the supported ‘de ce que’ construction; (cf. *être content que/de ce que* – ‘be happy/content that’) but unlike Modern French, complements of emotive verbs tend to take the indicative rather than the subjunctive in Old French (Jensen 1974:45-47, Jensen 1984:285).

deceptive in this case. Note that, unlike the sentences in (235–238), these two instances do not feature an overt subject. Secondly, both verbs are impersonal constructions which do not assign an external theta role and which therefore do not combine with a referential subject. It is therefore very likely to be the case that examples like (240) and (241) do not feature inversion and V2 at all, but rather an entirely different construction.

This phenomenon, which was briefly mentioned in section 4.1 and which is particularly prevalent in relative clauses, bears a strong resemblance to the construction ‘Stylistic Fronting’ (hereafter also SF) found in modern Icelandic (Maling 1990). This construction has several salient features that sets it apart from verb second, the most notable being the ‘subject gap requirement’ which rules out the construction in clauses with overt preverbal subjects. This explains why the construction is so prevalent in subject relatives, since there is an empty subject position in the clause. Furthermore, SF is constrained by quite strict locality conditions on the element that is to be fronted; unlike V2, SF can only front clause-mate constituents, and furthermore, the choice of constituent is governed by a strict ‘Accessibility Hierarchy’, the original version of which is given in (242):

- (242) *The Accessibility Hierarchy* of SF in Icelandic, according to Maling (1990:81):
 negation > predicate adjective > participle/verbal particle

Because of this, SF tends to front rather different constituents than what is normally fronted in V2, since all of the constituents in (242) are relatively uncommon in the prefield of V2 clauses (cf. the corresponding figures in table 3.3 in chapter 3). It has also been suggested that SF only fronts heads, in stark contrast to V2, which is generally taken to only front maximal projections. However, this view has been modified, and it is now assumed by many that SF can front both heads and phrases (Holmberg 2000; Thráinsson 2007); as for Old French, both Salvesen (2011) and Ott (2018) have argued that SF is derived through phrasal movement, including remnant VP movement. In general, the derivation of SF has generated considerable debate, with some of the more central issues being the landing space for the fronted constituent, the question whether the operation takes place in narrow syntax or in the phonological component, and relatedly, whether SF has any interpretive effect (see Holmberg 2006 for discussion and references).

Mathieu has argued that Stylistic Fronting did in fact exist in Old French (Mathieu 2006a,b) and that this is one of the Germanic properties of the language together with verb-second (Mathieu 2009), a direct result of historical Germanic influence. Mathieu’s analysis has met with criticism from researchers who point out that the Old French construction differs in several respects from the syntax of Stylistic Fronting as found in modern Icelandic. Labelle and Hirschbühler argue that the Old French construction is much less constrained in its application than its alleged Icelandic counterpart, for instance by disobeying constraints such as the requirement on a subject-gap or locality requirements on the fronted element as stated in the Accessibility Hierarchy (Labelle and Hirschbühler 2014). Our corpus confirms this, containing 14 tokens of this construction with overt subjects. It should be noted, however, that but for one ambiguous case which might not involve SF at all, all examples involve pronominal subjects, a point I will return to; consider the following examples with the (assumed SF)-fronted element underlined :

- (243) *Et se [je] a force l’ en voloie oster, ne porroie je*
 and if I by force him.CL therefrom.CL wanted remove, NEG.CL could I
mie.
 not

‘And if I had wanted to move him by force, I could not have done it.’ (Tristan, p.66: 71.15-16)

- (244) ... *e ançois qu’ [il] a moi repairast, uns lions sailli del bois...*
 and before that he to me returned.SUBJ a lion came-out from-the woods

‘And before he could return to me, a lion came out of the woods...’ (Eustace, p.32: XXVIII. 28-29)

4.2.3 SF or V2?

Although Labelle and Hirschbühler’s objections are highly pertinent, I will retain the appellation Stylistic Fronting since it has acquired some status in the literature. The most important point to emphasize here is that SF generally creates CVX strings, a string type which is also produced by embedded verb second. This means that these strings are potentially ambiguous between V2 and SF. When there is overt inversion, we can generally be quite confident that we are dealing with embedded verb-second, and the pronominal inversion string CVSpX is an unambiguous V2 string. However, in the absence of an overt subject, there is no infallible criterion for distinguishing the two fronting operations.

Nonetheless, there are several prototypical differences based on the relative characteristics of the two constructions that were mentioned above. For one thing, we have seen that embedded V2 is generally only available in specific contexts such as the complements of viaduct verbs and in certain ‘peripheral’ adverbial clauses (Haegeman 2007, 2010; cf. section 2.3.5). This is the reason why examples (240)–(241) above were singled out as suspicious, since the embedding predicates were not of the types that permit embedded root phenomena. It is also interesting to observe that both cases involved non-referential subjects in the embedded clause.

In fact, it is highly probable that the same analysis applies to some of the other cases of embedded non-subject-initial linear V2 in complement clauses. Although we have just illustrated that all of the other examples are in fact embedded under the appropriate kind of verbs, there are other reasons to prefer an SF-analysis for some of these. The first thing to note is that several of the relevant examples lack an overt subject. Among this group, several are in fact impersonal constructions lacking referential subjects (245)–(248), completely parallel to the cases in (240) and (241):

- (245) ... *il dist que [trop] seroit granz criauté s’ il l’ ocioit de sa main*
 he says that too-much would.be great cruelty if he him.CL killed by his hand
 ‘... he says (to himself) that it would be too cruel if he were to kill him with own hands.’ (Tristan, p.48: 25.4-5)

- (246) *Et neporquant, por ce qu’ il voient bien qu’[a faire] lor covient ...*
 and nonetheless for that they see well that to do them.CL behooves
 ‘And still, since they realized that they had to do it ...’ (=to continue the battle)
 (Tristan, p.60 : 56.3-4)

- (247) *Et quant je vis que [ensi] me covint demorer ou je vossise ou non*
 and when I saw that like-this me.CL behooved stay either I wanted or not
 ...

‘And when I saw that I had to sojourn in this fashion whether I wanted or not ...’
 (Tristan, p. 52: 35. 5)

- (248) *La novele cort par la vile, et dient li un et li autre que [devant
the news runs through the city, and say the one and the other that in-front-of
le temple Venus] gisoit deus homes morz.
the temple Venus lie two men.OBL dead
'The news spread through the city, and people start telling that two men are lying
dead in front of the temple to Venus.'* (Tristan, p.65: 69. 3-4)
- (249) *... il s' en revindrent au roy Pelias e li content que
... they REFL.CL of.it-CL returned to-the king Pelias and him.CL tell that
[trouver] nel poent.
find.INF NEG.him-CL can.
'... they returned to king Pelias and told him that they could not find him.'* (Tristan,
p.64: 65.15-16)

The fact that the predicates are impersonal is in itself no strong argument against embedded V2. However, a further indication that we are dealing with SF rather than V2 here is provided by the fact that the clause-initial elements are also quite unusual in terms of categorial profile as well as information structure. It was mentioned above that the IS properties of SF are not clear, but the IS properties of V2 are at least better understood. In fact, all the examples in (245–247) provide new/rhematic information, which, although not impossible, is still the least frequent filler of the prefield in V2 inversions in terms of information structure (cf. section 2.2.1 and section 3.3.2). In Stylistic Fronting, on the other hand, it is quite common.⁵ Furthermore, (249) and (246) feature infinitives, a syntactic category which is very common in SF, but once again rare in V2.

In sum, the combined testimony from the evidence, in the form of a lack of overt subjects, the impersonal nature of the predicates, the somewhat idiosyncratic information structure and categorial profile, strongly suggests that (245–247) do not feature V2, but rather Stylistic Fronting. If this is correct, these patterns add an interesting piece of information to the discussion around SF as well as to the theory of null-subjects. A priori, there is no connection between SF and impersonal predicates. The basic requirement on SF is that the clause contain no subject, or probably more precisely no subject in Spec-IP. While this subject-gap condition is not always respected in Old French, as we have seen, there is nevertheless good reason to maintain that this mechanism still plays a role. It seems to be the case that, whenever the subject is dropped, SF is immediately triggered as in the examples in (245)–(247), since there is not a single example of a verb-initial complement clause.⁶

⁵Again, this is not because SF is associated with focal readings per se. Generally, SF would seem to be a prime candidate for so-called ‘formal’ movement (cf. Fanselow 2002, Frey 2004b) that is not triggered by information structure, but rather by some purely syntactic (although seemingly optional) principle. In this respect, labels like TopP+ (Mathieu 2006b) or SFTopP (Ingham 2014) for the supposed landing site are somewhat misleading. In Icelandic, as already mentioned, SF obeys quite strict locality conditions, such that the fronted element will generally be the closest available element, head or phrase (Thráinsson 2007:380-385), in the clause. While these locality conditions might not be applicable or at least not systematically respected in Old French, (Labelle and Hirschbühler 2014) the result is still often that rhematic information ends up to the left of the verb. In principle, it seems like the fronted element can represent both old and new information, contrastive and non-contrastive.

⁶This observation in fact extends to all embedded clauses. Although the tables 150–4.3 report some scattered instances of V1 in adverbial clauses, these are in fact all of the same kind, namely a comparative adverbial clause introduced by the copula:

- (i) *... sa chevalerie sera autresi redotee entre les chevaliers, com est li lyons entre les
his chivalry will-be just-as feared between the knights, as is the lion between the*

This reveals an interesting aspect of SF that, to the best of my knowledge, has not been pointed out before: while SF is optional in its primary domain, which is that of subject relative clauses, it seems to be compulsory in all other contexts. If the subject in Spec-IP is dropped, SF immediately occurs. It is tempting to interpret this obligatoriness as a kind of repair strategy to make sure the embedded clause does not start with the finite verb. Either this means that there is an EPP on I⁰ as well, or it simply means that there is a purely linear V2 constraint in embedded clauses.

In fact, the trigger of SF is very important. As for non-referential subjects, one arguably does not really have to assume an empty category *pro* in the syntax at all, such that the merger of an expletive in preverbal position (triggered by the EPP) and Stylistic Fronting can be considered two alternative strategies of assuring that the clause does not open with the verb; if for some reason the EPP does not trigger the merger of an expletive, SF kicks in as a Last Resort rescue operation. This avoids postulating a *pro* in the clause altogether. However, this would lead us to expect that non-referential subjects are consistently dropped when another element is in the prefield. While this prediction is borne out in the vast majority of cases, there are in fact two tokens of overt, non-referential subjects in inversion strings:

- (250) [*De ma vie*], (*fait li preudons*), *ne me covient il mie panser*
of my life does the preudome NEG.CL me.CL behooves it not think
‘For my life, says the *preudome*, I do not have to worry...’ (Tristan, p.50: 30. 6-7)

It is not clear how a non-referential subject could be inverted and appear in postverbal position if it is not part of the syntax. On the other hand, if non-referential null subjects are also *pro* in the syntax, this would seem to entail that they can somehow be dropped in preverbal position, against the general rule of the language, and that this immediately triggers SF. This is very important, for reasons which will become clearer soon.

4.2.3.1 Preverbal referential null subjects?

In fact, there is reason to suspect that even referential subjects can occasionally be dropped in preverbal position in embedded clauses. As we saw in chapter 3, this generally does not happen in main clauses. Yet examples like the following raise some doubts as to whether this rule is completely robust in embedded clauses:

- (251) *Celi jor que la nef ariva en Cornoaille, estoit li rois montez en une soe*
that day that the ship arrived in Cornwall was the king ascended in one his
tor. Et quant il la vit a la rive, il sot bien que [d’ estreng
towers. And when he her.CL saw on the shore, he knew well that from strange
païs] venoit.
land came
‘On that day when the ship arrived in Cornwall, the king had ascended to one of his
towers. And when he saw the ship on the shore, he knew that it come from foreign
lands.’ (Tristan, p.45: 18.7-8)

autres bestes.

other animals.

‘...his chivalric prowess will be just as feared among the knights as the lion is among the other beasts.’ (Tristan, p.47: 22.5-7)

This seems to be the only context where embedded V1 is possible in our corpus.

In this case, the embedding verb is of the appropriate kind, the predicate is not impersonal and the null subject is accordingly referential, and the syntactic category – a PP – is completely normal, so the only indication that we are dealing with SF rather than V2 is the information structure. The PP introduces rhematic and unexpected information, very much against the tendency in V2 inversions, and the corresponding clause seems to rather marginal at least in modern Mainland Scandinavian.

Since SF seems to apply whenever there is a subject-gap in Spec-IP, we might also hypothesize that it might be compatible with cases of overt DP subjects in so-called ‘non contiguous’ inversions. This has been shown to be the case in modern Icelandic (Thráinsson 2007), and Franco has argued for a similar analysis for Old Italian (Franco 2017). These inversion structures are not very common in main clauses and even more restricted in embedded clauses, but in principle one might argue for an SF analysis of cases like the following, as the initial element is once again rhematic, focal and ill-suited for V2:

- (252) *A piés i vint, e a piés s' en vait, molt*
 on feet there.CL he-came, and on feet REFL.CL therefrom.CL he-goes much
esbahiz e molt coreciez, e dit que [mauvés guerredon] li ont
 astonished and much angered and says that bad retribution him.CL have
rendu cil de Cornouille . . .
 rendered those of Cornwall
 ‘He came by foot, and by foot he left, greatly astounded and angry, saying to himself that the people of Cornwall had given him a poor treatment . . .’ (Tristan, p. 63: 63.3-5)

There is presumably no way to prove that we are dealing with SF in such cases, which accordingly must remain ambiguous. Since we are not primarily focusing on SF, this matters little; what is important, rather, is the consequences of these facts for our general theory of V2 and word order in Old French. The crucial thing to notice is that SF is available in various kinds of embedded contexts other than relative clauses, and that some of these clauses are demonstrably rather small from a structural perspective. In fact, all the evidence points towards a landing site for SF-fronted elements which is not left-peripheral at all. SF is available in all kinds of adverbial clauses, including in complement clauses which are not embedded under viaduct verbs. Since we can assume with reasonable certainty that the these clauses are no greater than FinPs, with the complementiser itself lexicalising Fin⁰, the availability of SF to the right of the complementiser in such contexts points to a very high position in the sentential core, made available by the absence of a subject in Spec-IP. In other words, *SF takes place below the CP*.

One might even be tempted to suggest that the landing site of SF-fronted elements is even below the subject position in Spec-IP as well, since SF occasionally occurs even in the presence of overt subjects, violating the ‘subject gap requirement’. However, it was already noted that this only ever happens with pronominal subjects. This is very unlikely to be due to chance, so I will suggest, following Roberts (1993:122-123), that the subject pronouns cliticize to C/Fin⁰ in such cases. There is independent evidence that this analysis is correct, as we shall see later on.

Let me now explain why the correct analysis of SF is so important to the understanding of V2. In section 3.7.6.1 of the previous chapter it was argued that SF can occur in main clauses as well. In our corpus, all reasonably clear cases are restricted to contexts of null expletives with impersonal predicates. However, Salvesen (2011) presents data obtained

from searching the online corpus *Corpus de la littérature médiévale* (CLM), arguing for an SF analysis of many main clauses featuring referential null subjects. If this analysis is correct, the general availability of SF in all kinds of embedded clauses in fact provides us with the answer to a long-standing controversy, namely the structural underpinnings of subject-initial main clauses.

The argument is simple. We have just established that SF takes place below FinP, yet to the left of the verb, which we must assume to be in I⁰. It matters little at present if this position is Spec-IP or some higher position between IP and FinP, although the latter hypothesis seems more natural in the case of referential null-subjects, since these must be assumed to feature a *pro* in Spec-IP, occupying that position. Crucially, when main clause subjects, expletive or referential, are dropped, SF raises an element to a position preceding the verb, just like in embedded clauses. We can therefore assume that the verb is no higher than I⁰ in such cases.

It should be noted that this argument hinges on the assumption that SF is a purely syntactic operation that targets a consistent landing site. This is the assumption of Mathieu (2006, 2009) and Ingham (2013). If it turns out to be a PF-phenomenon that just serves to prevent both main and embedded clauses from opening with the verb, then the preceding argument is not decisive, since SF would just be triggered by linearization concerns and perhaps not even need a dedicated landing place in syntax. I will assume that SF is an operation in the (narrow) syntax.

The evidence from SF therefore aligns neatly with that of initial subordinate clauses reviewed in section 3.7.4, where it was observed that inversion after initial embedded clauses is vanishingly rare. This is not surprising if the following non-inverted main clause in fact just an IP and the initial subordinate clause simply fails to trigger inversion. On the other hand, if the subject-initial main clause is a CP, the near-absence of inversion after initial subordinate clauses remains mysterious.

It was noted earlier that there is not a single case of overt inversion in *Eustace*. There are not many cases of non-subject-initial V2 in complement clauses either, but the ones which are found resemble SF rather than V2. (253) is embedded under a class D predicate, containing a factive verb of emotion, a very hostile environment for embedded root phenomena. The sentence in (254) on the other hand contains an appropriate matrix verb, belonging to the class of semi-factive predicates,⁷ but the information structure is singularly odd from the perspective of V2. In fact, the fronted element is old, yet non-topical information, what one might call ‘tail information’ in the sense of Engdahl and Vallduví (1996). Asserting a presupposition, if possible at all from a theoretical perspective, is at least not compatible with V2 in modern Germanic, and the corresponding clause would be infelicitous. And finally, regarding the example in (255), the focal reading of the fronted adjective also raises some suspicion, in particular since subject predicatives are highly prone to Stylistic Fronting, but the matter cannot be definitely settled. Note also that if these are indeed all cases of SF rather than V2, then (253) and (255) provide further evidence that even referential subjects can sometimes be dropped in preverbal position. Be that as it may, it is in either case clear that *Eustace* features very little if any inversion in complement clauses.

(253) *Molt s' en merveilloient tuit que [si sodainement] estoit adirez*
 Much REFL.CL thereof.CL marvelled all that so suddenly was lost

⁷Note that, unlike predicates of the assertive classes A and B, negation of the matrix verb in (254) is presumably not crucial for the availability of embedded verb-second. This is presumably related to the general fact that matrix negation does not alter the truth value of presuppositions (Kiparsky and Kiparsky 1970).

‘Everyone was greatly surprised that he vanished so quickly ...’ (Eustace, p.15 : XII.12-13)

- (254) ... *li dui enfant se herbergerent en la loge lor mere; mes*
 ... the two children REFL.CL accommodated en the lodge their mother. OBL; but
ne savoient que [lor mere] fust.
 NEG.CL knew that their mother was.

‘... the two children took up accommodation in the lodge of their mother; but they did not know that she was their mother.’ (Eustace, p.31 : XXVIII. 7-8)

- (255) *Lors reconut il que c’ estoit il ; de sa fame e de ses enfanz lors*
 Then admitted he that it was him; of his wife and of his children them. CL said
dist il que [mort] estoient.
 he that dead were.

‘Then he admitted to being him; of his wife and children he told that they were dead.’ (Eustace, p.27 : XXIV. 1-2)

4.2.4 Inversion in adverbial clauses

Although less discussed than complement clauses in the literature on embedded verb-second, it has been established that V2 can sometimes operate in certain adverbial clauses as well (cf. section 2.3.5). Of the 39 occurrences of non-subject-initial linear V2 in adverbial clauses *Tristan*, there are only a handful of cases of unambiguous V2, and they are all found in consecutive adverbial clauses, a domain which is known to permit root phenomena (cf. section 2.3.5):⁸

- (256) *Cele avoit embracié un fuissel de la nef, ou ele gisoit et*
 this-one had embraced a piece-of-wood of the ship, where she was-lying and
se tenoit desus, et l’ avoit estraint si fermement que [a pones] poïst
 REFL.CL held upon and it.CL had clasped si firmly that at pains could
ele estre ostee.
 she be removed.

‘She had clung to a piece of flotsam from the ship, upon which she was lying, and she had clasped so firmly onto it that she could barely be removed.’ (Tristan, p.41: 4-10-12)

- (257) *Et la meïsmes ou la tempeste et li orages avoit enbatue la*
 and there self where the storm and the thunderstorm had beaten the
nef en une roche estoit si merueilleusement qu’ [a pones] l’ en
 ship in a cliff was so marvelously that at pains it. therefrom. CL
poïst l’ en remuer...
 could man remove

‘And at that very place where the storm had smashed the ship into a cliff, chance had it so that one could barely remove it ...’ (Tristan, p.41: 4.12-14)

⁸By ‘unambiguous V2’, I mean any string *CVS(p)X*, regardless of the category of the initial constituent. In other words, I do not make a distinct between adjunct-initial or argument-initial embedded inversion as argued for by Haegeman (2012).

- (258) *Si descent maintenant et trait l' enfant de desoz son mantel, et le*
 SI descends now and pulls the child from beneath his mantle, and it.CL
voit si bele creature de son aage, que [por la biauté de li] l'
 sees so beautiful creature from his age, that or the beauty of him it.CL
en prent il totevoies si grant pitié.
 therefrom.CL takes he still so great pity.
 'He descends presently and takes the child out from beneath his mantle, and he finds him such a beautiful creature for his age that he takes great pity on him for his beauty ...' (Tristan, p.48: 25.2-5)

All other cases either clearly involve Stylistic Fronting or are ambiguous between V2 and SF, with the balance in favour of the latter analysis in most cases. It is worth noticing that SF sometimes shows tendencies to clustering effects, as in the following passage, where three consecutive embedded clauses arguably feature SF.

- (259) *Tant avon fait que [lassés] sommes outre mesure, et que [a*
 So-much have.1.PL done that tired are.1.PL beyond measure, and that to
morir] nus estuet, se [plus] en faisons.
 die us.CL behooves, if more thereof.CL do.1.PL.
 'We have toiled so much that we are exhausted beyond all measure and will die if we do more.' (Tristan, p.62: 59.6-7)

This might be interpreted as an indication that SF at least to some extent is a rhetoric or stylistic device available to a narrator, and we can only speculate if this means that SF was more common in writing than in spoken language. What seems clear, however, is that embedded verb-second is a quite restricted phenomenon, and such indisputable cases which are found in our corpus are all completely consistent and show a remarkable degree of overlap with reported cases of embedded V2 in the modern Scandinavian languages. To the extent that there exist different types of V2 languages, Old French is therefore firmly situated in the 'asymmetric' group. Moreover – and this cannot be emphasized enough – the general availability of the string CVX in embedded clauses neither gives reason to make an argument against verb-second syntax nor in favour of a 'generalized' or 'symmetric' V2 system, as these cases plausibly all involve Stylistic Fronting. Granted, SF in Old French is not identical to SF in modern Icelandic, but the reality of the construction and the fact that it is qualitatively distinct from verb-second seems beyond dispute.

4.2.4.1 Unexpected V2

It would not be right to pretend that the overlap between embedded V2 in our corpus and modern Germanic is complete or to ignore such modest counterexamples as can in fact be found. Even when disregarding structures which may possibly be explained by appealing to SF, there remains a couple of instances of seemingly 'unambiguous V2' in unexpected context. These are found in *Eustace*:

- (260) *Quant [ce] sorent e virent li mauvés voisin, il entrerent en sa*
 when this knew and saw the bad neighbours, they entered
meison par nuit ...
 into his house by night

‘when the wicked neighbours discovered and saw this, they entered into his house by night ...’ (Eustace, p.14: XII. 1-5)

(261) *Quant [ce] vit li tiranz*
When this saw the tyrant

‘When the tyrant saw this...’ (Eustace, p.39: XXXV. 18)

Embedded V2 in temporal adverbial clauses is not possible in modern Germanic.⁹ Out of a total of 159 clauses introduced by ‘quant’, these two are the only tokens of inversion. Since they both contain nominal subjects and only a simple main verb, it is in principle possible to argue that they could represent cases of ‘Romance inversion’, in which case the subject gap in Spec-IP makes room for SF-fronting. Such an argumentation is dangerous, however. Since we are dealing with no more than 2 isolated tokens, I will not go further into this here. I hypothesize that this is nothing more than some kind of literary reflex.

4.3 Embedded linear V3

In this section, we will briefly consider instances of embedded linear V3. In light of recent proposals for the clause structure of Old French and the mechanism of V2 at this stage of the language (Wolfe 2015b), these structures are in fact of particular theoretical importance and contribute more information about Old French V2 than do mere embedded linear V2 strings. We must distinguish between two different cases of embedded linear V3: those that are generally available in all kinds of embedded clauses, and those that are a product and a principled subset of embedded V2, found in exactly the same contexts as those already discussed in the previous section. We start out with the former.

4.3.1 Embedded linear V3 after the NPIs *onques*/*ja*(*mes*)

We already mentioned a case of embedded linear V3 above, namely those instances where Stylistic Fronting occurs even in the presence of a pronominal subject, thus producing the string *SpCVX*. In order to account for the apparent circumvention of the ‘subject gap requirement’ which is normally considered the *sine qua non* of SF in languages such as Icelandic, it was suggested that the pronominal subjects can cliticize to the complementiser in Fin⁰, thereby evacuating SpecIP and creating a subject gap. We will now see some evidence that this analysis might be correct.

In fact, the only other contexts where we find embedded linear V3 – apart from those cases involving verb-second inversion, to be discussed in section 4.3.2 – revolve around the NPIs *onques* and *ja*(*mes*). We already discussed their idiosyncratic behaviour in main clauses in section 3.7.1.3. In embedded clauses, these adverbs show a unique distribution in that they can appear in clause-initial position, preceding nominal subjects and giving rise to the string *CSVX*:

⁹It is true that V2 or more generally root phenomena can be found in Germanic in embedded clauses introduced by conjunctions like ‘while’ or its equivalent, but generally only if they have an adversative reading, in which case they are always placed after their matrix clause, cf. examples (46a)–(46b) in section 2.3.5. When the reading is purely temporal, V2 is strongly ungrammatical regardless of the position of the adverbial clause relative to its matrix clause.

- (262) *Vos m' avez osté de la greignor prison ou [onques] [chetis]*
 You me.CL have removed from the worst prison where ever captive
demorast.
 remained.SUBJ
 'You have freed me from the worst prison that a prisoner ever endured.' (Tristan, 52. 7-8)
- (263) *Mes Nostre Sires qui bien le pooit fere, la garda si bien qu' [onques] [li barbarins] n' ot en li ne part ne compaignie ...*
 But our lord who well it.CL can do, her.CL kept so well that ever the
 barbarian NEG.CL had in her nor part nor company ...
 'But Our Lord who has the power to do so, protected her so that the barbarian never had any company or intimacy from her ...' (Eustace, XVIII. 2-5)
- (264) *Et il dient que il le garderont si que [jamés] [nus] n' en orra parler devant ce que il veille.*
 And they said that they him.CL would-keep and that (n)ever somebody
 NEG.CL thereof.CL would-hear talk before this that he wants.SUBJ
 'And they say that they would guard him and that nobody would ever hear any talk of this before he wished it so.' (Tristan, 46. 9-11)
- (265) *Et sachiez qu' il est si preuz de son cors, que [ja] [li rois Pelyas] ne porra longuement a li durer.*
 And know that he is so able of his body, that JA the king Pelias NEG.CL
 can.FUT.3.SG long against him last.
 'And know that he is so skilled that king Pelias cannot last long against him.'
 (Tristan, 51. 15-16)

It was pointed out by Vance (1997) that these are the only phrases which are ever found to the left of the subject in non-inverted embedded clauses (once again excluding V2 contexts). Our corpus confirms this, also giving evidence that they are in principle available in all kinds of embedded clauses. Crucially, the very same adverbs always follow pronominal subjects, giving the string *SpCVX*. No examples of *ja(mes)* were found in this constellation in our corpus, but examples with *onques* are quite numerous :

- (266) *...je i ving par une dé merveilleuses aventures que [vos] [onques] oïssiez ...*
 I here came by one of marvelous adventures that you ever
 hear.IPFV.SUBJ
 'I came here by one of the greatest miracles that you will ever hear ...' (Tristan, p.50 : 28. 17-18)
- (267) *...li rois Pelias, qui bien conoist que ses compainz est le meillor chevalier que il onques trovast ...*
 ... the king Pelias, who well knew that his compaignon is the best knight
 that he ever find.IPFV.SUBJ
 '... king Pelias, who knew well that his compaignon was the greatest knight that he would ever find ...' (Tristan, p.60 : 57. 4-5)
- (268) *Sachiez que vos l' avrez, se [je] [onques] puis.*
 Know.IMP that you it.CL will-have, if I ever can.

‘Know that you shall have it, if I can ever help it.’ (Tristan, p.67 : 73. 7-8)

The reader might wonder why this is necessarily anything different from the already reported cases of SF without a subject gap. In one sense, there is no difference, as the strings in (266–268) are identical, but appealing to SF here would leave completely unexplained the examples with nominal subjects in (262–265), since no other cases of SF with preverbal nominal subjects are found. Yet, there is reason to believe that the structural position of *onques/ja(mes)* is in fact exactly the same as those occupied by SF-fronted elements, as has also been suggested by Ingham (2013). We have already seen that this position is below FinP, yet above IP. The difference in position between nominal and pronominal subjects further corroborates this, enabling us to state the following unmarked position for the subject in Old French:

Conclusion: The position of the subject in non-inverted main clauses in Old French is Spec-IP. In inversion structures and embedded clauses, pronominal subjects cliticize to C^0/Fin^0 .

In between FinP and IP there must be a projection which I will simply call SFP. In order for an XP to reach this projection, Spec-IP must be empty (the ‘subject gap requirement’). The NPIs *onques* and *ja(mes)* can circumvent the subject gap requirement, reaching this projection even in the presence of an overt nominal subject in Spec-IP.

4.3.2 Embedded linear V3 in verb-second contexts

We will now consider the other source of embedded linear V3, namely those strings resulting from embedding the same constructions that were discussed in section 3.7 of the previous chapter. Terminology becomes essential here; one must avoid talking about ‘embedded verb-second’ in these cases, since it was argued that most of these strings (more specifically, the *CS(p) VX* strings) do not feature V-to-C movement at all, but rather a normal IP preceded by a scene-setter. Indirectly, however, the existence of these strings in embedded clauses still shed some light on the structure of verb-second inversions, and in particular they pose serious problems to the recently developed ‘Force-V2’ hypothesis, which we will now briefly review.

4.3.3 The Force-V2 hypothesis

In Wolfe (2015), an analysis was developed whereby the relatively strict linear V2 order of Old French was explained by postulating a high locus of verb movement in main clauses. Whereas most Old Romance languages featured verb movement to the lowest head of an articulated left periphery, Fin^0 , Wolfe argues that Old French (together with Old Spanish and Old Venetian) had developed a syntax with verb movement to the high projection ForceP. This difference is taken to account for the long-observed difference between so-called ‘relaxed V2 languages’ where linear V3 and V4 orders are not uncommon, and the ‘strict V2 languages’, where only a rather restricted set of V3 contexts are possible. Descriptively speaking, my corpus has confirmed that Old French does in fact pattern this way and that linear V3 orders are both highly restricted and predictable.

In order to make the analysis work and to solve some theory-internal problems (see section 2.4.4), Wolfe had to derive verb-second inversions by appealing to ‘bottleneck effects’. Since these aspects of the analysis were thoroughly discussed in chapter 2, I will not repeat

the details here. Rather, I will focus on the empirical part of the analysis and its predictive power when confronted with the evidence from the corpus.

The first thing to notice is that Wolfe, building on work by among others Poletto (2002) and Beninca and Poletto (2004), accounts for V3 orders in main clauses by assuming a FrameP above the Force projection. This projection is in itself a shorthand for at least two different projections explicitly discussed by Wolfe, namely a projection hosting initial scene-setters, thereamong embedded adverbial clauses, and a projection hosting Hanging Topics. Since it is assumed that these elements may be first-merged in the left periphery rather than moved there, they are exempt from the blocking effects of the 'bottlenecks'.

The major empirical problem with this analysis comes from embedded data. As I interpret it, Wolfe's analysis makes two distinct predictions regarding embedded clauses, both of which turn out to be incorrect, although in completely opposite directions.

First, Wolfe assumes that normal embedded clauses feature a complementiser in Fin⁰. This is uncontroversial and explains why embedded V2 is not generally available. It is also the assumption adopted in this thesis. When faced with cases of embedded V2 in complement clauses, which are found in Wolfe's data under exactly the same matrix verbs as those reported from our corpus, he suggests that these feature a high complementiser in Force⁰, opening up the lion's share of the left periphery for XP movement. Still, since the complementiser sits in Force⁰, the finite verb cannot move there and must content itself, so to speak, with the lower position in Fin⁰.

In such a system, this would seem to give rise to the prediction that such complement clauses should feature 'relaxed' V2 syntax with perhaps non-negligible amounts of linear V3 and V4. At the very least, it should be possible to have the order: *Complementiser - topic - focus - verb*. These orders are neither attested in Wolfe's data or in our corpus. Notice that it will not work, within the set of assumptions adopted by Wolfe, to appeal to Locality Effects to explain the absence of such patterns, for instance by saying that the bottleneck in Spec-FinP blocks movement of a topic when a focus has already moved to Spec-FocP via FinP. The reason is that these orders are liberally attested in the relaxed V2 systems that Wolfe examines, leading him to suggest that either the topic is first-merged in the LP after the bottleneck has been closed by the focus, or that foci somehow do not count as interveners for topics from the perspective of Relativized Minimality. It matters little presently if any of these suggestions is correct, the point is that one would expect Old French embedded clauses to 'revert' to such a relaxed V2 syntax once the verb can move no higher than Fin⁰. But this prediction is not borne out; Old French embedded V2 is not more 'relaxed' than main clause V2.

The other problem is in essence the exact opposite. Being situated in the very highest portion of the left periphery, above ForceP, the Frame-field consisting of at least Hanging Topics and scene-setters should not be available in embedded clauses at all, since the high complementiser is assumed to sit in Force⁰. But this prediction does not turn out to be accurate, either. In particular, initial subordinate clauses of various kinds are quite liberally found in embedded clauses in *Tristan*, far too frequently (18 tokens) to brush them off as parentheticals or the like:¹⁰

¹⁰Nor is there any reason to assume that these cases involve parentheticals, since it is perfectly fine to truly embed subordinate clauses in modern Germanic as well. Yet there are some paradoxes at play here which almost resemble 'transitivity failures'. Bringing in data from main clause interrogatives and imperatives, which have been disregarded in this investigation, we observe that it is possible to encode the illocutionary force after an initial subordinate clause:

- (269) *Saiches que [se li oirs dont ele est encente vit longuement], [il] t'*
 Know that if the heir of-which she is pregnant lives long, he you.CL
ocirra.
 will-kill.
 'Know that if the heir that she is pregnant with lives long, he will kill you.' (Tristan, p.47: 22. 4-5)
- (270) *En ceste partie dit li contes que [quant li marinier orent gité Sador en la mer], il s' en alerent et il remest ...*
 En this part says the story that when the sailors had thrown Sador in the sea, they REFL.CL therefrom.CL went and he remained ...
 'Here the story tells that when the sailors had thrown Sador overboard, they departed and he remained ...' (Tristan, p.49 : 27. 1-2)
- (271) *Aprés ce se porpense que [s' il ocist cest roi, qui par sa cortoisie l' amena en sa meson] ... [ce] sera la greignor traïson ...*
 After this REFL.CL thinks that if he kills this king, who through his courtesy him.CL brough in his house ... it will-be the worst treason ...
 'Then he thinks to himself that if he kills the king, who in his courtesy has invited him into his home, it will be the worst treason ...' (Tristan, p.53 : 37. 13-15)

As for left-dislocated constituents (LDs), these are also found in embedded clauses, although not very frequently. Just like in main examples clauses, it is not possible to decide if these are Contrastive Left Dislocations (CLDs) or Hanging Topics (HTLDs). In the absence of any clear evidence, the very fact that they are embeddable suggests that they might be CLDs, as Hanging Topics are often considered to occupy a very high position in the left periphery, above ForceP (Poletto 2002; Benincà and Poletto 2004). If one adopts the assumption that they are in fact CLDs, their presence in embedded clauses is compatible with the V-to-Force analysis. However, this in turn causes problems in main clauses, since the same kind of expressions also turn up to the left of the V2 construction there (see section 3.7.2 and 3.7.2.1), causing linear V3. There is certainly no evidence for claiming that these LDs are Hanging Topics in main clauses, but CLDs in embedded clauses. The following (272) is the clearest example of an embedded LD:

- (272) *Tex est la vie e tex est la fins del bonaure saint Eustace e de ses compaignons, e bien sachiez que [tuit cil qui l' avront en memoire] [..] [il] avront hastif conseil ...*
 such is the life and such is the end of-the blessed saint Eustace and of his companions, and well know that all those who him.CL will-have in memory they will-have speedy counsel

(i) *[Se je m' en is], [ou] porroie je aler...?*
 If I me.CL herefrom.CL go, where could I go?
 'If I leave, where could I go?' (Tristan, p.44 : 16.6-7)

(ii) *... [se tu les reconois], di le moi...*
 If you them.CL recognze, say.IMP it me...
 'If you recognize them, tell me... ' (Eustace, p.35 : XXXI. 1-2)

On the standard assumption that illocutionary force is encoded in ForceP, these patterns on the one hand provide support to Wolfe's hypothesis that the position of initial subordinates precede ForceP, but on the other hand it is all more surprising to find them in embedded clauses (see (269) and (271)).

‘Such is the life and end of blessed Saint Eustace and his companions, and know that all those who remember him [...] they shall receive prompt guidance...’ (Eustace, p. XXXIX. 1-6)

There are also some cases where an LD co-occurs with an embedded clause. When this happens in a main clause, the order is always *LD-embedded clause*. The following example seems to suggest that the order can be reversed in embedded clauses:

- (273) *A celi tens avoit une costume en Gaule ... que [quant li home estoient pris*
 At that time had a custom in Gaul ... that when the men were taken
en aucun mesfait ou il eüssent deservi mort], [qui que ce fust], se ce
 in some misdeed where they had deserved death, whoever that it was, if it
fust rois meïsmes, [si] ne fust il par esparniez ...
 was king himself, SI NEG.CL was he PAR saved ...

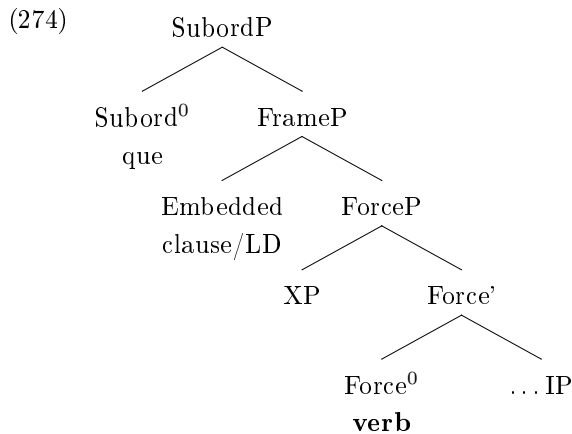
‘At that time there was a custom in Gaul, that whenever men were caught in misdeed for which they had deserved the death penalty, no matter who it was, even if it was the king himself, he would truly not be spared ...’ (Tristan, p.66 : 70. 1-4)

Admittedly, this passage is very dense with several complicating factors, such as the fact that the LD constituent is itself a free relative clause, there is much intervening material, and the resumptive pronoun is not in the prefield of the ensuing main clause, but rather postverbal in an inversion structure introduced by the particle *si*. There are other candidate structures of comparable or greater complexity; on the whole, the evidence is neither quantitatively robust enough or qualitatively clear enough to allow us to say that LDs may follow subordinate clauses in embedded left peripheries, so there is no strong case against the Principle of Transitivity, although the matters deserves more attention. What seems clear, on the other hand, is that left-dislocated phrases can in fact be embedded, and it is beyond dispute that subordinate clauses can themselves be embedded. This means that the combined evidence from main and embedded clauses is incompatible with the idea that verb moves to Force⁰.

To recapitulate the essentials: both embedded clauses and LDs of some kind can precede the verb in main clauses. If the verb is in Force⁰, this entails that these constituents must lexicalise projections above ForceP. In consequence, they should not be embeddable, since the highest complementiser must be assumed to be in Force⁰ (see also section 4.3.3.1 for evidence that it does not help to postulate an even higher complementiser). As this section has demonstrated, this prediction is not borne out. The facts once again strongly resemble the situation in modern Germanic. As it stands, the Force-V2 analysis does not make the right predictions, neither for modern Germanic, nor for Old French.

4.3.3.1 How to fix it

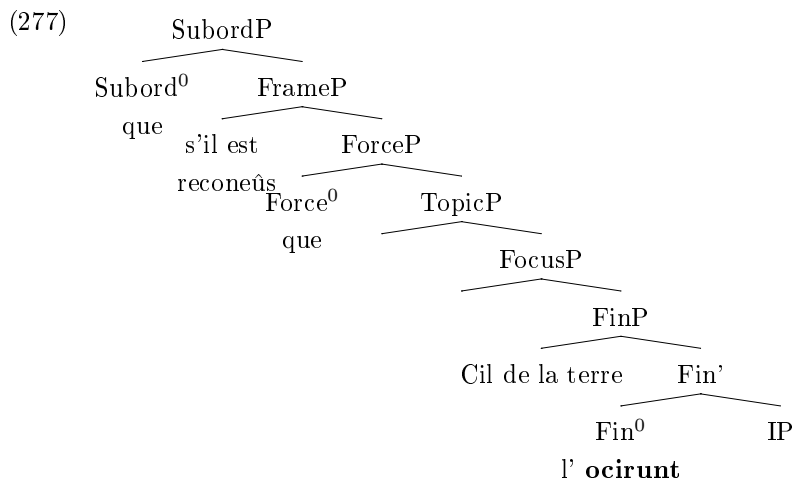
The question immediately arises as to how these problems can be addressed theoretically. I can at least two possible solutions. The first one would be to simply postulate an even higher projection for the complementiser in such clauses, perhaps a SubordP above the FrameP:



This is theoretically very unattractive, since it conjures up a third homophonous complementiser for which there is no evidence – the patterns discussed here can hardly be considered as such – and thereby dilutes the predictive potential of the cartographic approach to the Left Periphery. Unless one assumes that the complementiser is first merged in Force⁰ and then moves to Subord⁰, a solution which again just grafts the facts forcefully onto the model without any justification, this solution would also completely drop the idea that there is a competition between the complementiser and the verb at all for the position in Force⁰, undermining a central claim of Wolfe’s analysis. More seriously, there is direct evidence against this hypothesis. The relevant data are provided by cases of so-called ‘recomplementation’, where the embedded left periphery displays two overt complementisers, one on each side of an embedded subordinate. In the following examples, the complementisers are underlined.

- (275) *Je vos comant [...] que [maintenant que li enfes, qui de la roïne
I you.CL command that now that the child who from the queen
istra, sera nez], que [vos] le m’ **aportez** ...
will-come-out will-be born that you it.CL me.CL bring*
‘I command you [...] that, as soon as the child of the queen is born, you bring it to me.’ (Tristan, p. 47: 23. 3-5)
- (276) *... quar il pense que [s’ il est reconeüs], que [cil de la terre] l’
for he thinks that if he is recognized that those of the land him.CL
ocirunt ...
will-kill
‘... for he thinks that the people of the country will kill him...’ (Tristan, p.63: 64, 2-3)*

Such double complementisers, of which there are many instances in the corpus, à priori provide neat evidence in favour of a cartographic approach to the left periphery. However, in order to reconcile this phenomenon with the structure in (274), we are forced to construe the first of them in Subordinate⁰ and the second in Force⁰, like this:



But if that is the case, we still have several left-peripheral A-bar positions available below the lowest complementiser, and we would expect to find inversion structures here, contrary to fact: every single instance of recomplementation has the subject following directly after the lowest complementiser. The exact same state of affairs is reported in Salvesen and Walkden (2017) and Wolfe (2015), providing strong evidence that the lowest complementiser must sit in Fin^0 , not Force^0 . This is also the consensus view in the literature on Romance recomplementation (Ledgeway 2005; Paoli 2007; Villa-García 2012). We can therefore reject the solution based on a third complementiser above ForceP .

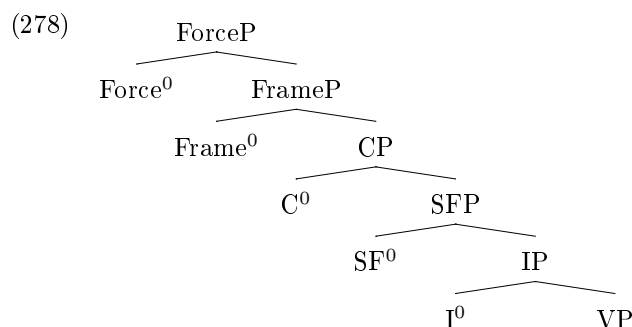
A more natural solution would be to suggest that the structure of the left periphery in Old French must be somewhat different from what is assumed in Wolfe (2015). The Frame-field consisting of scene-setters must simply be situated below ForceP . This gives a better description of the facts, but it also has a clear consequence, namely that the verb in Old French never moves as high as Force^0 . Rather, the evidence clearly seems to indicate that Fin^0 is the appropriate locus of verb movement. This is a very welcome result from the perspective adopted in this thesis, since the SSAP-principle states that only the minimal structure will be constructed to account for the various strings. However, assuming that Wolfe's analysis of the 'relaxed' V2 systems in Old Romance is correct, the corollary is that Old French does not distinguish itself through any higher verb movement than its sister languages. The relative strictness of the linear V2 rule in Old French must therefore be accounted for in some other way. One can either accept that there might be some linear constraint at play after all, or one can imagine a partially collapsed or 'syncretised' CP in the sense of Hsu (2017). I will adopt this latter solution, thereby rejecting (the universality of) the strong cartographic tenet of 'One-Feature-One-Head.'

4.4 A formal analysis of Old French clausal syntax

Having reviewed the syntax of both main and embedded clauses in considerable detail, we are now in a position to develop a concrete formal analysis of Old French clausal structure. I will focus on the representation or 'end result' which represents the competence acquired by the learner.

The global evidence from main and embedded clauses leaves no doubt that Old French featured V-to-C movement as a pervasive feature of the language. In this sense, Old French

was a ‘V2 language’, although I repeat from the introduction that this concept seems to me of limited theoretical interest. The evidence points towards the very lowest position in the Left Periphery: a single, syncretised A-bar projection able to host a great variety of different constituents with all the different IS-readings normally associated with the Left Periphery in cartographic work. Although I have been referring to this position as FinP, it is clear now that it is rather a feature bundle containing many different features, and FinP is therefore not an appropriate name for it. Rather than calling this projection something cumbersome like ‘Scene-setterP/TopicP/FocusP/FinP’, or something opaque like ‘FP’, I will simply call it by its traditional name, CP. Above this projection, there is another projection, able to host initial subordinate clauses and LD structures. Perhaps it would be most correct to name this projection ‘LDP’, since all elements here are dislocated in some sense, I retain the term FrameP. This makes for the following representation:



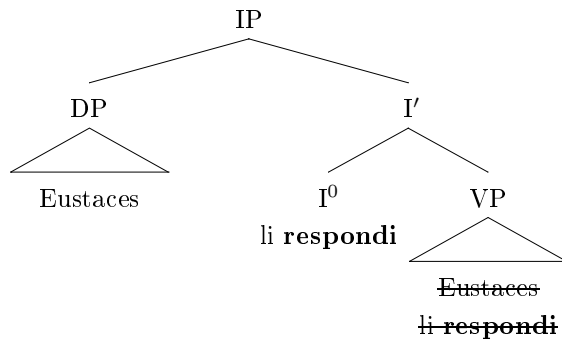
I believe this much, but no more, is required by the evidence. If and only if it turns out that initial subordinate clauses and left-dislocation structures (LDs) can co-occur in the left periphery, it will be necessary to consider FrameP a shorthand for two projections. The evidence from our corpus was not conclusive in this respect, since the possible candidate examples were few, highly complex and ambiguous, arguably featuring parentheticals rather than lexicalising clausal projections. Given the structure in (278), I will now briefly run through the structural underpinnings of the major string patterns in the corpus. I will leave aside all strings featuring ‘Romance inversion’, in other words strings where nominal subjects follow non-finite verbs, although I emphasize that the position of the verb is the same in these strings as in normal V2 inversion structures. The difference is either related to a particular position of the subject, or alternatively, to a short scrambling movement of the non-finite verb above the subject in Spec-vP.

4.4.1 Main clause strings

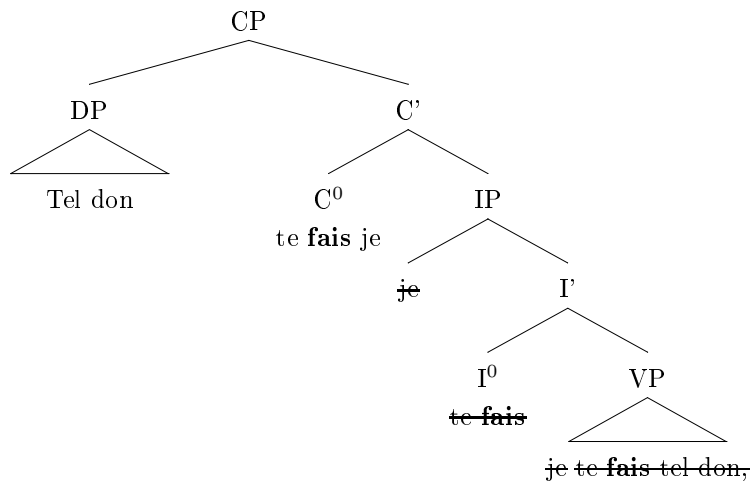
The unmarked word order of Old French is the one represented by the subject-initial S(p)VX strings. Stylistic Fronting provided evidence that these are in fact mere IPs, in accordance with the suggestion offered by Vance (1997). This can also be considered evidence in favour of well-foundedness of the SSAP-principle. In contrast, inversion strings must be analysed as involving V-to-C movement and topicalisation of an XP to Spec-CP. The exact structure of the inversion string depends on whether the subject is pronominal or nominal. This gives the three different structures depicted in (279–281):¹¹

¹¹‘VP’ in the following structures is used to refer to the vP/VP complex, the structure of which is only spelled out properly where necessary.

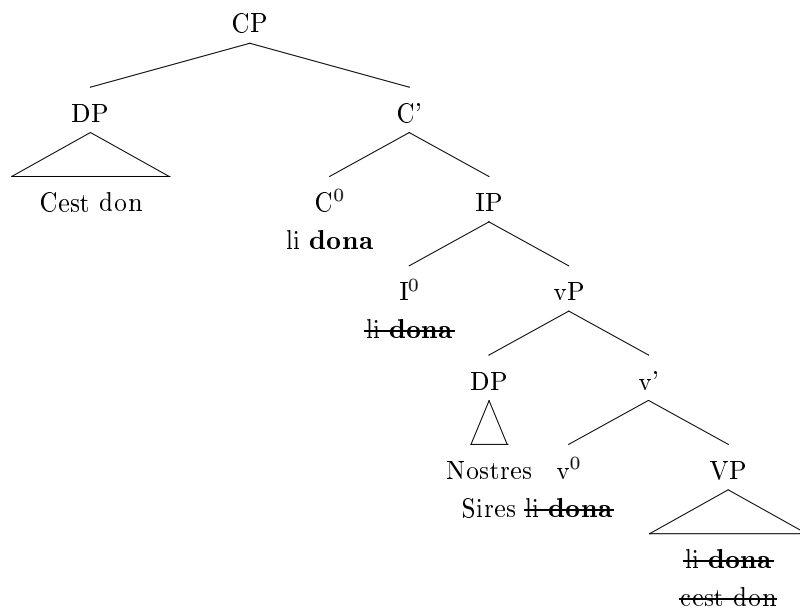
- (279) [*Eustaces*] *li respondi...*
 Eustace him.CL answered...
 'Eustace replied to him...' (Eustace, p.13 : X. 5-6.)



- (280) [*Tel don*] *te fais je, biaux amis.*
 Such gift.ACC you make I good friend.
 'Such a gift I give to you, my good friend.' (Tristan, p.40: 2. 23)



- (281) *Car [cest don] li dona Nostre Sires...*
 For this gift.ACC him.CL gave Our Lord...
 'For Our Lord gave him this gift...' (Eustace, p.45 : XXXIX. 7-8)



Pronominal subjects (280) are always adjacent to the verb in inversion structures (*CVSpX*). This fact alone could be accounted for by construing the pronominal subject in Spec-IP. But in fact, the apparent incapacity of the postverbal pronominal subject to participate in coordination structures suggests the bond between the verb and the pronominal subject is even tighter. In addition, the position of the pronominal subject relative to the NPIs *onques/ja(mes)* in embedded clauses indicates that pronominal subjects may in fact cliticize to C⁰, a suggestion that was already put forth by Vance (1997). Note also that the structure in (280) is the configuration where *pro-drop* is licensed as an alternative to overt subjects, in accordance with Foulet's Generalization.

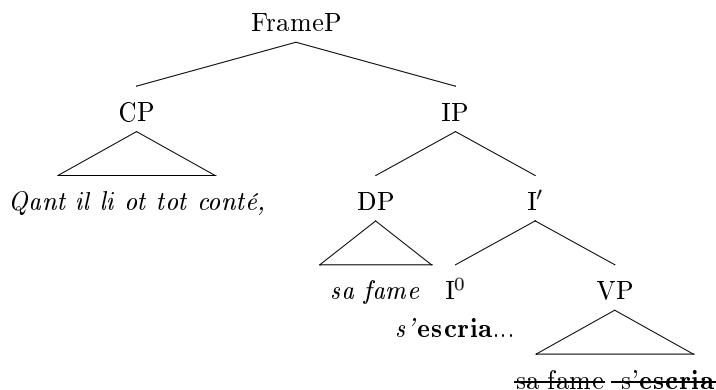
The hypothesis that S(p)VX strings do not feature V-to-C movement might also go some way towards explaining why initial subordinate clauses: (1) generally do not trigger inversion, and (2) are almost invariably followed by a non-inverted main clause.¹² I suggest that V-to-C movement is triggered by topicalisation to Spec-CP, and that initial subordinate clauses normally fail to be analysed as a case of topicalisation, thereby leaving the CP unprojected. This seems more satisfying than having to deal with the conundrum of why there should all of a sudden be (near-)obligatory topicalisation of the subject whenever the clause opens with an initial subordinate clause.¹³ It should be emphasized again that it is not initial scene-setters *per se* that fail to trigger inversion, but rather *initial subordinate clauses*: the syntactic category *CP* is not fully integrated into the V2 inversion pattern.

¹²I remind the reader that these two claims are subtly different: the first is a negative claim, saying that the subordinate normally does not trigger inversion (the string *CVS(p)X*). The second is a positive claim, saying that the subordinate is normally not followed by an inversion structure afterwards either; in other words, we normally get *CS(p) VX* rather than *CCVS(p)X*; See sections 3.7.3 and 3.7.4 for details.

¹³Seen in a derivational perspective, this conundrum might even be more challenging, since the merger of the embedded clauses must be considered the final operation in a bottom-up Minimalist derivation of the clause; the near-obligatory topicalisation of the subject to Spec-CP therefore also involves some strange kind of syntax-internal look-ahead. An alternative would be to suggest that, in the absence of topicalisation, an EPP-feature on C⁰ is capable of triggering 'Formal Movement' in the sense of Bhatt (1999) or Frey (2004), blindly attracting the nearest XP (the subject) to Spec-CP. In either case, the fundamental evidence remains that the word order facts of Stylistic Fronting reveal that subject-initial clauses do not feature V-to-C.

Initial scene-setters with any other categorial status quite consistently move to Spec-CP, triggering inversion. As was pointed out in section 3.7.5, *Eustace* is not as robust as *Tristan* in this respect and sometimes features other categories in FrameP, in violation of the V2 grammar:

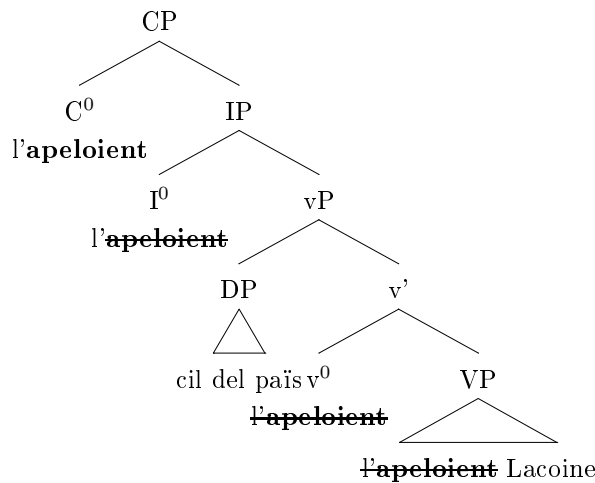
- (282) *Qant il li ot tot conté, [sa feme] s' escria e li dist...*
 When he her.CL had all told, his wife REFL.CL cried and him.CL said...
 'When he had told her everything, his wife cried out and said to him... ' (*Eustace*, p.8 : VI. 3-4)



The unavailability of these constructions in normal embedded clause is therefore not the result of the verb being denied access to the Left Periphery, but simply because the complementiser in C⁰ closes off access to FrameP. Of course, in the odd case where the initial subordinate clause does in fact trigger inversion, one must assume that it has moved to Spec-CP, just like it is reasonable to suppose that Spec-CP is the position of resumptive particles/adverbs like *si* or *adonc*.

As for *et/ne-V* clauses, it is unclear to me whether they should all be treated alike. Some of them might perhaps feature 'topic-drop' or some kind of loco-temporal expletive in Spec-CP, but as the discussion in 3.6 should have made clear, there are numerous examples which cannot be analysed in this way. For such strings, I suggest the following structure:

- (283) *Et estoit li chastiax mout forz et mout bons, et l' apeloient*
 And was the castle.NOM very strong and very good and it.CL called
cil del país Lacoine.
 those of-the land Lacoine.
 'And the castle was very strong and good, and the people of the land called it Lacoine.' (*Tristan*, p.53 : 36.9-10)

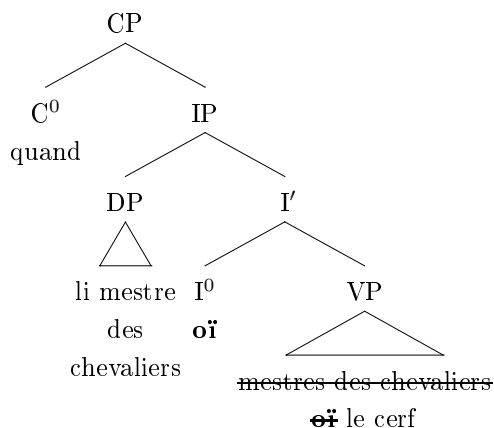


I hypothesize that this is a very old construction, a remnant of a verb-initial stage that has been hypothesized for Late Latin (Salvi 2004; Ledgeway 2017) and which we shall see more of in the Late Latin data in chapter 5. It stems from a time when there was no subject position in Spec-IP, and is perhaps historically related to the verb-initial grammar of Old Sardinian as reported in Wolfe (2015). In a sense, it is a parallel to the verb-initial clauses still found in some Germanic languages, which have also been convincingly analysed by Öttenfors (1997) as historical remnants of a very old pattern. Although it is tempting to speculate that this construction might have found little favour in spoken language in the 13th century, there is no way we can tell for sure. It seems reasonable to assume that it represented some kind of island, a part of the ‘periphery’ of the grammar, and not yet entirely suppressed by the new SVO/V2 system.

4.4.2 Stylistic Fronting in main clauses

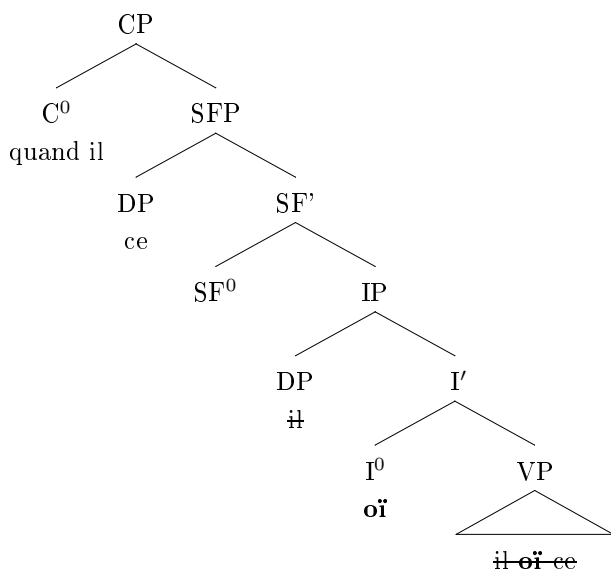
It was demonstrated in the previous chapter that Foulet’s Generalization is robust and generally gives accurate predictions in main clauses. Exempting some hard cases like *et/ne-V* clauses or clauses with initial *onques/(ja(mes))*, null-subjects are generally only possible in postverbal position. However, there is reason to believe that this only applies in full to referential subjects, as impersonal predicates occasionally feature fronting operations which have a distinct flair of Stylistic Fronting, witness the following:

- (284) *Ta fuie ne te vaut riens, dealeal feme. [A morir] te*
 Your flight NEG.CL you.CL is-worth thing, disloyal woman. To die you.CL
covient après ton lecheor.
 suits after your adultery.
 ‘Your escape will not help you, unfaithful woman. You deserve to die after your
 adulterous act.’ (Tristan, p.65 : 67. 23-24)



The presence of an overt subject in Spec-IP also blocks Stylistic Fronting, but if this subject is moved to the Left Periphery, as is the case in relative clauses, or if non-referential subjects are left unexpressed, SF can optionally occur. In some cases, SF can even occur in the presence of an overt subject. This apparent violation of the ‘subject gap requirement’ only takes place with pronominal subjects, suggesting that in such cases, the pronoun has cliticized to the complementiser (286). This analysis is supported by the evidence from the NPIs *onques/ja(mes)*. These are the only elements in our corpus which are capable of preceding nominal subjects in normal embedded CPs (287), but yet they consistently follow pronominal subjects (288):

- (286) *Eustaces, quant [il] [ce] oï...*
 Eustaces, when he this heard...
 ‘When Eustace heard this...’ (Eustace, p.24 : XXI. 7)

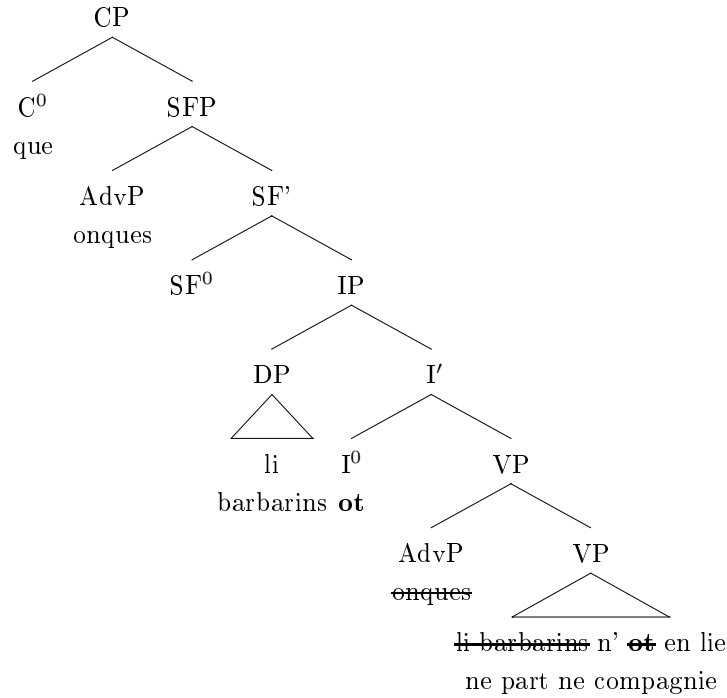


- (287) *Mes Nostre Sires qui bien le pooit fere, la garda si bien qu’ [onques] [li]*
 but our lord who well it.CL can do, her.CL kept so well that ever the

barbarins] n' ot en li ne part ne compaignie...

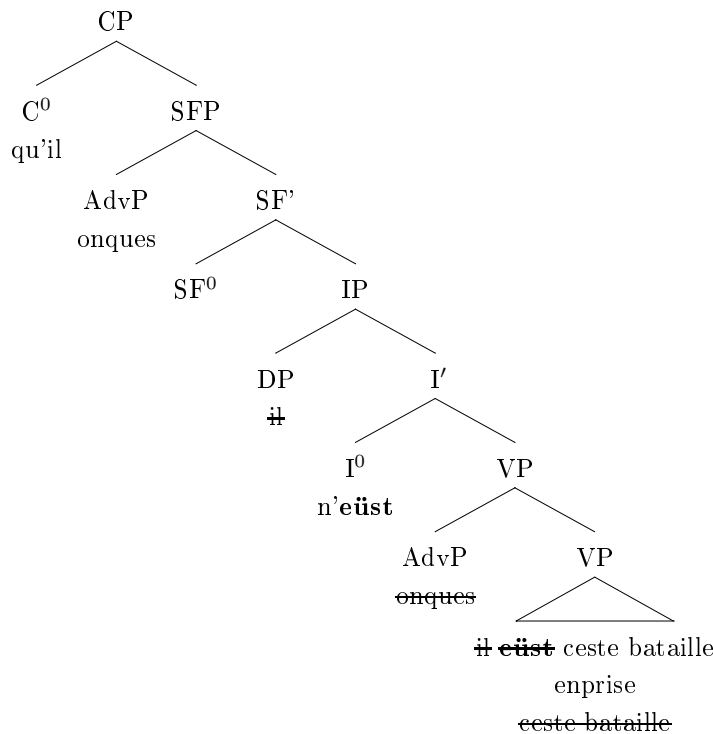
barbarian NEG.CL had in her nor part nor company

'But Our Lord who has the power to do so, protected her so that the barbarian never had any company or intimacy from her... ' (Eustace, p.21 : XVIII. 2-5)



- (288) *Et por ce voudroit il avoir doné la moitié de son reaume par covent qu'*
 and for this would he have given the half of his kingdom by covenant that
[il] [onques] n' eüst ceste bataille enprise...
 he ever NEG.CL had.SUBJ this battle undertaken

'He would have given half of his kingdom in return for this, that he had never undertaken this battle... ' (Tristan, p.61 : 58. 6-8)



Interestingly, Stylistic Fronting is optional in all cases except when called upon to prevent the clause from starting with the verb.¹⁵ This happens when an expletive subject for some reason is not merged, or alternatively not pronounced, in Spec-IP; in these cases SF seems to be automatically triggered. We also argued that even referential subjects can sometimes be left out in Spec-IP in embedded clauses with the same result. This suggests that Foulet's Generalization is not entirely without exceptions in embedded clauses.

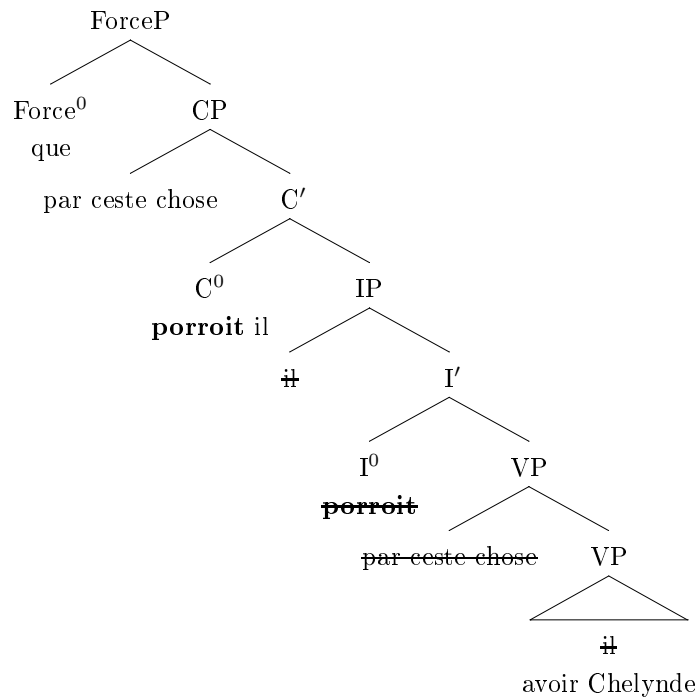
4.4.4 Embedded V2 and embedded V3

Rounding off this chapter, I repeat the important conclusion that embedded V2 is possible in Old French like in modern Germanic in complement clauses embedded under 'viaduct verbs' (Walkden and Booth to appear), in other words classes A, B and E from Hooper and Thompson's (1973) seminal study of embedded root phenomena. These verbs select a high complementiser in Force⁰, thereby giving access to the embedded left periphery and permitting inversion (289). In the same contexts, it is also possible to have embedded linear V3, for instance initial subordinate clauses (290) or LDs followed by an SVX string. It is somewhat misleading to talk about 'embedded V2' in these cases, since it assumed here that these feature a left-dislocated XP on top of a regular, subject initial IP. In other words, the CP projection is not engaged in these structures:

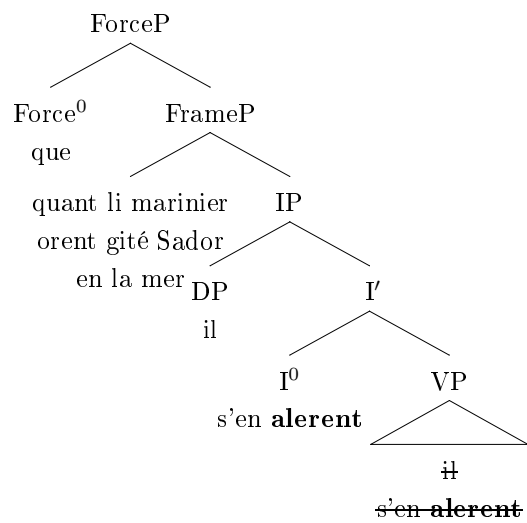
¹⁵In (288), the object 'ceeste bataille' has been scrambled to a position preceding the past participle, as indicated by the strikethrough. This position might well be external to the VP, but for simplicity I leave it here.

- (289) *Naburzadan, quant il vit que ses freres n' estoit mie venuz, il Naburzadan, when he saw that his brother NEG.CL was not come, he en fu auques liez, car il pensa que [par ceste chose] porroit il thereof.CL was quite happy, for he though that by this thing could he avoir Chelynde have Chelynde.*

‘When Naburzadan saw that his brother had not come, he was very happy, thinking that through this chance he might have Chelynde.’ (Tristan, p.42 : 9. 1-3)



- (290) *En ceste partie dit li contes que [quant li marinier orent gité Sador en En this part says the story that when the sailors had thrown Sador in la mer], [il] s' en alerent et il remest... the sea, they REFL.CL therefrom.CL went and he remained...*
- ‘Here the story tells that when the sailors had thrown Sador overboard, they departed and he remained ...’ (Tristan, p.49 : 27. 1-2)



Chapter 5

Late Latin

5.1 Introduction

In this chapter I will conduct an investigation into the syntax and word order of a Late Latin text.¹ The objective of this investigation is to pursue the hypothesis presented in chapter 1, namely that the frequent use of subject-verb inversion in the Romance languages of the high medieval period might derive from Late Latin syntax. I will test this hypothesis by providing and analysing quantitative and qualitative data from a Latin prose itinerary dating from the late 4th century, the *Itinerarium Egeriae* (henceforth also *Egeria*). Departing from the assumption that this specific text may provide information not only about the spoken language of its author, but about the syntax and word order of Late Latin in general, I will focus particularly on the following research questions:

- (291) a. Had Late Latin already developed subject-inversion structures of the Old Romance kind?
b. Had Late Latin already developed V-to-C movement?
c. Had Late Latin already developed into a V2 language?

These questions are distinct, but not unrelated, since they are phrased in terms of increasing specificity; a positive answer to (291a) makes it possible to consider the stronger hypothesis in (291b) which, if confirmed, in turn makes it possible to consider the hypothesis in (291c). On the other hand, a negative answer to (291a) will automatically entail a negative answer to (291b)–(291c) as well. Importantly, the definition of V2 adopted in chapter 2 makes a clear distinction between (291b) and (291c), making it possible in principle to answer the first in the positive and the latter in the negative.

Phrasing these research questions as polar interrogatives is, however, somewhat naive, as we may expect the answers to be rather quantitative in nature, rather than absolute. For this reason we will modify them into (292). With this modification, the entailment relations

¹I wish to stress that I use the term ‘Late Latin’ for convenience to refer to the spoken varieties of Latin in Egeria’s day and later. It is not meant to signal that this stage of the language is a cohesive entity which is qualitatively very distinct from what preceded it. For a discussion of some of the problems in delimiting ‘Late Latin’ as a specific linguistic entity, see Adams 2011. Although I have followed the common practice of rendering the epithet ‘late’ in ‘Late Latin’ with an initial capital letter, it follows from Adam’s observations that it might be more prudent to simply talk of ‘late Latin’, to avoid giving an impression of internal cohesion or unity that is not warranted from closer scrutiny of the textual evidence.

still hold, but on a quantitative level. This is a welcome outcome, since it makes it possible to provide answers which are largely objective.

- (292) a. To what extent had Late Latin already developed subject-inversion structures of the Old Romance kind?
b. To what extent had Late Latin already developed V-to-C movement?
c. To what extent had Late Latin already developed into a V2 language?

Structure Before we can approach these research questions, there are some matters that deserve our attention. First of all, the quantity and range of textual evidence from Latin is very large and to a considerable extent contradictory. In all periods of the language, there is significant synchronic variation between different texts and across many different variables, particularly with respect to word order. Secondly, over and beyond this intertextual synchronic variation, there is another kind of variation that is a property of the language itself and hence unfolds *within the texts*, namely the phenomenon of free word order. This particular feature of Latin has attracted much attention in the literature and is so relevant to the present study that it would be an omission not to address it all, since free word order must be assumed to have a considerable impact on the acquisition of phrase structure.

These two interrelated issues will be addressed in section (5.2), where I will argue that backward projection from Romance and some other considerations lead to the conclusion that some texts are far better witnesses of the evolution of the language than others. This section also introduces the text chosen for investigation and offers some internal linguistic arguments why this text should be considered trustworthy in this respect; finally, the phenomenon of free word order and its consequences for the upcoming analysis will be discussed.

The rest of the chapter is devoted to a quantitative and qualitative assessment of the syntax of *Egeria*. Sections 5.3–5.5 and 5.6–5.8 focus on the syntax of main and embedded clause, respectively. A short summary is offered in section 5.9; the question of how to interpret the findings in the larger picture of the Latin-Romance diachrony is deferred until chapter 6.

5.2 The transition from Latin to Romance and the nature of the evidence

‘Scarcely any written text can bear much resemblance to the spoken language of its writer.’ (Adams 1976b:11)

With these hearty words of encouragement from Adams, I open this section devoted to a discussion of the textual sources and their relation to the actual object of our study, namely the spoken variety or varieties of Latin that may plausibly be considered the predecessor of the Romance languages.

When investigating the Latin language, there are in general two different kinds of evidence available to the historical linguist: direct evidence in the form of the written sources, and indirect evidence through the insights provided by backward projection from the Romance daughter languages through the so-called *comparative method* (CM); for a nice discussion of the CM and its benefits and shortcomings, see Weiss (2015). While there is no

reason to restrict oneself to only one of these kinds of evidence, it seems clear that direct evidence is preferable in the sense that it can provide us with details and chronology generally unavailable to reconstruction.

More often than not, diachronic investigation has recourse to both kinds of evidence. Concretely, the hypothesis that we seek to explore here is informed by backward projection. It is not reconstruction from the Modern Romance languages, however, as the predominant SVO word order of these would not lead anyone to suspect a V2 stage in their historical development. Yet the Old Romance sources solidly testify to the existence of a rather lengthy historical stage featuring widespread inversion, and it is the generalized character of this inversion pattern across Old Romance that has inspired both the synchronic hypothesis of a V2 grammar in Old Romance as well as the diachronic hypothesis of an internal origin of this V2 syntax within the Latin/Romance family²(see references in the introduction).

We now turn to the written sources for confirmation of this hypothesis, as we seek to ‘connect’ the syntax of Old Romance with Latin. The problem is that a millennium of written Latinity before the first written manifestations of Romance in the Middle Ages has provided us with an abundance of texts, and even if we restrict ourselves to a ‘late’ period, say after 400 AD, we are still forced to make a selection. This is of course not a problem in itself. The real difficulty arises from the texts, as many, if not most of them, do not show the kind of evolution that reconstruction would lead us to expect (Vincent 2000:27). Moreover, although many of the innovative patterns that would ultimately become part of the grammar of the Romance languages can be discerned in the written corpus as a whole, their distribution often does not follow a clear logic or pattern that would allow us trace their evolution in time or space with much accuracy; we can at the very best establish some patterns of distributional statistics (Adams 2011).

5.2.1 Classical Latin and the ‘submerged’ spoken language

This rather surprising inertia in the written material has led philologists to conclude that, at some point in the history of the Latin language, a written standard of ‘proper style’ must have emerged which most subsequent authors have sought faithfully to replicate (Burguy 1869b; Hofmann 1951; Pulgram 1958; Elcock 1975; Panhuis 1984). It is the persistence of this literary standard, what we know as Classical Latin, that obscures the linguistic data by making the real linguistic changes ‘go underground’, in the words of Clackson and Horrocks (Clackson and Horrocks 2007:265). A similar metaphor was employed by Palmer, who spoke of the ‘underground stream of the living language’, hidden beneath an ‘artificial’ and ‘distorted’ literary language. (Palmer 1954/2001:147) Adams also uses strong words, speaking of the ‘gulf . . . between learned written forms of the language and the speech of ordinary people’ (1976:94) and more recently of ‘submerged Latin’ that rarely makes it into written documents preserved for posterity (Adams 2013).³

²I will not engage in the discussion of the utility or lack thereof of such reconstructed entities such as ‘Proto-Romance’, nor in the extremely complicated debate on how the transition from Latin to Romance should be conceptualized, nor if or where it makes sense to place it on a timeline. One prominent view, associated first and foremost with the work of Wright and Banniard, holds that the linguistic situation in Western Romania was one of prolonged and complex monolingualism rather than bilingualism or even diglossia until well into the High Middle Ages, by which time the effects of the Alcuinian reforms and the ‘twelfth century renaissance’ had separated learned *latinitas* and spoken vernaculars so much, both in practice and in the minds of the people, that the way was ready paved for the new emergent national scripta. (Banniard 1992; Wright 2002). See also Varvaro (2013) for a mostly negative appraisal of the theory.

³A closely related (but not coextensive) issue is the hypothesis, originally proposed by Marx (1909), that many of the particular linguistic features of Late Latin are the unbroken continuation of early spoken

As a case in point, which is also highly relevant to the present dissertation, consider the change from SOV to SVO. There is considerable consensus that the modern Romance languages are basically SVO, although there is some variation across the family regarding the possibility of other orders (see the various contributions in Harris 2000b), and it is often assumed, more controversially this time, that the basic unmarked word order of Latin in the Classical period was SOV (Linde 1923; Hofmann and Szantyr 1965; Vincent 1977:56-58; Harris 1977:36)⁴ This change has generally taken place by the time of the emergence of texts written in the different Romance vernaculars in the High Middle Ages, as little if anything remains at that time of the verb-final pattern.⁵ It has become common wisdom in handbooks and manuals on Latin and Romance diachrony to assume a steadfast decrease of verb-finality in favour of the emergent SVO order, (Linde 1923, Lehmann 1972:272, Harris 1978:7, Anderson and Rochet 1979:353, Bauer 1995:7) often illustrated by citing relevant texts from archaic and late Latin, respectively, as evidence of the change (Adams 1976a:93). However, Danckaert's (2017) recent and thorough diachronic study of the OV/VO-alternation, based on a corpus of 39 Latin prose authors from 200 BC to AD 600, draws a quite different picture, as the surface VO/OV alternation is relatively stable diachronically when all clauses are taken into consideration without distinction. The picture is one of significant variation at all times between different authors and different texts. It is revealing that the oldest texts in Danckaert's corpus, the comedies of Plautus (ca. 200 BC) show much higher proportions of SVO than the latest text, the *Historia Francorum* of Gregory of Tours (ca. AD 590) (Danckaert 2017b:112). Indeed, Adams (1976) has shown that VO orders were common in substandard registers from a very early date, and in a more general vein, that many morpho-syntactic features often subsumed under the rubric of 'Late Latin' indeed turn up very early in the written material, and that it is rather their frequency that increases over time (Adams 2011). It is unclear how much of this increased frequency derives from actual linguistic evolution, and how much is the result of a gradually less rigorous normative standard that allows spoken language to 'surface'.

It should be immediately clear why the existence of a literary standard, which even seems to have been particularly perseverant in the case of the SOV pattern, poses serious problems to the current investigation, which aims to reveal the development in spoken language of the placement of the finite verb. However, the silence of the sources should not cast any doubt on the historical reality of the development. As far as I am aware, we have no similar reason to distrust the Old Romance sources, which unmistakably show a placement of the finite verb that is not only very different from that of Classical Latin, but also very familiar from the modern Romance languages. As regards the development of the finite verb, then,

Latin as evidenced in the plays of early popular playwrights like Plautus, subsequently suppressed by the strong standard of Classical Latin. This hypothesis has been modified or challenged in recent years; see in particular the various contributions in Adams and Vincent 2016.

⁴See Panhuis (1984) for a critique of this view. According to Panhuis, word order in preclassical Latin was governed by a principle of Communicative Dynamism which placed constituents according to a theme-rheme partitioning of the information (Panhuis defines thematic elements as the ones about which the rheme constitutes a comment - in other words more closely to a traditional topic-comment distinction), with thematic and rhematic elements gravitating respectively towards the left and right edge of the clause. Already early on, sometime in the third century BC, the verb-final order was established as a written norm, thereby exempting the verb from the principle of Communicative Dynamism, which continued to operate on the arguments of the verb and other constituents of the clause.

⁵Bauer briefly discusses SOV-orders in embedded clauses in Old and Middle French. (Bauer 1995:110) These orders were common in relative clauses and in early texts, Bauer claims, but since she does not cite any examples, it does not seem unlikely that what she refers to might be instances of Stylistic Fronting, (see 4.2.2) which should be kept distinct from real, head-final SOV orders. The same applies to the subordinate SOV orders discussed in Dardel and Haadsma 1976.

the corollary of this state of affairs clearly seems to be that some sources cannot be trusted. Different sources must therefore be sought, even if they are a minority on purely statistical grounds.

5.2.2 Which sources can we trust?

Since the hypothesis we are pursuing in this chapter is that the Old Romance inversion systems were syntactic cognates of a Latin antecedent, it seems only natural that we should seek confirmation of this hypothesis in texts not too far removed in time from the Old Romance languages. Ideally, we would also like to restrict the corpus in space as well, as the Roman Empire included vast areas outside of what was eventually to become the medieval Romance world, areas where linguistic evolution took a different direction and Latin was ultimately lost altogether. Other things being equal, a text from AD 600 is better than a text from AD 100, and a text written by a native of Gaul is better than a text written by, say, a native of Northern Africa who perhaps only acquired Latin as a second language. Guided by these criteria, I have decided not to include in the corpus a text that otherwise might present itself as a natural candidate: *The letters of Claudius Terentianus* are written sometime in the first half of the second century CE by a soldier in the Roman army in Egypt. Of the letters he sent to his father, which were with all likelihood dictated to a scribe, both Latin and Greek specimens survive. Adams suggests his native language was Greek (Adams 1977:3). This is the principal reason for excluding the letters from the corpus, although their great antiquity also plays a role in this respect.

Furthermore, while the corpus should be areally constrained, we do not need it to be local to any particular region; the nature of the ‘pan-Romance V2 hypothesis’ suggests we can include texts from at least all of Western and Southern Romània. This presupposes that we are capable of locating the text at all; the *Mulomedicina Chironis* and the *De re coquinaria*, while in other respects highly interesting from a diachronic perspective, are two such texts that cannot be located properly in either time or space (Danckaert 2017b:85). Besides, both the *Mulomedicina Chironis* and *De re coquinaria* display quite classical word order patterns with very high frequencies of verb-final (Cabrillana 1999).

As already mentioned, the persistence of a literary standard complicates the otherwise rather simple selectional criteria. Judging by temporal and spatial criteria alone, the *Historia Francorum* of Gregory of Tours would seem an ideal choice, written in Latin by a native Gaul shortly before the turn of the sixth century. But although several features of Romance morpho-syntax can be been verified in the text (Hofmann and Szantyr 1965:319-321, Adams 2013:643-644), the placement of the finite verb still strongly adheres to the classical verb-final pattern (Danckaert 2017b). The same applies to other roughly contemporary sources such the writings of Isidor of Seville or the somewhat later *Chronicle of Fredegar*. At the beginning of the ninth century, the *Paderborner Epic*, of unknown but almost certainly Frankish authorship, is written in perfect classicizing Latin (and in verse) and is therefore useless as a source, and the same applies to any text in Latin thereafter.⁶

In the light of these considerations, the guiding principle for selection must be the texts themselves; in our case texts that show the kind of word order that might plausibly be considered a forerunner of Old Romance word order. But since these texts as already mentioned

⁶Indeed, the Carolingian renaissance and the educational reform ushered in by Alcuin finally breaks all bonds between spoken and written; the history of Latin in the following millennium is one of a ‘cultural artefact’, in the words of Clackson and Horrocks (2007:266). The earliest texts in the Romance vernaculars are generally written in verse, and are therefore also problematic, see chapter 6.

might turn out to be a minority, and because it might seem methodologically unsound to consider as true evidence only those texts that would seem to confirm a hypothesis, I will in the next section rather briefly show that the *Itinerarium Egeriae* quite generally evinces a host of morpho-syntactic and lexical properties of unmistakably Romance nature. This suggests that we are dealing with a reliable witness of the linguistic evolution and that we are not entirely unjustified in assuming that this might apply to the position of the finite verb as well.

Although the bigger issue that we seek to explore in this chapter is the historical origins of the inversion structures of Old Romance, in other words a matter of central importance to the diachrony of the Romance family, the method of investigation adopted here is not a diachronic one, since I am not considering the impact of time on the evolution of the position of the finite verb. At the same time, it is not really a synchronic investigation either in anything but a rather trivial sense of the word, since I am only considering a single text. In principle, of course, this text could be recruited as part of a synchronic investigation, an investigation which would need to take into consideration other texts from the same time as well. It is therefore more correct to describe the method employed here as a kind of *selective case study* guided by a specific interpretation of the socio-historical context of composition of Late Latin texts (strongly simplified: some are reliable, others are not). As the preceding discussion should have made clear, this is more than anything a virtue of necessity.

5.2.3 *Itinerarium Egeriae*

The *Itinerarium Egeriae*,⁷ (henceforth also *Egeria*) is written by and recounts the journey of a devout Christian woman, earlier assumed to be a nun, to Palestine, and is generally considered to have been composed towards the end of the fourth century. The text, of which both the beginning and the end are lacking, is preserved in one single manuscript, the 11th century Codex Aretinus from the Montecassino Abbey. The exact provenance of the author is somewhat disputed, the south of Gaul and the north of Spain having both been suggested; for a discussion, see Väänänen (1987:153-157). Einar Löfstedt (1936 [1911]) was sceptical about the possibility of assigning a native country to Egeria on linguistic grounds, claiming she did not represent any particular dialect, but for the purpose of the present study, the general consensus that she was a native Latin speaker of the western Roman world is sufficient (Adams 2007:342).

Egeria enjoys a very special position in Latin-Romance diachrony and has become something of a usual suspect for studies looking for early signs of Romance phonology, lexis or morphosyntax. Several features of unmistakably Romance character have been identified in the text, including but not limited to high proportions of SVO, presentational constructions involving the auxiliary *habere* – to have – (cf. French *il y a* or Spanish *hay* – ‘there is’), occasional overt subject pronouns, and not least an almost consequent tendency to buttress nominal phrases with the determinatives ‘ille’ or ‘ipse’ in seemingly article-like fashion (Adams 2013:512–520).⁸ Due to these features, the consensus view is that *Egeria* is an extraordinary, not to say *unique*, witness of the diachronic evolution. In particular the word order of the text, which differs strongly from Classical patterns, is generally perceived as an

⁷In earlier philology also variably called *Peregrinatio Aetheriae*, *Peregrinatio Egeriae*, *Peregrinatio Aetheriae vel Silvae ad Loca Sancta*

⁸The status of ‘ipse’ and ‘ille’ in *Egeria* has generated considerable debate, see Herzenberg 2015 and references therein.

authentic glimpse of the vernacular of the author. Thomas approvingly cites Wilkinson’s assessment that ‘Egeria wrote much as she spoke’ (Thomas 1981:53). Spevak considers that the text is written in a register ‘close to the spoken language’ (Spevak 2005:1). Clackson and Horrocks, after discussing some caveats in the interpretation of the text, adds that ‘[i]n one area, however, this text substantially increases our knowledge of what is going on ‘beneath the standard’, and that is the order of the major sentence constituents’ (Clackson and Horrocks 2007:291). In a similar vein, Cuzzolin and Haverling express the opinion that *Egeria* is ‘written in a language with almost no literary ambitions, providing us with a lot of interesting information regarding the language of everyday conversation at the time’ (Cuzzolin and Haverling 2009:55).⁹

Crucially, *Egeria* has also featured prominently in the recent debate on the internal evolution of Romance V2 and has been invoked by several researchers as evidence for Late Latin verb-second syntax (Salvi 2004; Clackson and Horrocks 2007; Wolfe 2015c; Ledgeway 2012, 2017). For complete philological discussion of the text, see Lofstedt (1936 [1911] and Vaananen (1987), and for previous studies devoted to the word order, see Hinojo (1986), Spevak (2005) and Ledgeway (2017).

The edition used is that of Maraval (1982); in order to reduce mistakes in manual transmission, the version of the text (the edition of Heraeus 1908) which is available from the PROIEL treebank (Haug and Jøhndal 2008) was extracted and then manually corrected.¹⁰

5.2.4 A worst case scenario: Christian Latin as a ‘Sondersprache’

There exists a possible historical scenario that constitutes a kind of ‘worst case’ for our corpus, and which I must therefore briefly address. This is the hypothesis that Christian Latin constitutes a language apart (*Sondersprache*), a written register that is heavily influenced by the language of the Latin Bible translations in particular and early Christian writings in general. This idea is primarily associated with the Nijmegen scholars Johannes Schrijnen (Schrijnen 1932) and Christine Mohrmann (Mohrmann 1958-1977), and is therefore often referred to as the *Schrijnen-Mohrmann-hypothesis*; see Coleman (1987), Clackson and Horrocks (2007:284-290) and Burton (2011) for discussion. Beyond lexical and idiomatic influence, it is reportedly hard to pinpoint what is Christian from what is just late Latin in general. Little discussion is made of word-order, although one could in principle imagine that the Semitic VSO order which underlies the Old Testament order may have exerted an influence on *Egeria*’s written language through the intermediary of the early Latin Bible translations.

Indeed, in Danckaert’s (2017) diachronic study of the evolution of the word order of the Latin clause, *Egeria* alongside some other ‘Christian’ texts of Late Latin stands out as evincing quite idiosyncratic properties. One could easily imagine that this reflects the fact that these texts were written without any attempt to emulate the norms of Classical Latin, in other words that they are specimens of vernacular language exceptionally ‘surfacing’ in the historical corpus. On the other hand, one could in principle also argue that these texts staunchly follow another standard, namely the standard of Christian Latin as promulgated by the Church and in the Christian communities. Danckaert wishes to distance himself from

⁹Palmer, on the other hand, expresses more reserve, stating that the language of the text is ‘simple and unaffected, but not without some anxious concessions to the grammarians’, pointing out that her written Latin, like other specimens of sub-standard Latin, cannot be considered ‘a true and undistorting mirror of the spoken language.’ (Palmer 1954/2001:163)

¹⁰It should be mentioned that the differences only exceptionally involved word order.

all such attempts at interpreting the material, emphasizing that such hypotheses can never be truly evaluated before they are tested against an explicit, multivariate, statistical model which systematically compares Christian and non-Christian texts (Danckaert 2017b:86).

I absolutely concur with Danckaert's assessment that one should try to construct an explicit statistical model to test the hypotheses of the distorting effects of literary standards, whether Classical or Christian, rather than just taking the latter for granted. However, in the absence of such a model, we have no choice but to make inferences of the traditional kind to make sense of the Latin-Romance diachrony. After all, the Old Romance languages are on the whole much more like their modern descendants than like Classical Latin, and we cannot plausibly assume that all of these profound changes took place overnight during the period that separates, for instance, the end point of Danckaert's corpus (AD 600) and the first written manifestations of Romance. In this chapter, I will therefore inspect one of the most extreme outliers in Danckaert's diagrams, in the hope and belief that its great distance from the regression line plotted by Danckaert is indeed indicative of its unusual reliability as a witness of the linguistic evolution.

5.2.5 Free word order and its consequences

In typological terms, the Latin language could be described as a fusional synthetic, dependent-marking language with rich morphology and great flexibility in word order. In this section I will focus on this latter property, the free word order, a trait which Latin shares with other ancient Indo-European languages and which (in relative terms) distinguishes it from most modern Indo-European languages, including the Romance languages which descend from Latin. As an illustration, consider a transitive verb selecting an internal and external argument. All six possible permutations SVO, SOV, VSO, VOS, OSV, OVS are in principle well-formed (Ledgeway 2012:61-62) and widely attested in Latin texts. Although they are certainly not on even footing in statistical terms, it cannot really be said that any of these orders is truly marginal. For instance, in the corpus used in Haug (2017), all of the six word order patterns exceeded a frequency of 5%. Furthermore, as illustrated by Danckaert, this apparent freedom is maintained when an auxiliary verb is added to the mix, as all 24 possible permutations of the sequence *Subject-auxiliary verb-main verb-object* are indeed also attested (Danckaert 2017b:4-7).

It should also be noted that Latin does not only permit clausal constituents to be arranged in such liberal fashion, but also allows constituents to be broken up, such that heads and modifiers appear in linearly discontinuous positions of the clause. This phenomenon is traditionally referred to as *hyperbaton*; an example from our corpus is given in (293), where the discontinuous subject is underlined:¹¹

¹¹Note that in (293), the subject is discontinuous, appearing both to the left and to the right of the finite verb. This creates a practical problem when counting the linear position of the verb and determining the string type. The solution adopted is to assume that movement is always to the left, and that the fronted element must necessarily target some syntactic position and should therefore be counted as a constituent in front of the verb. The clause in (293) is therefore annotated as linear V3, since both the head of the subject NP, *signa*, and the adverb *ibi* precede the verb. Furthermore, the type is CCVSX, meaning the clause is in fact an inversion structure, since the original position of the subject is to the right of the verb. This approach is adopted for simplicity and consistency, and is of course not the only way to do it; one could suggest with Elerick (1992) that movement can be both left and rightward, and that is is rather the modifier *castrorum* in (293) that has been moved to a right-peripheral position from a subject-initial string. Notice that no matter which approach is adopted, the evidence suggests that both the modifier and the head noun of the NP can be moved, suggesting hyperbaton includes both phrasal movement and head movement, something which in turn could be interpreted to mean that prosody is involved.

- (293) *signa ibi parebant castrorum.* (Egeria, 12.9)
 signs-NOM here be.visible-IPFV-3PL camp-GEN
 ‘The signs of the camp were visible here.’

Although the free word order of Latin never fails to be mentioned in traditional grammars and manuals on the language (Kühner and Stegmann 1955; Hofmann and Szantyr 1965), there is no consensus on how to properly capture this phenomenon on a theoretical level. From a modern linguistic perspective, the crucial question is where in the formal apparatus of any given model the phenomenon should be localized, or put differently, which domain of the overall grammar should be held responsible for the attested variation. To start with a point of consensus, it is widely held that Latin word order is sensitive to the dynamics of discourse, what is now often referred to as information structure. Thus, the word order permutations do not occur randomly, but are motivated by the communicative needs of the situation, what is common ground among speaker and hearer, etc.. In addition, there might of course be prosodic (and in the case of poetry, metric) factors influencing the choice of one variant over another.

The more problematic question is what role this leaves for syntax in the organization of the clause. Ever since the publication of Hale’s extremely influential paper on Warlpiri (Hale 1983), much ink has been spilled on the issue of *non-configurationality*, the hypothesis that certain languages do not exhibit hierarchical phrase-structure; for discussion, see Hale (1989), Baker (2001), and Pensalfini (2004). Since languages which exhibit such pronounced word order freedom are generally morphologically rich case languages, some linguists claim that morphology might take over the role of expressing the important dependency relations that exist within a clause. This view is nicely captured by Bresnan’s dictum ‘morphology competes with syntax’ (Bresnan et al. 2015:5). The syntax of non-configurational languages might therefore lack the phrase-structural articulation that characterizes many languages, employing either exocentric, n-ary branching ‘flat structures’ or a combination of such flat structures and configurationally (phrase-structurally) defined sub-domains of the clause¹². For instance, Hale claimed that Warlpiri possesses a single configurational rule, namely that an auxiliary element must always occupy the second position of the clause (Hale 1983).¹³

Clearly, if syntax does not constrain word order or impose any patterns at all, the observed tendencies must be accounted for in some other component of grammar, for instance in an independent module that relates information structure to linearization. The challenge for a non-configurational approach is to develop a formally explicit model of such a module that makes good predictions; for proposals along these lines, see Panhuis (1984) and Spevak (2010).¹⁴

¹²Danckaert (2017, pp.18-22) briefly discusses such ‘hybrid’ systems combining configurationality at some levels of the clauses with nonconfigurationality at other levels. As an example, he discusses a structure where there is a configurational, endocentric CP on top of an otherwise flat clausal structure. While this model (which is also discussed in Ledgeway 2012:78-80) correctly predicts some observed ordering constraints in the Latin left periphery, Danckaert concludes that it does not capture the existence of higher order constituents like VPs. This is true for the model chosen as an example, of course, but it is also possible in principle to argue for the existence of a VP without necessarily assuming internal structure in that VP (i.e. that it contains several binary branching sub-constituents or ‘shells’). In other words, there are other ‘hybrid’ systems on offer than the example discussed by Danckaert.

¹³This in fact makes Warlpiri a kind of V2-language, but not one that would qualify for the label as defined in this thesis.

¹⁴The notion that Latin might be a nonconfigurational language could also, although anachronistically, be attributed to many traditional philologists, who generally emphasize that word order in Latin is in principle free, but largely determined by discourse factors Weil 1887; Marouzeau 1922.

In transformational generative grammar, these views imply a departure from the standard conception of the T-model, where inflectional morphology is not allowed to interface directly with meaning without the mediation of syntax. To some extent, the same has also applied to notions of information structure, which to varying degrees have been incorporated into syntax itself as so-called A'-projections. This tendency has of course been considerably strengthened with the advent of cartography, which is often described as a 'syntactization of discourse' (Rizzi 2013). Furthermore, exo-centricity and multiple dominance is viewed with skepticism in some generative camps, and in Minimalism, binary branching is (at least without extra assumptions) even dictated by the way Merge works in building up complex constituents.¹⁵

As a consequence, some linguists maintain that Latin like all other languages are configurational, exhibiting the same kind of structural layers as other languages. Evidence in favour of this view is adduced by Ledgeway (2012, chapter 5) and Danckaert (2017, chapter 1) who point out that grammatical processes in Latin make reference to higher order constituents like VPs or IPs (see also Elerick 1992, Oniga (2014, chapter 18) and Oniga and Cecchetto 2014). The challenge for such a fully-configurational view on Latin is first to determine exactly what the syntactic organization of the Latin clause is, and secondly what triggers the displacement processes give rise to the great variety of surface word order patterns.

I should like to point out that although configurational and non-configurational theories might seem to be almost diametrically opposed, the difference between them is sometimes considerably slighter in actual practice.¹⁶ The reason is simply that configurational approaches like for instance Danckaert's are forced to assume a significant number of different syntactic structures to be able to do justice to the surface variation (Danckaert 2017b,a). Unless supplemented with an explicit theory of how these different structures differ systematically at the level of information structure, we end up with a considerable amount of optionality in syntax.¹⁷ If so, the question in the end becomes what is really the difference between a non-configurational approach and a fully configurational approach with high degrees of optionality in the syntax, or how these might be distinguished empirically.¹⁸

To understand what is at stake, consider some examples from *Egeria*, which is replete with near-minimal pairs varying minutely in some aspect of word order. In the following examples, the only kind of highlighting used is underlining to signal the elements of the minimal pair whose relative order is reversed from the (a) to the (b) examples. Witness the

¹⁵It is worth pointing out that Chomsky has explicitly stressed the tentative nature of this hypothesis:

'If Merge is binary, then generated X and Y can intersect only if one is a term of the other. If n-ary operations are added for $n > 2$, other options arise, including those studied in *multidominance* theories [...] *It is in fact likely that binary Merge in its simplest form is insufficient* (italics added), and that some extensions of Merge are licensed by UG, an interesting topic I will not try to pursue here.'(Chomsky 2012:3-4)

¹⁶The hypothesis that the languages of the world displays various degrees of configurationality in the syntax is quite prevalent in LFG and often invoked as an argument for the architecture of the model in that framework (Falk 2001; Bresnan et al. 2015).

¹⁷Such a theory has been developed for instance by Devine and Stephens (2006), but it is based on a somewhat restricted set of data.

¹⁸Perhaps the difference between a configurational grammar with high degrees of optionality and a non-configurational grammar lies in the capacity of the former to refer to higher order constituents like VPs, for instance by VP-fronting operations or by pronominalisation of the VP, phenomena which are clearly attested in Latin. The process of pronominalisation is also blind to internal structure and just targets the VP-node directly.

alternations between adverb and clitic (294), verb and locative PP (295), verb+object and manner PP (296), unaccusative verb and subject (297), transitive verb and subject (298), and subject and dative (299).

- (294) a. *Ostenderunt etiam nobis locum, ubi...* (Egeria, 5.7)
 Show-*PRF-3PL* also us-*DAT* place-*ACC* where
 ‘They also showed us the place where...’
- b. *Nam ostenderunt nobis etiam et illum locum, qui...* (Egeria, 5.7)
 For show-*PRF-3PL* us-*DAT* also and the place-*ACC* which...
 “For they also showed us the place that...”
- (295) a. ... *qui scriptum est in euangelio...* (Egeria, 29.5)
 which-*NOM* written is in gospel-*ABL*
 ...“which is written in the Gospel...”
- b. ... *sicut in euangelio scriptum est...* (Egeria, 29.6)
 as in gospel-*ABL* written is
 “... as is written in the Gospel...”
- (296) a. ... *iam omnis populus et omnes apotactitae*
 now all people-*NOM* and all apotactitae-*NOM*
deducunt episcopum cum ymnis usque ad Anastase. (Egeria, 40.1)
 lead-*3PL* bishop-*ACC* with hymns-*ABL* until to Anastasis
 “and presently all the people and all the apotactitae lead the bishop with hymns to the Anastasis.”
- b. ... *et inde omnis populus usque ad unum cum ymnis*
 and therefrom all people-*NOM* until to one with hymns-*ABL*
ducunt episcopum usque ad Syon. (Egeria, 40.2)
 lead-*3PL* bishop-*ACC* until to Sion.
 “... and from there all the people down to a man lead the bishop with hymns to Sion.”
- (297) a. ... *id est in eo loco, de quo Dominus ascendit in*
 that is in that place-*ABL* from which-*ABL* Lord-*NOM* ascend-*PRF-3SG* in
caelis. (Egeria, 39.3)
 heavens-*ABL*
 “... that is at the place where the Lord ascended into Heaven.”
- b. ... *id est in eo loco, unde ascendit Dominus in*
 that is in the place-*ABL* from-which ascend-*PRF-3SG* Lord-*NOM* in
caelis... (Egeria, 43.5)
 heavens-*ABL*
 “that is at the place where the Lord ascended into Heaven...”
- (298) a. *In ea ergo die et in ea hora, qua auerterant*
 In that day-*ABL* and in that hour-*ABL* when divert-*PLPRF-3PL*
Persae aquam... (Egeria, 19.12)
 Persians-*NOM* water-*ACC*
 “On that day and on that hour when the Persians had diverted the water...”

- b. *Illa autem aqua, quam Persae auerterant.* . . . (Egeria, 19.12)
 That water-NOM which-ACC Persians-NOM
 divert-PLPRF-3PL
 “The water which the Persians had diverted. . .”
- (299) a. . . . *ponitur episcopo cathedra media ecclesia maiore.* . . . (Egeria, 45.2)
 place-PASS-3SG bishop-DAT chair-NOM middle-ABL church-ABL
 major
 “A chair is placed for the bishop in the center of the great church. . .”
- b. *Et statim ponitur cathedra episcopo ad Martyrium in ecclesia maiore.* . . . (Egeria, 46.1)
 and immediately place-PASS-3SG chair-NOM bishop-DAT at
 martyrium-ACC in church-ABL major
 “And at once a chair is placed for the bishop at the martyrium in the great church.”

Granted, for reasons of space not enough context is provided in these examples to determine to what extent information structure might play a role in the observed alternations. Let it suffice to say that the present author is at a loss in finding a proper generalization in terms of discourse properties. Note however that many of the alternations are very close to each other in the text, separated sometimes only by a few lines. This suggests that, at least in some cases, the author is simply seeking some stylistic variation. I believe this is a crucial point, since we should ask ourselves what this suggests about the discourse properties of the elements so manipulated. Take the example of a modern Germanic V2 language. In expressing a series of temporally ordered events to an interlocutor, the language affords some leeway for stylistic variation to the narrator, who can open the clause sometimes with the subject, sometimes with an adverb, sometimes with an expletive. This is possible because these are all, in some sense, unmarked word orders. However, there are clear limits to this variation, as the narrator cannot just place the object or any other non-subject argument in the prefield for the sake of stylistic variation, since the information structural effects that inevitably accompany such fronting operations would lead to incoherence at the level of discourse. In Latin, on the other hand, it seems like these effects are largely lacking or at least much more moderate.

The intention behind these observations is not to plead in favour of a non-configurational approach. The phenomenon under study, V-to-C movement, is a structural phenomenon that tautologically requires structure, and as we have already seen, the syntax of Old French is unmistakably configurationally defined. This means that, at least in descriptive terms, the evolution towards Romance has brought about a rigidification of word order patterns. Presumably, this change would not have come about without the participation of the language acquirers in re-analysing the input. I will therefore assume that children do not resign themselves to the morphological cues alone, but proceed to assign articulated phrase structure to the input strings as well. What is important to emphasize here is not so much the cause, but rather the consequence of the above variation. The input to the child in a language like Latin must have contained substantial amount of *noise* in the form of *structural ambiguity*, meaning that the strings are seemingly contradictory and often not even reconcilable with

a single syntactic structure.¹⁹ Furthermore, if the information structural effects of the displacements were in fact as subtle as they might seem – and we shall see more evidence to support this conclusion when analysing the data – it is not at all inconceivable that the child might conclude that the variation truly belongs in the syntax, and that there is some degree of optionality in the grammar with respect to variables like the position of the subject or other arguments relative to the verb, the headedness of the verbal projections, and the like.

5.3 Itinerarium Egeriae: the main clause

In this section, we will try to establish a concrete hypothesis of the syntax of main clauses in *Egeria*.²⁰ We will proceed in similar fashion to what was done for Old French in chapters 3 and 4, starting out with some rather rough quantitative data and then gradually working our way towards the more detailed kind of quantitative and qualitative evidence that ultimately offers the best probe into underlying syntactic structure. In the remainder of section 5.3, we will consider the linear distribution of the verb and the make-up of the prefield. Section 5.4 addresses the issue of inversion in main clauses, while section 5.5 is devoted to the position of the subject, an issue which will be of the utmost importance to the general analysis of the text.

5.3.1 Linear distribution of the verb and the prefield

It is natural to start by considering the linear distribution of the verb, although the latter provides extremely limited information about the syntax of the clause. This information is given in table 5.1. Observe that the percentages in each column are calculated downwards, so it is possible to see immediately the relative linear distribution of the finite verb within each predicate class. The ‘Total’ column on the right gives the overall linear distribution patterns in main clauses. Under the table, the amount of verb-final strings is indicated.²¹

¹⁹A good illustration of how serious the problem of *structural ambiguity* can be is provided by Danckaert, who claims that a two-word clause consisting of nothing more than the subject followed by a verb is in fact *sevenfold* ambiguous (Danckaert 2017a:126-127).

²⁰Main clauses include clauses which are introduced by a connecting relative phrase, so-called ‘pseudo-relatives’ (Oniga 2014:287-288). This phenomenon is prevalent in Latin, where seemingly any non-selected clause, finite or non-finite, may in principle be ‘relativized’. Needless to say, since I am following the punctuation of the editors, and since this punctuation does not derive from the source text, there is a chance that some ‘true relatives’ have made their way into the data considered here. It is unclear to what extent true relative clauses and pseudo relative clauses behave differently in syntax. Although the issue merits closer scrutiny, the impression given by the corpus is that there is indeed a difference between the two with respect to word order in that relativised main clauses behave rather like normal, un-introduced main clauses rather than like adjunct relative clause attached to a NP.

²¹The count of verb-final strings only includes strings which are $V \geq 2$, in other words it excludes cases where the clause consists of only a single verb (although these have not been excluded *tout court*, but enter the data as V1 clauses (only 4 tokens).

Table 5.1: Linear order of the finite verb in main clauses in *Egeria*

	Transitive	Unaccusative	Copula	Functional	Total
V1	40 (15.44%)	111 (21.02%)	41 (17.67%)	4 (7.02%)	196 (18.22%)
V2	100 (38.61%)	174 (32.95%)	78 (33.62%)	20 (35.09%)	372 (34.57%)
V3	57 (22.01%)	125 (23.67%)	67 (28.88%)	10 (17.54%)	259 (24.07%)
V4	39 (15.06%)	76 (14.39%)	31 (13.36%)	11 (19.30%)	157 (14.59%)
V5	12 (4.63%)	31 (5.87%)	13 (5.60%)	11 (19.30%)	67 (6.23%)
V6	9 (3.47%)	9 (1.70%)	1 (0.43%)	– (0.00%)	19 (1.77%)
V7	– (0.00%)	1 (0.19%)	– (0.00%)	1 (1.75%)	2 (0.19%)
V8	2 (0.77%)	1 (0.19%)	1 (0.43%)	– (0.00%)	4 (0.37%)
Total	259 (100.00%)	528 (100.00%)	232 (100.00%)	57 (100.00%)	1076 (100.00%)

Average number of constituents $\approx 3,77$

Verb-final strings: $304/1076 = 28.25\%$

Null-subjects: $380/1076 = 35.32\%$

These figures reveal a considerable degree of variation in the surface word order patterns.²² This is already enough to conclude that there is no V2 constraint at play in this text as defined in chapter 2, since there is clearly neither any requirement for a constituent to precede the verb, as attested by the quite frequent verb-initial strings, nor any constraint

²²A note on the different predicate classes is in order. In the annotated data set, further distinctions were made. The predicate class *transitive* includes 4 different categories in the data set; first, predicates which were classed as *transitive* during the annotation, including not only canonical transitive and ditransitive verbs (verbs selecting an internal, accusative NP or clausal argument), but also all verbs that take an internal oblique argument (PP or non-accusative NP) and an external agentive subject, such as for instance *cum aliquo* (ABL) *loqui* – ‘to talk to someone’ or grammaticalised, complex predicates like *gratias agere alicui*(DAT) – ‘to give thanks to someone’; second, *unergative* verbs, which also have an agentive external (subject) argument, but lack internal arguments, like *natate* – ‘to swim’; third, *reflexive verbs* (only transitive ones) like *se lavare* – ‘to wash oneself’; and finally, *causatives* (which are almost completely absent from the corpus). As for the predicate class *unaccusatives*, it comprises canonical unaccusatives (which are assumed to assign a patient or theme theta-role to their internal argument), passives and impersonal constructions. The class *copula* includes all non-auxiliary uses of the verb *esse* – ‘to be’ – whether used to connect a subject to a predicative complement or in presentational constructions to introduce new discourse referents, plus the verb *feri* – ‘to happen’, which is also used presentationally. The class of *functional* predicates is certainly the most heterogeneous and also the theoretically most questionable class. It includes all verbs which select a non-ACI, non-NCI infinitival complement (in the X-bar sense of the word), and hence includes verbs like *possum* – ‘to be able to’ – *debeo* – ‘to must, to be under the obligation to’ – *volo* – ‘to want’ – *nolo* – ‘to not want to’ – *malo* – ‘to prefer/want more to’ – but also *incipio*, *coepi* – ‘to begin’ – *dignor* – ‘to deign’ (to do), as well as some other, less frequent ones like *festino* – ‘to hurry’ (to do) – *sufficio* – ‘be able’ (to do) – *soleo*, *consuo* – ‘to be wont’ (to do) – *desidero* (to desire), *conor* (try), *cesso* (cease), *audeo* (to dare/risk), as well as some complex predicates like *necesse habeo* (to need to do), *libenter habeo* (to like/take pleasure in doing). In other words, no attempt is made to make a distinction between mono-clausal raising predicates and bi-clausal control predicates. Needless to say, such a distinction is theoretically well-founded and should in principle be made, but since the matter is both complex and somewhat peripheral to our concerns here, I decided to lump them together, for the time being, in this single class. The important thing is that the infinitival constructions selected by these verbs vary between preverbal and postverbal position, a fact which might shed significant light on the headedness of the IP.

against having multiple constituents in the prefield. On our definition of V2, this conclusion was strictly speaking given in advance, since no known Latin text presents a V2 grammar in this strict sense. Even on a more permissive definition of verb second which allows for the existence of ‘relaxed V2’ languages, the output of linear V2 in table 5.1 is very low, well below the figures adduced for main clauses in all Old Romance languages in Wolfe (2015).²³

The fundamental question is to what extent this text features V-to-C movement. The data in table 5.1 does not allow us to say anything about this, containing as it does only quantitative, surface-oriented data. Furthermore, although the low figures for $V \geq 4$ in the table at first seem to indicate that the verb tends occupy a leftish position in the clause, even this conclusion might be premature, given the fact that the average main clause contains no more than 3.77 constituents, as indicated below the table. The low frequencies of verb-late strings might therefore to some extent be an artifact of clause length.

Still, it might be the case that there is V-to-C movement combined with a Romance style rich use of the left periphery for information structural purposes, as argued by Ledgeway (2017). Furthermore, there is a particular feature of Egeria’s language that has a considerable impact on the figures in table 5.1 and that might lead to a somewhat distorted first impression of the syntax of the text, namely her predilection for stacking heavy, clausal constituents at the beginning of the sentence. These heavy constituents include adverbial clauses of various kinds, conjunct participial clauses, and ablative absolute constructions. A natural interpretation would be to consider these as instances of initial scene-setters, possibly occupying a high left peripheral position. An example like (300), although the verb surfaces in linear fourth position, contains a transitive verb and inversion, making it a plausible candidate for V-to-C movement:

- (300) *[Lecto ergo ipso loco omni de libro Moysi*
 read-*PST-PTCP-ABL* thus same passage-*ABL* all from book-*ABL* Moses-*GEN*
et facta oblatione ordine suo], [hac sic
 ad made-*PST-PTCP-ABL* oblation-*ABL* order-*ABL* REFL there thus
communicantibus nobis], [iam ut exiremus de
 communicate-*PRS-PTCP-ABL* us-*ABL* now as go.out-*IPFV-SBJV-1PL* from
aeclesia], dederunt nobis presbyteri loci ipsius
 church-*ABL* give-*PRF-3PL* us-*DAT* priests-*NOM* place-*GEN* same
eulogias... (3.6)
eulogiae-ACC

‘Having read that entire passage from the book of Moses and made oblation as customary, then communicating there, just as were about to leave the church, the priests of the place gave us *eulogiae*...’

Given that we find such examples, the V-to-C hypothesis cannot be discarded on the basis of linear order alone and must be elucidated by considering more evidence.

²³It is also worth noting that there is a considerable discrepancy between some of these figures and those presented in Ledgeway (2017, p.169). The most noteworthy difference lies in the fact that Ledgeway reports 133 more main clauses than table 5.1. The lion’s share of this divergence stems from the fact that the expression *id est* – ‘that is’ – was not annotated in my corpus. This formulaic explicative is particularly cherished by Egeria, who resorts to it no less than 118 times. If included, the difference between Ledgeway’s count of main clauses and ours would have sunk to 15, a difference that should not be considered surprising given the length of the text, the fact that different editions were used, and the sometimes rather blurry line between parataxis and hypotaxis in the text. This might also to a large extent account for the second important discrepancy between table 5.1 and Ledgeway’s data, namely the relatively higher amount of linear V2 (ca. 5.5% difference) in the latter.

5.3.1.1 The prefield in linear V2 strings

One possible approach is to consider the V2 strings more in detail, to see what kind of elements appear in the prefield. As was illustrated in chapter 2, this is one crucial domain where the syntax of V2 languages, as exemplified by the modern Germanic languages, differs substantially from non-V2 languages like English or the modern Romance languages. In the former, the prefield is an A-bar position able to host a great variety of different constituents, whereas in the latter, the prefield is generally reserved for the subject.

It is clear that, descriptively speaking, the prefield of V2 strings in *Egeria* behaves like V2 languages in this respect, since it does not only host the nominal or pronominal subject (301), but also adverbs and adverbial expression of various kinds (302), direct objects (303), predicative complements (304), oblique arguments (305), infinitives (306), sentential negation (307), as well as clausal constituents like adverbial clauses (308), conjunct participle clauses (309) and absolute clauses (310).²⁴

- (301) a. *[Monachi autem plurimi] commanent ibi uere sancti...* (10.9)
 Monks-NOM several stay-3PL here truly holy-NOM...
 ‘Several monks live here, truly holy men...’
- b. *Domine Iesu, [tu] promiseras nobis, ne aliquis hostium ingrederetur ciuitatem istam...* (19.9)
 Lord-VOC Jesus you-NOM promise-PLPRF-2SG us-DAT that-not
 anyone-NOM enemies-GEN enter-IPFV-SUBJ-3SG city-ACC this...
 ‘Lord Jesus, you promised us that no enemy would enter this city...’
- (302) a. *...nam et aeclesia ibi est cum presbytero. [Ibi] ergo mansimus in ea nocte...* (3.1)
 ...for also church-NOM there is with priest-ABL there
 remain-PRF-1PL in that night-ABL ...
 ‘...for there is a church there with a priest. We took up lodgings there for the night...’
- b. *...[hora ergo quarta] peruenimus in summitatem illam montis Dei sancti Sina...* (3.2)
 hour-ABL fourth arrive-PRF-1PL in summit-ACC that
 mountain-GEN God-GEN holy-GEN Sinai...
 ‘...in the fourth hour we thus arrived at the summit of the holy mountain of God, the Sinai...’
- c. *[In eo ergo loco] est nunc ecclesia non grandis...* (3.3)
 In that place-ABL is now church-NOM not big ...
 ‘In that place there is now a small church...’

²⁴There are also numerous cases where a past participle appears in first position directly followed by the auxiliary. However, it is clear that participle and auxiliary are virtually inseparable in the grammar of *Egeria*, as they always appear as a unit, no matter where in the clause the verb turns up. This must be taken as strong evidence that participle and auxiliary somehow form a complex verbal projection, and they are there annotated together as the finite verb. Clauses opening with the participle and the verb therefore enter the statistics as V1 clauses. Note that an unfortunate consequence of this state of affairs is that the distinction between G-inversion and R-inversion disappears, and with it a potentially important surface diagnostic for V-to-C movement. The distinction between G-inversion and R-inversion can in principle still be maintained in the case of modals and infinitives, although the text does not provide many interesting cases.

- (303) [*Multos enim sanctos monachos*] **uidebam** *inde*
 many holy monks-ACC see-IPFV-1SG therefrom
uenientes in Ierusalimam... (13.1)
 come-PRS-PTCP-ACC in Jerusalem-ACC
 ‘For I saw many holy monks coming from there to Jerusalem...’
- (304) [*Carneas*] *autem* **dicitur** *nunc ciuitas Iob* (13.1)
 Carneas call-PASS-3SG now city-NOM Job
 ‘The city of Job is now called Carneas’
- (305) [*Retro in absida post altarium*] **ponitur** *cathedra*
 Behind in apse-ABL past altar-ACC place-PASS-3SG chair-NOM
episcopo... (46.5)
 bishop-DAT...
 ‘In the apse behind the altar the chair is placed for the bishop...’
- (306) *et ecce* [*occurrere*] **dignatus est** *sanctus presbyter ipsius*
 and behold approach-INF deigned is holy priest-NOM same-GEN
loci et clerici... (14.1)
 place-GEN and clerics-NOM
 ‘... and behold! The holy priest and clergy of the place deigned to meet us.’
- (307) [*non*] *enim* **putabam** *hoc sine causa esse.* (16.3)
 not think-IPFV-1SG this-ACC without reason-ABL be-INF
 ‘For I did not think this could be without some reason.’
- (308) [*Statim ergo ut haec audiui*], **descendimus** *de*
 Immediately as these-ACC hear-PRF-1SG descend-PRF-1PL from
animalibus... (14.1)
 animals-ABL...
 ‘As soon as I heard these words, we got down from the animals...’
- (309) [*Transeuntes ergo fluuium*] **peruenimus** *ad ciuitatem,*
 Cross-PRS-PTCP-NOM river-ACC arrive-PRF-1PL to city-ACC
qui appellatur Libiada... (10.4)
 which-NOM call-PASS-3SG Libidia
 ‘Having cross the river we arrived at a city which is called Livias...’
- (310) [*Facta ergo et ibi oblatione*] **accessimus** *denuo ad*
 Make-PST-PTCP-ABL also there oblation-ABL procede-PRF-1PL again to
alium locum non longe inde... (4.4)
 other place-ACC not far therefrom...
 ‘After having made oblation there as well, we set forth again to another place not far from there...’

The prefield is therefore not specialised for hosting subjects, a fact which is also pointed out by Ledgeway and recruited as part of the evidence for V-to-C movement (Ledgeway 2017:172-175). An examination of the actual quantitative distribution of preverbal elements in linear V2 strings (see table 5.2 below) further underscores this, since subjects constitute

less than 20% of the overall amount of initial constituents.²⁵ This is indeed a remarkably low percentage, much less than that what seems natural to expect from a basic SVO or SOV language. In comparison, the corresponding figures adduced for six different Old Romance varieties in Wolfe (2015) range from slightly above 35% for Old Spanish to over 69% in Old Venetian, with the notable exception of Old Occitan, which featured no more than 23.78% subject-initial V2 strings. This leads us to raise the question if there is a (non-peripheral) subject position in front of the verb at all in main clauses.

Table 5.2: Preverbal constituent in V2 strings in main clauses in *Egeria*

Constituent	Tokens
Adverbial	187 (50.27%)
Subject	71 (19.09%)
Adverbial clause	53 (14.25%)
Participial clause	18 (4.84%)
Absolute clause	15 (4.03%)
Direct object	11 (2.96%)
Predicative complement	5 (1.34%)
Free relative clause	2 (0.54%)
Oblique object	2 (0.54%)
Negation	2 (0.54%)
Complement clause	2 (0.54%)
Infinitive	1 (0.27%)
Total	372 (100.00%)

While the evidence clearly shows that the prefield is not a subject position, this does not say anything about the structural position of the verb, and as a consequence, what structural domain of the clause the surface term ‘prefield’ really covers. The evidence considered so far is compatible with a V-to-C grammar, but it is also perfectly compatible with various grammars with verb raising only as far as I^0 . For instance, we cannot say in descriptive terms that the prefield is specialised for the subject in the modern Romance languages either, although this is at least generally the case in French (disregarding left or right peripheral arguments, see Harris 2000a:235-236), since most of these languages quite liberally accept V2 strings with initial non-subject constituents, both arguments of the verb and adjuncts, in other words strings which are altogether parallel to the examples in (3.7.3)–(310). Furthermore, any verb-initial grammar should also be able to produce the strings in (3.7.3)–(310); while it is not unlikely that some of the initial XPs occupy a position in the left periphery, that does of course not entail that the verb itself moves to the left periphery.

²⁵Ledgeway’s corresponding number is twice as high, 38.8% (Ledgeway 2017:172). This difference is once again due to the fact that the expression *id est* has been left out of my corpus. If included, the share of subject-initial V2 strings would rise to 41.04%. In this case, the difference is in fact quite relevant, since a percentage of less than 20% subjects reveals a pattern which in fact differs quite markedly from similar figure adduced for all Old Romance languages in Wolfe (2015).

5.4 Inversion

In order to make headway with these questions, we must consider inversion strings. The first thing to notice in this respect is that main clauses in *Egeria* feature a very considerable degree of inversion: $350/1076=32.53\%$. Given the fact that Latin is a Consistent Null-Subject language in the sense of Roberts and Holmberg (2010:5-13), always omitting referential subjects whenever their reference is retrievable from the context, this figure is important and must have provided a strong acquisitional cue. The question is how to interpret it.

The prerequisite for establishing a credible V-to-C hypothesis must be considered the existence of inversion strings featuring transitive predicates, since inversion with unaccusative predicates may just as well reflect a low position of the subject as a high position of the verb. To get a clearer picture of this, it is instructive to consider the distribution of the subject across different predicate classes. This information is provided in table 5.3.

Table 5.3: The position of the subject (S) distributed over different predicate classes in main clauses in *Egeria*

	Transitive	Unaccusative	Copula	Functional	Total
Preverbal S	73 (28.19%)	144 (27.27%)	112 (48.28%)	14 (24.56%)	343 (31.88%)
Postverbal S	66 (25.48%)	193 (36.55%)	88 (37.93%)	3 (5.26%)	350 (32.53%)
Null S	120 (46.33%)	191 (36.17%)	32 (13.79%)	40 (70.18%)	383 (35.59%)
Total	259 (100.00%)	528 (100.00%)	232 (100.00%)	57 (100.00%)	1076 (100.00%)

The first thing to notice is that the three possible options *preverbal subject*, *postverbal subject* and *null subject* show a very balanced overall distribution in main clauses, as can be gleaned from the ‘Total’ column.²⁶ On the other hand, inversion strings/postverbal subjects display a somewhat skewed distribution, as they seem to interact to some extent with the predicate class variable. If we leave aside the functional class, which for some reason shuns postverbal subjects, as well as the copula, confining our attention to transitive and unaccusative predicates, there is a preference for inversion with unaccusatives (36.81% vs. 25.86%), in line with previous observations on word order in the text (Väänänen 1987). The difference is significant (p-value 0.0238, d.f. 1, Chi-square 5.112).

On closer inspection, however, this difference is not primarily set off by a preferentially preverbal position of the subject of transitives compared to the subject of unaccusatives, as this divergence is very slight (28.19% vs. 27.27% respectively) and not statistically significant. Rather, transitive predicates tend to feature null-subjects more often than unaccusatives (46.33% vs. 36.17%). This difference is significant (p-value 0.0062, d.f. 1, Chi-square 7.501). This is important from an acquisitional perspective, since the children will have to deduce information about the relative position of the verb and the subject based

²⁶Spevak’s claim that the normal position for the subject (when expressed) is preverbal (Spevak 2005:1) must therefore be somewhat qualified; when all possible postverbal positions, including string final ones, are included, the subject is equally frequent in postverbal position.

on strings with overt subjects. Looking at the behaviour of the transitive predicates alone, the data in table 3.3 indicate that inversion is almost as frequent as non-inversion (25.48% vs. 28.19%).

We need an explanation for the somewhat stronger tendency for inversion under unaccusative verbs. One might hypothesize that unaccusative verbs in main clauses often perform the function of introducing discourse-new subjects, and that focal information is preferentially realized in postverbal position in Latin like in many languages (cf. the *given-new contract* of Clark and Haviland 1977). These assumptions find some support in the quantitative data. Starting with the assumption that focal information tends to be realized in postverbal position, we observe that from a total of 1076 main clauses, 80 (7.43%) introduce new discourse referents by way of the subject position. 57.50% of these are realized in postverbal position vs. 42.50% in preverbal position. Table 5.3 has shown that the global distribution of overt subjects relative to the verb is almost identical (31.88% preverbal vs. 32.53% postverbal subjects), so the focal status of the subject does seem to be an extra effect in inversion. The difference is not statistically significant, however (p-value 0.2359).²⁷

The second assumption, that unaccusatives introduce new discourse-referents more often than transitives, is much easier to confirm. The data in table 5.4 reveal that unaccusatives along with the copula are the only predicates that serve the function of introducing new discourse referents into the text through the subject position, as only two tokens are found of discourse-new subjects with transitive verbs.²⁸

²⁷In fact, new discourse referents are not infrequently introduced directly into the prefield (ia), and moreover, there are many near-minimal pairs, cf. (ia)–(ib), where the subject is underlined:

- (i) a. *Haec est autem vallis, in qua factus est vitulus, qui usque in*
 This-NOM is valley-NOM in which-ABL made is calf-NOM which-NOM until in
hodie ostenditur: nam lapis grandis [ibi] fixus stat in ipso loco. (2.2)
 today show-PASS-3SG for stone-NOM big there fixed stand-3SG in same place-ABL
 ‘For this is the valley where the (golden) calf was made, which can be seen even today: for a big rock stands firmly there on that very spot.’
- b. *Mostrauerunt etiam locum, ubi factus est vitulus ille; nam [in eo loco]*
 Show-PRF-3PL also place-ACC where made is calf-NOM that for in that place-ABL
fixus est usque in hodie lapis grandis. (5.2)
 fixed is until in today stone-NOM big
 ‘They also showed us the place where the calf was made; for in that place a big rock stands until this day.’

²⁸However, it is worth mentioning that the annotation contained a discourse category which was called ‘new, anchored’, and which is not included in table 3.4. This category was used for discourse-new subjects which are ‘anchored’ by some other element that provides more information about that subject; in most cases, this is a relative clause. With such subjects, inversion is equally frequent with transitive verbs and unaccusative verbs. Note however that the postverbal subject in these cases is very often string-final, as in (i):

- (i) *... et [peruenientes ad monasteria quaedam] susceperunt nos [ibi]*
 ... and come-PTCP.PRS-NOM to monasteries-ACC some receive-PRF-3PL us-ACC there
[satis humane] monachi, qui ibi commorabantur (3.1)
 very humanely monks-NOM who-NOM there stay-IPFV-3PL
 “and arriving at some hermitages, the monks who resided there received us very warmly.”

It is tempting to see the position of the subject here as dictated by the need to provide adjacency between the antecedent and the relative pronoun. This is presumably a syntactic principle, but at the same time and on a functional level, this anchorage through a relative clause seems to facilitate the introduction of new-discourse subjects with transitive verbs.

Table 5.4: Discourse-new subjects distributed across predicate classes in main clauses in *Egeria*

Transitive	Unaccusative	Copula	Total
2 (2.50%)	48 (60.00%)	30 (37.50%)	80 (100.00%)

The conclusion must clearly be that new subjects are normally first introduced by means of the copula or an unaccusative verb, and only in the next discourse move are these subjects combined with transitive verbs, a fact which also goes some way towards accounting for why null-subjects are appreciably more common with transitive verbs (see table 5.3). I suggest that these natural information-structural principles go a long way towards accounting for the relatively stronger tendency for inversion with unaccusative predicates, without any need to invoke any extra syntactic difference. This being said, it is of course also possible to imagine that a lower first-merge position of the arguments of unaccusative verbs may play a role. I will not pursue this further here.

Let us return to transitive predicates and considering some examples of inversion (311–(317). Notice the strong tendency for oblique pronominal arguments to intervene between the verb and the postverbal subject. This might suggest that the pronouns cliticize to the verb, but they patterns revealed here are not in adherence with the Tobler-Mussafia Law (Mussafia 1898) and suggest proclisis rather than the enclitic position in non-verb-initial clauses in Old Romance:

(311) *Ac [sic] ergo [aliquo biduo] [ibi] tenuit nos sanctus*
 and thus some two-day-period-ABL there keep-PRF-3SG us-ACC holy
episcopus... (9.1)
 bishop-NOM...

‘And thus the holy bishop lodged us there for a couple of days...’

(312) *[Euntibus nobis] commonuit presbyter loci*
 Walk-PTCP.PRS-ABL us-ABL advise-PRF-3SG priest-NOM place-GEN
ipsius... (10.8)
 same

‘While we were walking, the priest of the place gave us advice...’

(313) *[Tunc] dixerunt nobis sancti, qui nobiscum iter*
 than say-PRF-3PL us-DAT holy-NOM who-NOM us-with road-ACC
faciebant... (16.3)
 make-IPFV-3PL

‘Then the holy men who where travelling with us told us...’

(314) *Ostendit etiam nobis sanctus episcopus memoriam Aggari uel*
 Show-PRF-3SG also us-DAT holy bishop-NOM tomb-ACC Abgar-GEN or
totius familiae ipsius... (19.18)
 whole family-GEN same

‘The holy bishop also showed us the tomb of Abgar and of his whole family...’

- (315) [*Nachor autem cum suis uel Bathuhelem*] [*non*] *dicit*
 Nachor-ACC with REFL.ADJ.ABL or Bathuhelem-ACC NEG say-3SG
scriptura canonis, quo tempore transierint. (20.10)
 scripture-NOM canonical what time-ABL cross-PRF-SBJV-3PL
 29

‘Nachor with his people, or Bathuhelem, the canonical scriptures do not mention at what time they passed this way.’

- (316) [*Post biduo autem quam ibi feceram,*] *duxit*
 After two.day.period-ABL that here make-PLPRF-1SG lead-PRF-3SG
nos episcopus ad puteum illum, ubi... (21.1)
 us-ACC bishop-NOM to well-ACC that where ...

‘After the two days that I stayed there, the bishop took us to that well where...’

- (317) *atque iterata oratione] benedixit nos episcopus.*
 and repeat-PST.PTCP-ABL prayer-ABL bless-PRF-3SG us-ACC bishop-NOM
 (21.1)

‘And after another prayer, the bishop blessed us.’

If we trust the testimony of the text with respect to the word order of spoken language at the time, we may conclude, on the strength of the quantitative evidence in table 5.3 and the qualitative evidence in (311–317), that the weakest of our initial hypotheses has been confirmed; namely that the Late Latin grammar reflected in *Egeria* supports inversion under any kind of predicate. In this respect, this grammar differs both from Classical Latin on the hand and from the modern Romance languages on the other,³⁰ a very promising finding for our hypothesis that Old Romance inversion stems diachronically from an internal Latin-Romance development. We seem to be dealing with a transitional phase here, a potential link between a (perhaps predominantly SOV) language with great word order freedom and a more configurational system with widespread inversion.

Having rejected the strongest hypothesis (the grammar of *Egeria* was not that of a V2 language) and confirmed the weakest (the grammar of *Egeria* featured widespread inversion), the rest of this chapter will focus on the middle hypothesis, namely whether Late Latin as exemplified by this text had developed a syntax with V-to-C movement in declaratives.

5.5 The position of the subject

We cannot jump to the conclusion that the inversion structures just demonstrated necessarily feature V-to-C movement. The reason why this conclusion would be premature is that we simply do not know what the basic position of the subject is, and in consequence, we do not

²⁹This is one of several cases of *proleptic accusative* constructions in *Egeria* and in the corpus in general. According to the view of many traditional philologists, the proleptic accusative is a feature of spoken language (Löfstedt 1936 [1911], Hofmann and Szantyr 1965:471-472). The same conclusion is also reached by Serbat in more recent work (Serbat 1996:181).

³⁰Of course, it is not entirely correct to say that Classical Latin does not permit inversion with transitive verbs. This word order pattern, along with any other possible permutation of the verb and its arguments, is attested in Classical Latin as well (Ledgeway 2012:61-62, Haug 2017, Danckaert 2017b:4-7); the point here is that there seems to be a systematic and relevant difference here between the latter and the grammar of *Egeria*.

know what the position of the verb is either. Inversion can only be considered strong evidence for V-to-C movement given certain other conditions, namely when the subject occupies a high position in the clause. It is with all likelihood the abundant and unambiguous evidence for such a high subject position in modern Germanic and Old French that creates a strong pressure on children to analyse inversion strings as featuring V-to-C movement.³¹ Recall from chapter 3 that we considered the potential of both a V-to-I and a V-to-C parse to account for the inversion strings in Old French, and that it was not until the unmarked subject position in Spec-IP was established beyond doubt in chapter 4 that the V-to-C parse clearly outperformed the V-to-I parse. We must therefore establish the position(s) of the subject in the grammar of *Egeria*.

Secondly, even if we were to find evidence for a high subject position, we would still have to make sure that the inversion structures actually arise by movement of the finite verb across this high subject position, rather than for instance by the subject targeting some low position such as the topic or focus positions of a lower left periphery argued for in Italian by Belletti (2004), or the rightward position targeted in narrow focus reading on the subject, whatever the proper analysis of such strings, or finally a high, right-peripheral position, which might seem plausible in cases where the subject is modified by a relative clause or otherwise structurally or prosodically complex (cf. the issue of string-final Heavy Inversion in Old French (Vance 1997)). It is not clear that the evidence considered so far has been conclusive in this regard; note for instance that many of the examples in (311)–(317) are in fact string-final.

In other words, alongside the question of the subject position, there is the equally crucial question how similar the inversion structures in *Egeria* are to those found in the Old Romance languages. One of the hallmarks of the Old Romance inversion structures is that they are cases of *G-inversion*, meaning that the postverbal subject intervenes between the finite auxiliary and non-finite main verbs. This is important because it shows that it is indeed the verb that occupies a high structural position above the subject, rather than the latter occupying a low position in the clause.

What this means is that uncontroversial evidence for V-to-C movement is presumably dependent on at least three factors: inversion strings which are not restricted to a particular type of predicate (which we have already seen to be the case in *Egeria*), an unmarked subject position in a high position, preferably the highest non-peripheral position for which there is evidence (which we have not established), and G-inversion or any other equivalent type of evidence demonstrating that it is in fact the verb that moves above the subject (which we have not established.)

We are therefore led to ask if there is any evidence of this kind in *Egeria*, starting with G-inversion.

5.5.0.1 G-inversion and functional predicates

G-inversion arises in two different guises in Old Romance and Modern Germanic, namely with temporal and modal auxiliaries. Since Late Latin still has not developed the active periphrastic perfect and pluperfect of the Romance languages,³² we are therefore left with

³¹That is not to say that there is only a high subject position in modern Germanic; while there is some variation across the family in this respect, at least the Scandinavian languages also allow for (presumably several) lower subject positions than Spec-IP, cf. example (15 in section 2.2.2. The same applies to Old French, as we have seen in chapter 3.

³²In *Egeria*, there is a handful of constructions which resemble such incipient periphrastic constructions:

constructions featuring the past participle and a form of the copula for the first system. These only turn up in passives and the perfects of deponent verbs. However, we run into a problem here, since the auxiliary and the participle in our corpus are virtually inseparable, as also observed by Ledgeway (2017, p.173 fn.) and Danckaert (Danckaert 2017a:147-149).³³I interpret this as evidence that the participle has combined with the auxiliary, presumably by left-adjunction, since the order *participle-esse* is almost exceptionless.

- (318) *Lectus est ergo et ibi ipse locus de libro Moysi...* (4.4)
 Read is thus also there the passage-NOM from book-ABL Moses-GEN
 ‘There also the passage from the book of Moses was read.’

The corollary of this state of affairs is that in main clauses, all cases of inversion under complex predicates feature unaccusative verbs and R-inversion, structures which are completely licit in modern Romance as well. As for transitive verbs, there are no cases of complex inversion at all. The structural information of the clause is therefore somewhat impoverished, since only a single head position is lexicalised at a time. The only cases where we do in fact have multiple heads overtly spelled out are provided by the group of functional predicates, but as table 5.3 above has illustrated, these are for some reason highly prone to appear in clauses that lack overt subjects. Furthermore, when they do in fact turn up with overt subjects, the latter are consistently non-inverted. Only a single potential case of complex inversion is attested (319) and this example is not even clear at all, since the predicate *incipere* – ‘to begin’ might just as well be a control verb that selects a non-finite clausal complement, in which case we are not dealing with a restructured or complex predicate at all. Besides, the subject (if it is indeed the subject and not just a free predicative) is string final.

- (319) *Et incipient episcopo ad manum accedere singuli.* (24.6)
 and begin-3PL bishop-DAT to hand-ACC proceed-INF each-NOM.PL

-
- (i) a. *Tunc videntes hoc Persae auerterunt ipsam*
 Then see-PTCP.PRS-NOM this-ACC Persians-textitNOM divert-PRF-3PL the
aquam a ciuitate et fecerunt ei decursum contra ipso loco,
 water-ACC from city-ABL and make-PRF-3PL it-DAT detour-ACC towards same place-ABL
ubi ipsi castra posita habebant. (19.11)
 where they-NOM camp-ACC placed have-IPFV-3PL.
 ‘Seeing this, the Persians diverted the water from the city and made it flow towards the place they themselves had put up camp.’
- b. *Ipsam ergam uallem nos trauersare habebamus...*
 Same valley-ACC we-NOM cross-INF have-IPFV-1PL
 ‘For we had to cross that (same) valley...’

As for the first example, it is hard to make any strong case for an incipient Romance structure, since periphrastic constructions of this kind is attested since early Latinity with telic predicates that give rise to resultative readings of the kind in (ia). In other words, the participle is arguably not (fully) verbal here, but rather a nominalised adjective functioning as a predicative complement of the object. In *Egeria*, the construction still shows no signs of expanding beyond the Classical pattern. (ib) is a periphrastic construction featuring an infinitive and an inflected form of the verb *habere* – ‘to have’, the structure that would ultimately become the new Romance future and conditionals tenses. In *Egeria*, it is used in total 3 times, the example in (ib) being the closest to a temporal reading. None of these cases involve inversion of any kind, but rather a head-final auxiliary.

³³As is well documented by Danckaert, the participle and the auxiliary tend not to split in ‘early’ Latin as well (200 BC - 200 AD), but this tendency becomes vastly stronger and approaches a rule in Late Latin (200 AD - 600 AD). Note however that the data provided by Danckaert concern participles and auxiliaries in constructions featuring transitive deponent verbs, not passives (Danckaert 2017b:137-138).

‘And they start approaching the bishop in turn to kiss his hand.’

Laying aside the issue of inversion for a minute, functional verbs are still very useful in determining the structural organization of the clause, both for the child and the linguist, as these verbs lexicalise more head positions simultaneously. This point is emphasized most thoroughly by Danckaert (2017), who uses sequences of (*possum* and *debeo*) auxiliaries and VP as the primary diagnostic for the evolution of clausal structure. In our text, it is clear that the story told by these richer structures is hardly one of a generalized, across-the-board V-to-C movement:

- (320) [*Itaque*] *ergo* [*Deus noster Iesus*], (*qui sperantes in*
Therefore God-NOM our Jesus who believe-PTCP.PRS-ACC in
se non deseret,) [*etiam et in hoc*] [*uoluntati meae*]
REFL-ACC NEG desert-3SG, also and in this-ABL will-DAT my
[*effectum*] [*praestare*] ***dignatus est.*** (10.2)
effect-ACC lend-INF digned is.
‘Thus our Lord Jesus, who does not abandon those who believe in him, digned to give effect to my will in this matter as well.’

As always, the bracketed notation indicates the maximum number of constituents. It is of course possible to suggest that *itaque* is somehow an external connective, or to argue for more complex constituents, perhaps by suggesting that *effectum praestare* – to give will (to) – is somehow a complex predicate, or even that *uoluntati meae effectum praestare* is the entire VP that is fronted.³⁴ But even with all of these assumptions, it is not clear how this sequence would map on to the left peripheral roadmaps that have been proposed in the literature. The only solution that comes to mind would be to suggest that *Deus noster Iesus* is a topic, followed by another topic *etiam et in hoc*, finally followed by the VP fronted to the lowest left peripheral focus position. If this were an isolated example in a text that otherwise showed robust evidence for V-to-C movement in main clauses, perhaps such a solution could be warranted. But this is not the case, as the examples (321)–(326) serve to demonstrate (parenthetical clauses are enclosed in parentheses):

- (321) [*Cum autem ingressi fuissetus ad eos*],
When entered be-PLPRF-SBJV-1PL to them-ACC
[*facta oratione cum ipsis*] [*eulogias*] *nobis*
made-PTCP.PST-ABL prayer-ABL with them-ABL eulogiae-ACC us-DAT
[*dare*] ***dignati sunt.*** (11.1)
give-INF digned are

³⁴Notice also that the VP is head-final, and the same applies to (321), (324) and (325) as well. The persistence of head-final verbal projections should not come as a great surprise, and it has indeed been noted before in the literature that the new Romance future and conditional paradigms are the grammaticalisation of a verb-final projection. The same applies to some complex predicates such as *certify*:

- (i) [*certas uos facere*] ***debi*** (24.1)
sure.ACC.FEM you.ACC.CL make.INF I-should
‘I should inform you. . .’

This provides more evidence that the stable head-initial patterns of Romance are not fully in place yet. On the other hand, notice that the VP in (326) is seemingly head-initial, in violation of the Final-over-Final Constraint (Biberauer et al. 2014), which (somewhat simplified) states that head-final projections can only dominate other head-final projections. For a discussion of violations of the FOFC in the history of Latin and a proposal for an analysis, see Danckaert (2017).

‘When we had entered and prayer had been made together with them, they deigned to give us *eulogiae*.’

- (322) [*Tunc*] [*ego*], (*ut sum satis curiosa*), [*requirere*] **coepi** (16.3)

Then I, as am very curious, ask-*INF* begin-*PRF-1SG*

‘Then, since I am very curious of nature, I began to ask. . .’

- (323) [*Quae me cum uidisset*], [*quod gaudium illius*
Who-*NOM* me-*ACC* when see-*PLPRF-SBJV-3SG* what pleasure-*NOM* her-*GEN*
uel meum esse potuerit], [*nunquid*] *uel* [*scribere*]
or mine-*ACC* be-*INF* be.able-*PRF-SBJV-3SG*, surely-not write-*INF*
possum? (23.3)

be.able-*1SG*

‘And when she had seen me, how could I possibly write her joy or my own?’

- (324) [*quia adhuc catechumini estis*], [*misteria Dei*
because until-now catechumens-*NOM* BE-*2PL* mysteries.-*ACC* god.-*GEN*
secretiora] [*dici*] *uobis* [*non*] **possunt.** (46.6)

more-secret-*ACC* tell-*INF-PASS* you-*DAT* NEG be.able-*3PL*

‘And since you are still catechumens, the most secret of God’s mysteries cannot be told to you.’

- (325) *Et* [*illud*] [*etiam*] [*scribere*] **debui** (45.1)

and this-*ACC* also write-*INF* should-*PRF-1SG*

‘And I should write this as well’

- (326) *et* [*ideo*] [*fallere*] *uos* [*super hanc rem*] [*non*] **possum.**
and therefore fool-*INF* you-*ACC.PL* over this thing-*ACC* NEG be.able-*1SG*
(12.7)

‘And therefore I cannot lie to you on this matter. . .’

In order to maintain a V-to-C analysis of such cases, one would have to assume a quite frequent VP-fronting to topic and focus positions, often in combination with other constituent fronting operations. In (325), a VP-fronting analysis might not seem entirely implausible, but in (326), for instance, one would have to assume VP-fronting of *fallere vos* coupled with other XP fronting operations on either side; notice also that the candidate VP *fallere vos* is clearly focal information and must accordingly be expected to target Spec-FocP on such an analysis, and yet the following PP *super hanc rem* is clearly topical information. The sequence therefore does not match well with the view that all topics precede the left peripheral focus projection (Benincà and Poletto 2004). In a more general vein, although VP-fronting operations clearly exist, they are at least not very frequent in either the modern Germanic or Romance languages, and one is justified in asking why Latin should behave so differently.

To emphasize this point further, consider the examples (327)–(330). These only feature a simplex verb and are therefore structurally somewhat less informative, but on the other hand, the lack of a non-finite verb makes it much harder, if not impossible, to argue for the fronting of any higher order constituent. The sheer amount of constituents in front of the verb in some of these examples would seem to strain even the most elaborately structured left periphery:

- (327) *[seculares] autem (tam uiri quam feminae) [fidei]*
 Lay.people-NOM both men-NOM and women-NOM faithful-ABL
animo] [propter diem sanctum] [similiter] se [de omnibus]
 mind-ABL because-of day-ACC holy likewise REFL from all
prouinciis] [isdem diebus] [Ierusalima] colligunt. (49.1)
 provinces-ABL these days-ABL Jerusalem-ABL gather-3PL
 ‘Lay-people, both men and women, likewise in these days gather faithfully in Jerusalem from all provinces because of the holy day.’
- (328) *Sed [statim] [Aggarus] [epistolam Domini ferens] ad*
 But immediately Abgar-NOM letter-ACC lord-GEN carry-PTCP.PRS-NOM to
portam] [cum omni exercitu suo] [publice] orauit. (19.9)
 gate-ACC with all army-ABL REFL.ADJ-ABL publicly pray-PRF-3SG

‘But immediately Agbar, carrying the letter to the gate, held a public prayer together with all of his army.’

- (329) *[Sane] [dominica die per pascha] [post missa lucernarii], (id est*
 Truly sunday day-ABL at Easter after mass-ACC vespers-GEN] that is
de Anastase), [omnis populus] [episcopum] [cum ymnis] [in Syon]
 from Anastasis all people-NOM bishop-ACC with hymns-ABL in Syon
ducet. (39.4)
 lead-3SG
 ‘For on Easter sunday, after the mass of vesper – that is at the Anastasis – all of the people conduct the bishop to Syon with hymns.’
- (330) *Et [iam] [inde] [cum ymnis] [usque ad minimus infans] [in*
 and now therefrom with hymns-ABL until the smallest child-NOM in
Gessamani] [pedibus] [cum episcopo] descendunt (36.2)
 Gethsemani feet-ABL with bishop-ABL descend-3PL
 ‘And then everyone, even down to the smallest child, goes down on foot with the bishop to Gethsemani with hymns.’

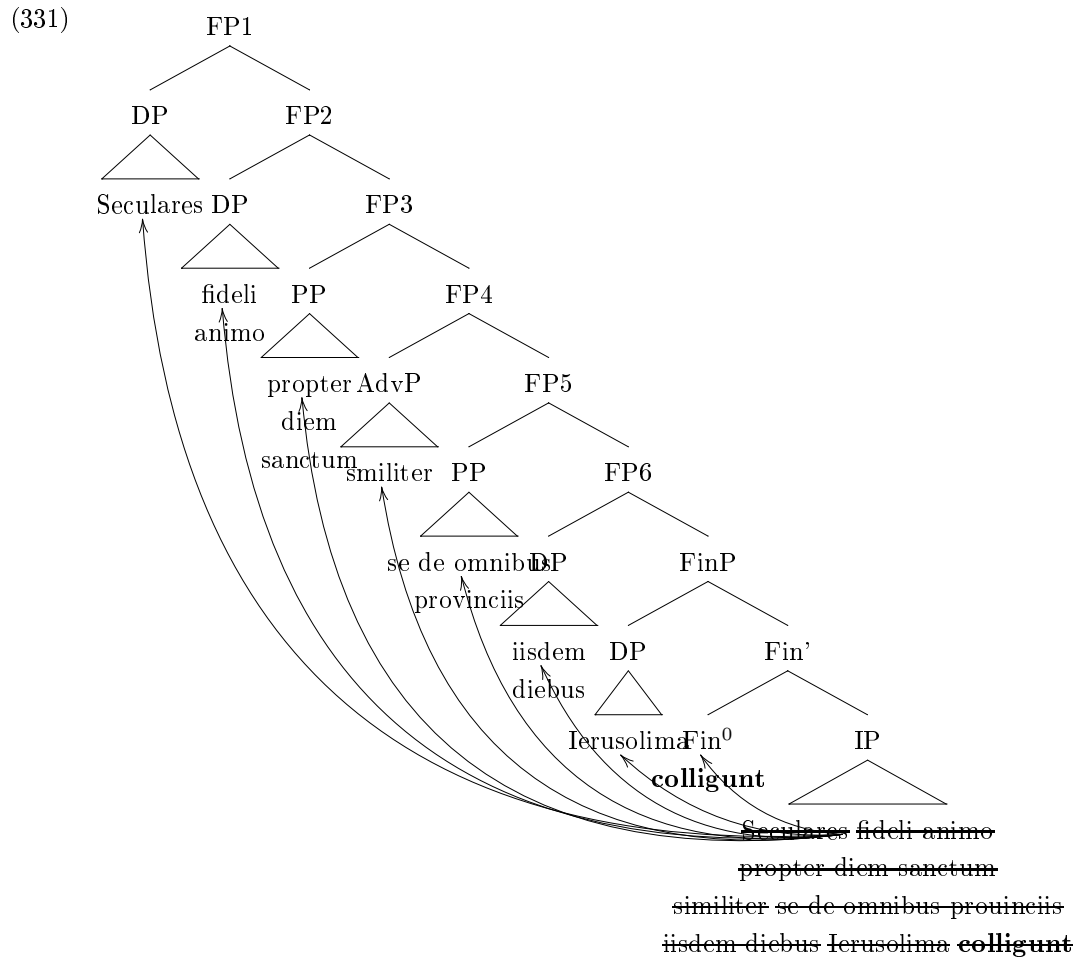
Once again, it is possible to argue for more complex constituents, but there are limits on how far this can be pushed, and even on a very inclusive bracketing the number of constituents in front of the string-final verb is very high.³⁵ It does not seem reasonable to

³⁵In a more general vein, we should ask how frequent such multiple fronting operations to the left-periphery really are. Presumably, the use of the left-periphery per se is very frequent, since scene-setters, linking devices and topicalisation are all very common strategies in discourse, much more frequent than all-focus sentences. But even in a language like modern Italian, which seems to make quite substantial use of the left periphery, cases like Rizzi’s by now famous example (i), in which the verb is pushed into a linear 4th position by multiple left-peripheral constituents, are presumably not *highly frequent*, at least not if compared to the figures for $V \geq 4$ in table 5.1:

- (i) *[A Gianni], [QUESTO], [domani], gli dovrete dire.*
 To Gianni, THIS, tomorrow, him.CL should-COND-2PL say-INF
 Literally: ‘To Gianni, this, tomorrow, you should tell him.’
 (From Rizzi 1997:291)

If anything, it does not seem unreasonable to suppose that the left periphery is exploited more in everyday

suppose that the entire clause for some reason has been ‘evacuated’ to the left periphery, but unless we want to assume that the verb occupies a head-final projection in the left periphery – for which there is no evidence – that is indeed what the hypothesis of generalized V-to-C forces us to do in examples like (327):



I can only think of one other parse of (327) that is consistent with V-to-C movement, and that would be to assume that there is Remnant IP-fronting after the verb has moved to Fin⁰. I cannot see any reason to consider such options, not only given the well-documented immobility of the IP (Abels 2003; Wurmbrand 2014; Bošković 2018), but also because it seems highly unlikely that children should prefer such an ingenious parse over the simple alternative of a normal, head-final IP.³⁶

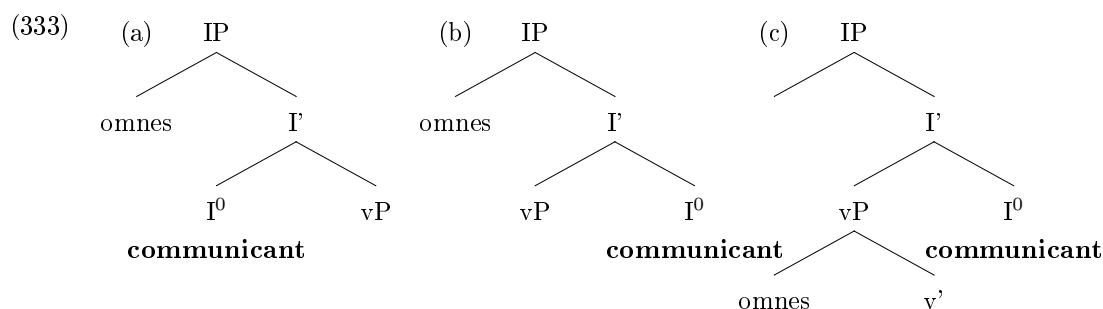
conversation that in a written genre like our itinerary, now matter how unpretentious and colloquial its language might be.

³⁶It also seems to me that, once such possibilities are admitted along other more conservative ones, the predictive power of the model is seriously diluted. To take just one example, nothing presumably prevents us from assuming that Classical Latin SOV orders were derived by obligatory V-to-C movement followed by remnant-IP movement.

The hypothesis of a generalized V-to-C movement therefore does not seem sustainable. Verb-final strings in *Egeria* are not just the accidental result of there being no phonologically overt material following the verb, but are rather the expression of a syntactic structure that puts the verb in string-final position. I will therefore from now on assume that the verb may reside in a head-final IP in some cases.³⁷ Furthermore, we are forced to recognize that the verb in I⁰ can still both precede and follow the VP, just like in the classical language. This is in accordance with a quite common view of Latin syntax, suggested among others by Bauer (1995) and more recently Danckaert, who states that ‘[w]e know independently that at all stages of the Latin language, the T-node can be either head-final or head-initial’ (Danckaert 2017a:126). It follows from this that shorter strings are indeed syntactically ambiguous, since the IP can be both head-initial and head-final.

Here, I will adopt a base-generation analysis of this variation. To illustrate very briefly, consider a very short sentence like (332), consisting only of the subject and the verb. Even if we disregard the possibility that there might be V-to-C movement or that the subject may have been topicalised to the left periphery, the clause is at least threefold ambiguous (333):

- (332) ... *sed [omnes] communicant*
 but all-NOM communicate-3PL
 ‘but all communicate’.



This means that we can assume that long clauses with string-final verbs like (331) involve a finite verb in a head-final IP. However, this does not allow us to draw up the exact structure of the clause with complete certainty. The fact that the subject is initial, preceding various adverbial expressions, shows that it does not occupy a low position, and a priori points towards a position in Spec-IP. However, it could of course also be that it has been topicalised to the left periphery. In other words, we still need to find out more about the position of

³⁷Harris suggests that verb-final patterns are ‘inflated by the attempts of the authoress in a number of ways to imitate classical usage.’ (Harris 1977:36). This argument does not seem convincing, for the simple reason that *Egeria* does not on the whole seem to imitate Classical patterns. If the author was aware that verb-final was a norm of proper Classical Latin and wanted to imitate it, why does she not do it more often? There is no a priori reason to assume that verb-final projections are not part of the grammar of *Egeria*, or indeed that verb-finality in general had vanished completely from spoken language in *Egeria*’s day. This is pointed out by Clackson and Horrocks, who also call attention to the relatively late grammaticalisation of the new Romance future and conditional tenses from an erstwhile *head-final* pattern: ‘It seems then that despite a trend towards head-first structures, different word order patterns are still possible in the middle of the first millennium’ (Clackson and Horrocks 2007:281). I concur with this assessment. On the other hand, Clackson and Horrocks also claim that the verb in *Egeria* ‘is never situated at the end of a main clause’ (Clackson and Horrocks 2007:291), a claim which is falsified by the data.

the subject in the grammar of Egeria. G-inversion did not provide any answer, so we are forced to look for other kinds of evidence.

5.5.1 Adverbs

A potentially useful probe into the structural position of the subject is provided by IP-cartography. We will follow the line of reasoning established by Cinque (1999), and assume that non-circumstantial adverbs do not move clause-internally, but rather reside in the left-leaning specifiers of a strictly ordered hierarchy of functional projections in the middle field of the clause. However, adverbs may clearly move to operator positions in the left-periphery, and furthermore, homophonous adverbs may sometimes represent the spell-out of different functional projections. We must therefore keep these caveats in mind when using adverbs as evidence for syntactic structure.

Unfortunately, much like G-inversion, the story told by sentential adverbs in main clauses is not very illuminating. There are not too many adverbs in general and the ones which appear are largely restricted to a handful of temporal adverbs like *denuo* ('again'), *iterato* ('again'), *primum* ('first'), *deinde* ('thereafter', 'next') *statim* ('immediately'), *iam* ('now'), the latter often assuming a discourse particle-like role (cf. *ja* in Old French, chapter 3). A few locative adverbs are also encountered, most notably *inde* ('from there', 'thither') and *ibi* ('there'), plus the manner adverbs *sic/ita* ('thus', 'in such a way'), *item* ('likewise') and *similiter* ('similarly', 'likewise'), as well as the causal adverb *ideo* ('therefore').

All of these adverbs tend more than anything to appear in very high positions, often as the first constituent of the clause or after an initial clausal constituent which could possibly be interpreted as an initial scene-setter. They therefore generally outscope both the verb and the subject, making them all but useless as diagnostics for the relative position of the latter. In many of these cases there is inversion below the adverb, and if this inversion is produced by V-to-C movement of the verb, we would have to conclude that these adverbs are in fact very often moved to operator positions in the left periphery, presumably to function as scene-setters or connectives. While this hypothesis is appealing in some contexts (334, 336), it is arguably less so in the not infrequent cases where these adverbs are stacked together in initial position (337)–(339). Not because stacking in itself is an uncommon phenomenon, but because a general prerequisite for stacking seems to be that the adverbial expressions are of the same kind (temporal, locative), so that they together form a kind of complex constituent. In (337) a locative adverb is apparently stacked together with a temporal adverb, and in (339), a manner adverb has joined the group:

- (334) *[Item] fit oratio et [denuo] mittet diaconus*
 Also happen-3SG prayer-NOM and again send-3SG deacon-NOM
uocem. . . (24.6)
 voice-ACC

'There is also a prayer and again the deacon lifts his voice. . .'

- (335) *[Item] hora sexta [denuo] descendit omnes similiter ad*
 Also hour-ABL sixth again descend-3PL all-NOM likewise to
Anastasim. . . (24.3)
 Anastasis-ACC

'Also at the sixth hour everybody again descends in similar fashion to the Anastasis. . .'

- (336) *Et [statim] leuat se episcopus et omnis populus...* (31.2)
 and at-once lift-3SG REFL bishop-NOM and all people-NOM
 ‘and at once the bishop and all of the people get up...’
- (337) *[Ibi] [denuo] legitur ille locus de euangelio ubi...* (36.4)
 there again read-PASS-3SG that passage-NOM from gospel-ABL where
 ‘And there again that passage is read from the gospel where...’
- (338) *et [ideo] [ante quartam horam aut forte quintam] [missa] [non] fit.* (25.1)
 and therefore before fourth hour-ACC or perhaps fifth mass-NOM NEG
 happen-3SG
 ‘and therefore mass does not take place until the fourth or perhaps the fifth hour.’
- (339) *Et [ibi] [denuo] [similiter] [lectiones et ymni et antiphonae aptae diei] dicuntur* (35.4)
 and there again likewise readings-NOM and hymns-NOM and antiphons-NOM suitable-NOM day-DAT say-PASS-3PL
 ‘And there again in similar fashion lessons, hymns and antiphons suitable for the day are recited...’

Still, it is possible to come across a few examples where some of these adverbs in fact intervene between the verb and the subject in inversion strings (340–(341), and there are also cases where phrasal adverbial expressions intervene (342). This shows that there is no adjacency between the verb and the subject in inversion structures. Such non-contiguous inversion strings, which we also witnessed in Old French, *a priori* point to a rather low position for the subject. Ledgeway, following Cinque (1999), argues that adverbs like *denuo* demarcate the edge of the vP (Ledgeway 2017:185). If this is correct, we can assume that the subjects of the following examples are vP-internal:

- (340) *et fit [denuo] oratio ad Crucem et dimittitur populus.* (31.4)
 and happen-3SG again prayer-NOM at cross-ACC and dismiss-PASS-3SG people-NOM
 ‘Again there is a prayer to the cross and the people is dismissed.’
- (341) *Cum ergo peruentum fuerit in Gethsemani, fit [primum] oratio apta...* (36.3)
 when arrived be-PRF-SBJV-3SG in Gethsemani, happen-3SG first prayer-NOM suitable
 ‘... when the crowd has arrived in Gethsemani, there is first a suitable prayer...’
- (342) *[Nam si dominica dies est], [primum] leget [de pullo primo] episcopus euangelium...* (44.2)
 for if sunday day-NOM is first read-3SG from cock-ABL first-ABL bishop-NOM gospel-ACC
 ‘for if it is Sunday, from the first crow of the cock the bishop first reads the gospel...’

Unfortunately, all cases involve subjects of unaccusative verbs. In principle, it could be that these are situated lower in the tree than agentive subjects of transitive verbs, for

instance in a first-merged complement position inside the VP.³⁸ The evidence does not allow us to decide; the cases where the subject outscopes sentential adverbs are invariably non-inverted clauses with the subject in a preverbal position, and in such cases it is impossible to tell if the subject is clause-internal or left-peripheral.

At this point, there is presumably not much more information to cull from main clauses. It is simply not possible to determine the position of the subject with certainty. The quantitative data show that subjects are equally frequent in preverbal and postverbal position, but qualitative analysis fails to reveal which of these positions is the unmarked and which is the derived, since neither complex inversion nor adverb positions provided a clear result in this respect. Still, the weight of the evidence tends toward a low position for the subject, as witnessed by certain instances of non-contiguous inversion. The same conclusion is reached by Ledgeway (2017:186).

5.5.1.1 Wide-focus clauses

Before moving on to a consideration of embedded clauses, it is worth mentioning an observation made by Ledgeway in his analysis of the text. Ledgeway points out that so-called thetic clauses, where the focus scopes over the entire event, generally feature the verb in initial position (Ledgeway 2017:183-184). This phenomenon is particularly frequent in the second half of the text, where the liturgical practices in Jerusalem are described, often in somewhat enumerative fashion. Notice the succession of verb initial clauses in the following sequence (343), sometimes supported by a semantically bleached adverb *sic*, already marking little more than temporal progression (one of the major functions of its Old French descendant *si*):

- (343) *Intrat episcopus intro cancellos Anastasis, dicitur unus*
 enter-3SG bishop-NOM inside railings-ACC Anastasis say-PASS-3SG one
ymnus, et [sic] facit orationem episcopus pro eis, et
 hymns-NOM and thus make-3SG prayer-ACC bishop-NOM for them-ABL and
[sic] uenit ad ecclesiam maiorem cum eis... (38.2)
 thus come-3SG to church-ACC great with them-ABL
 ‘The bishop steps inside the railings of the Anastasis, a hymn is sung, and then the bishop makes a prayer for them, and then he returns to the great church together with them...’

It is often assumed in theoretical syntax that wide-focus clauses of this kind are particularly revealing with respect to unmarked word order, since there is neither narrow focus, nor a topic-comment articulation which could serve to displace any constituent, and Ledgeway therefore concludes that the unmarked word order in *Egeria* is VSO, which is consistent

³⁸On a couple of occasions, *Egeria* employs the accusative case rather than the nominative for the postverbal subject of a passive verb:

- (i) ... *et [sic] fit orationem pro omnibus...*
 and such happen-3SG prayer-ACC for all-ABL
 ‘...and then a prayer is made for everyone...’

Adams suggests that this phenomenon, not infrequent in later Latin, signals a psychological affinity between direct objects and the subjects of passives (Adams 1976b). This fits nicely with the assumption in transformational grammar that the subjects of passives and unaccusatives receive a theme or patient theta-role.

with his analysis of the text as featuring V-to-C movement (Ledgeway 2017:182-183). This is essentially the same analysis as the one adopted for early Old High German in Hinterhölzl and Petrova (2010). However, while the fact that wide-focus clauses tend to feature the verb in initial position must clearly be considered another argument in favour of a VSO-setup, it does not in itself strengthen the V-to-C-hypothesis, since a verb in I⁰ would also yield these strings, as long as the subject is low in the structure.

Interestingly, there are also several examples of what may plausibly be considered wide focus with the subject in preverbal position, with or without an initial scene-setting element. In (344), there are two consecutive clauses featuring a preverbal, indefinite and discourse-new subject. While these cannot really be interpreted as topics, it is possible to argue that these clauses feature narrow (constituent) focus on the subject, an analysis that receives some support from the discontinuous structure of the first subject *ecclesia*, whose adjectival modifier *pisinna* is directly postverbal. If this is a case of head-fronting under hyperbaton, the stranded adjectival modifier in fact reveals the unmarked postverbal position of the subject. In the second clause of (344), or in (345), no such argument can be used:

- (344) *In eo ergo loco [ecclesia est pisinna] subter montem, non*
 in that thus place-ABL church-NOM is small-NOM under mountain-ACC not
Nabou, sed alterum interiorem: sed nec ipse longe est de Nabou.
 Nabo but other more-interior-ACC but not same far is from Nabo
[Monachi autem plurimi] commanent ibi uere sancti...
 monks-NOM several remain-3PL here truly holy
 ‘In that place there is a small church under a mountain, not the Nebo, but another
 one further in, yet not far away from the Nebo. Many monks reside there, truly holy
 men...’
- (345) *Sed ut redeam ad rem, [monasteria ergo plurima] sunt*
 but that return-SBJV-1SG to thing-ACC monasteries-NOM thus several are
ibi per ipsum collem et in medio murus ingens... (23.4)
 there at same hill-ACC and in middle-ABL wall-NOM huge
 ‘but to return to my story: there are several monasteries there on the hill and in the
 middle a great wall...’

In general, it is extremely difficult to tease apart readings with narrow focus on the subject from clause-wide focus. The argumentation easily becomes circular, since it is the very preverbal position of the subject that is used to support the claim that such narrow focus movement has taken place. At the very least, we must envisage the possibility that there might be more than one unmarked position for the subject.

The conclusion from main clauses is therefore that inversion with transitive verbs is far from uncommon, but the data we have been considering so far does not make it possible to say anything definite about the structural position of the verb. Furthermore, the main clause shows at least as strong tendencies for an unmarked VSO pattern as for a SVO pattern. The question is how to interpret these surface data in terms of syntactic structure, which amounts to asking what kind of structure the children acquiring the language would assign to such strings. This question cannot be answered by restricting our attention to main clauses, as the global input must be taken into consideration. We will therefore proceed to consider embedded clauses.

5.6 Itinerarium Egeriae: embedded clauses

Before we consider any data from embedded clauses, it is relevant to recall in more general terms what we might expect from embedded data as opposed to main clauses. First, as a general synchronic insight, it is widely held that embedded clauses are somewhat *pragmatically impoverished*, containing fewer displacement operations than the root clause (Hooper and Thompson 1973; Cruschina 2010). As a consequence, we might expect that the unmarked word order(s) might appear more clearly than what is the case in main clauses. In chapter 4, this was shown to be the case for Old French, where the unmarked SVO order is particularly dominant in embedded clauses.

Second, this time at a diachronic level, it has been observed that embedded clauses are conservative in that they may preserve for a longer period of time old word order patterns that have declined or disappeared completely in main clauses (Givón 1971, Dixon 1994:206-207, Harris and Campbell 1995:27). In the research literature on Latin, this argument has been used to account for the fact that SOV structures seem to endure longer in embedded clauses (Adams 1976a; Bauer 1995; see also Danckaert 2017b:113).³⁹

And finally, it is very relevant to our concerns that V-to-C movement is much more restricted in embedded clauses than in main clauses, although this asymmetry is not categorical, since the evidence from modern Germanic V2 languages as well as from Old French in chapter 4 clearly shows that V-to-C movement is possible in a narrowly definable subset of embedded clause types, in particular complement clauses under viaduct verbs (cf. section 2.3.3) and certain ‘peripheral’ adverbial clauses which permit high syntactic attachment to their matrix clause (cf. section 2.3.5).

With these considerations in mind, we now turn to the data.

5.6.1 Linear distribution of the verb

The different categories of embedded clause that were annotated were adverbial clause, complement clauses, interrogative and relative clause. The relative frequency of these is very unevenly distributed with 450 adverbial clauses, 95 complement clauses, 25 interrogatives and 688 relative clauses. The main focus will be on adverbial and complement clauses, since relative and interrogative clauses have rather different syntax, presumably creating the bond to their matrix clause through movement of a phrase to the left periphery rather than by lexicalising a C-head, a fact which complicates their analysis considerably.⁴⁰ Recall also from chapter 2 that relative and embedded interrogative clauses are generally staunch non-V2 domains across Germanic, with the notable exception of relative clauses in certain

³⁹In a very influential survey (according to Panhuis (1984), *too influential* – since it has been central in propagating the claim that Latin word order is basically SOV), Linde (1923) collected data on verb-final orders from around 20 different authors ranging from Cato (*De re agri cultura*) to Victor Vitensis (*Historia pers.*); in every single work examined (including *Egeria*), verb-finality was more frequent in embedded clauses than in main clauses; see Ledgeway 2012:226 for the results presented in table form.

⁴⁰At least two problems arise when considering the syntax of relative (and interrogative) clauses, related respectively to the filler and the gap of the dependency created by the relativised phrase. The first is to decide where in the left periphery the fronted XP lands, if this position is the same for subjects and non-subjects, and on a more practical level, if this constituent should be counted when considering linear order. Secondly, there is the question where inside the clause the gap of the moved element is situated. This is particularly challenging in a free word-order language like Latin, and yet the question is completely crucial in the case of subject relatives, since the information is needed to establish a ‘type’ of the kind used in the annotation adopted here, or to calculate the rate of clause-internal inversion. A thorough discussion of these problems will be available in the user manual that follows the data files in the TROLLing Repository (Klævik-Pettersen 2018).

varieties of German (see section 2.3.5). The same was true of the Old French corpus in chapters 3 and 4, where there was no sign of V2 in relative or interrogative clauses.

Table 5.5 shows the overall linear distribution of the verb in complement and adverbial clauses. When comparing the figures in the table with the corresponding figures in table 5.1 from main clauses, it is immediately clear that there is a certain quantitative asymmetry between main and embedded clauses. The most salient difference from main clauses is that V1 clauses have increased significantly from 18.22% to 30.46%. V2 clauses have in fact also increased somewhat from 34.57% to 40.00%, whereas all $V \geq 3$ orders have decreased significantly.

Table 5.5: Linear order of the finite verb in complement clauses (95 tokens) and adverbial clauses (450 tokens) in *Egeria*

	Transitive	Unaccusative	Copula	Functional	Total
V1	54 (34.18%)	74 (32.31%)	15 (13.89%)	23 (46.00%)	166 (30.46%)
V2	66 (41.77%)	96 (41.92%)	48 (44.44%)	8 (16.00%)	218 (40.00%)
V3	29 (18.35%)	29 (12.66%)	34 (31.48%)	12 (24.00%)	104 (19.08%)
V4	6 (3.80%)	17 (7.42%)	9 (8.33%)	4 (8.00%)	36 (6.61%)
V5	3 (1.90%)	10 (4.37%)	2 (1.85%)	3 (6.00%)	18 (3.30%)
V6	– (0.00%)	3 (1.31%)	– (0.00%)	– (0.00%)	3 (0.55%)
Total	158 (100.00%)	229 (100.00%)	108 (100.00%)	50 (100.00%)	545 (100.00%)

Average constituent count: ≈ 2.92

Verb-final strings: $237/545 = 43.49\%$

Null-subjects: $330/545 = 60.00\%$

It is also interesting to observe that, alongside the considerable increase in V1 strings and the concomitant drop in $V \geq 3$ strings, the amount of verb-final strings has increased from 28.25% to 43.49%. Some of the explanation for this apparent paradox can already be found in the average constituent count, which reveals that adverbial and embedded clauses are simply much shorter than main clauses. The fact that verb-early strings (V1, V2) and verb-final strings may both statistically increase considerably in the passage from one clause type to another shows how extremely dangerous it is to draw any kind of conclusion, however cursory, from the examination of linear distribution patterns alone. For instance, Ledgeway interprets the paucity of $V \geq 3$ strings as an indication that SOV is hardly productive in embedded clauses (Ledgeway 2017:195). It is clear from the passage that what Ledgeway has in mind is not the string SOV as such (which clearly cannot exist in V1 or V2 strings), but in more general terms head-final verbal projections. Such a conclusion cannot be drawn on the basis of linear evidence, however. Admittedly, it cannot be drawn on the basis of verb-final strings either, since verb-finality in linear terms does not equal head-finality in syntactic terms.

5.7 Head-finality in the IP

I will in fact start by addressing the question of head-finality, since it is completely fundamental to the overall understanding of the embedded syntax. A more detailed and reliable assessment of the distribution patterns of head-finality therefore becomes necessary. We will try to approach the matter from two slightly different angles. First, we consider the alternation between VO-OV (with or without the subject expressed) in transitive clauses in adverbial and complement clauses. Secondly, we will look at the relative order of infinitives (excluding the AcI and NcI constructions) and their selecting verbs, the group which has (somewhat dubiously) been called ‘functional’ predicates.

Starting with the VO/OV alternation, it must be emphasized that this is a rather uncertain probe into head-directionality. First, we must exclude all objects which are either pronominal or clausal, since the former might well be clitics and in any case show a strong tendency to gravitate towards the left, while the latter on the other hand are virtually always right-dislocated. Even with this precaution, it is clear that these string types are ambiguous and that one can neither equate VO with a head-initial or OV with a head-final IP. For instance, surface VO order could also arise from a head-final IP either through V-to-C movement of the verb or by movement of the object to a right-peripheral position (for instance in cases of narrow focus or for prosodic reasons/reasons of ‘heavyness’). In either case we would not be able to say anything about the headedness of the IP. However, both of these options are somewhat marked, in particular V-to-C movement, which by default is expected not to happen very frequently in embedded clauses. Therefore, VO is *likely* to be the expression of a head-initial IP in many cases.⁴¹

In a similar vein, surface OV order could arise in a head-initial IP through topicalisation of the object to the left periphery, combined or not with V-to-C movement of the verb, or possibly also by scrambling of the object to a high, clause-internal position above the IP (this time necessarily without concomitant V-to-C movement). Once again, these must be considered marked cases, particularly left-peripheral topicalisation, which by default should be blocked in embedded clauses. As for the scrambling operation, we shall see some evidence suggesting it exists, but it certainly does not seem to be frequent. OV strings might therefore also be expected to correlate statistically with a head-final IP, or at least give us a rough first impression of the matter.

The second probe in principle works the same way, as it is based on the assumption of a certain statistical connection between the order functional verb (hence abbreviated *Aux* – without implying true auxiliary status) - infinitive and head-initial IPs, and vice versa a connection between infinitive – functional verb and head-final IPs. The same caveats as before apply, but with very different associated probabilities. *Aux-Infinitive* order could also be produced by V-to-C movement of the verb with presumably the same probability as in VO strings. However, *Aux-Inf* could also be produced by movement of the infinitive to a right-peripheral position. This is much more likely to be the case than with the corresponding VO string, simply because infinitival constructions are very heavy constituents that are likely

⁴¹Note that in this section (and in this chapter in general), I am making the assumption that finite verbs always raise as high as I⁰. The same assumption was not made for Old French, where it was in fact possible to demonstrate empirically that all finite verbs move as high as I⁰. In *Egeria*, it is possible (with some minimal assumptions) to prove that finite verbs at least sometimes move as high as I⁰, and we will therefore assume that they always do so. This is nothing but a deed of necessity; if we were to consider the possibility that the verb sometimes does not raise out of the VP, the amount of combinatorial possibilities and hence the structural ambiguity would simply grow out of control, in particular since the VP itself can demonstrably be both head-initial and head-final.

to be placed in extraposition. Furthermore, the fact that the class of ‘functional’ verbs is highly heterogeneous and almost certainly includes constructions which are biclausal control structures rather than monoclausal raising structures makes this probability even higher, as clausal constituents very often tend to be placed in extraposition even regardless of their weight. For instance, finite complement clauses in Latin generally follow their selecting verb even when they are very light, a tendency already established in Classical Latin among authors with a predilection for verb-final patterns. In *Egeria*, not a single complement clause is incorporated into the matrix clause, not even when they consist of a single word, like (346):⁴²

- (346) ... *et perfecta sunt singula, quae iusserat Deus in*
 and made were all-things-NOM which-ACC order-PLPRF-3SG God-NOM in
montem Moysi, ut fierent. (5.9)
 mountain-ACC Moses-DAT that happen-IPFV-SUBJ-3PL
 ‘and all things were accomplished that God had bidden Moses on the mountain that
 they should be made.’

As for the order *Inf-Aux*, the situation is quite different. This string should be a reasonably reliable indicator of head-finality in the IP. The only way it could fail to be so, is through topicalisation of the infinitive to the left periphery (with or without concomitant V-to-C movement), piecemeal ‘evacuation’ of every constituent in the VP (or whatever constituent the infinitival construction corresponds to in each case) to the left periphery (cf. example (331), or scrambling of the VP/Infinitival construction to a position above the verb in I⁰ (without V-to-C movement). All of these scenarios must be considered extremely marked, particularly in embedded clauses, where the left periphery in the default case is not accessible. Although one certainly cannot categorically exclude the possibility that such structures might arise (although the ‘evacuation’ scenario certainly borders on the impossible), they will not occur very frequently. As a quite sturdy generalisation, amply documented from studies of modern languages, simple constituents often move towards the left (since they are often targeted by discourse-related fronting operations like topicalisation, focalisation or scrambling, plus possibly for prosodic reasons), while heavy, higher-order constituents like VPs do not tend to move as easily towards the left (since they are much more rarely targeted by fronting operations), while they do often move to the right for reasons related to prosody or ‘heaviness’ (cf. the ‘Gesetz der wachsenden Glieder’ (Behagel 1909)). Schematically, then, we could suggest the following:

⁴²The AcI construction is also generally placed in extraposition, although it is sometimes incorporated, and very frequently (multiply) discontinuous.

Table 5.6: Reliability of different probes into the head-directionality of IP in embedded clauses

	Head-initial IP	Head-final IP
VO	Moderately reliable	–
OV	–	Moderately reliable
Aux-Inf	Unreliable	–
Inf-Aux	–	Reliable

Table 5.7: VO-OV alternation in embedded clauses

Clause type	VO	OV	Total
Adv. & Comp.	56 (53.85%)	48 (46.15%)	104 (100.00%)
Relative	58 (68.24%)	27 (31.76%)	85 (100.00%)

Table 5.8: Aux-Inf, Inf-Aux alternation in embedded clauses

Clause type	Aux-Inf	Inf-Aux	Total
Adv. & Comp.	29 (65.91%)	15 (34.09%)	44 (100.00%)
Relative	31 (67.39%)	15 (32.61%)	46 (100.00%)

With this in mind, we can now consider the actual distribution, which is given in table 5.7 and table 5.8. Corresponding data from relative clauses are also provided, since these reveal a first clear indication of an asymmetry between different types of clauses that will prove important for the general understanding of the embedded syntax. As table 5.7 indicates, the order VO has only a rather slight edge on the order OV in adverbial and complement clauses, while VO is more than twice as common as OV in relative clauses. Table 5.8 does not show any such asymmetry; on the other hand it clearly suggests that head-finality of the IP is quite robust, if indeed the order *infinitive-aux* is as reliable as was suggested above. On a qualitative level, it is certainly not hard to find good examples of strings which qualify as *bona fide* head-final clauses of a completely classical pattern (as word order goes), witness (347)–(351):

(347) *sicut [et illi sancti] [singula] [nobis] [ostendere] dignabantur*
as also the holy-NOM all-things-ACC us-DAT show-INF deig-IPFV-3PL

(4.2)

‘thus the holy men deigned to show us every single place.’

- (348) *Et licet [semper] [Deo] [in omnibus] [gratias agere]*
and although always God-DAT in all-things-ABL thanks-ACC give-INF
debeam (5.12)
should-SBJV-1SG

‘and although I ought always to give thanks to God in all things...’

- (349) *si [qua] [preterea] [loca] [cognoscere] potuero...* (23.10)
if some-ACC thereafter places-ACC know-INF be.able-FUTPRF-1SG
‘if later I shall be able to see some other places...’

- (350) *diligentius et securius iam in eo loco ex consuetudine*
More-diligently and more-safely now in that place-ABL from habit-ABL
Faranitae ambulant nocte quam [aliqui hominum] [ambulare]
Faranites-NOM travel-3PL night-ABL than some-NOM men-GEN travel-INF
potest in his locis, ubi uia aperta est. (6.2)
be.able-3SG in those places-ABL where road-NOM open is

‘through experience, the Faranites travel with greater precision and with more safety at night than other people can travel in those places, where the road is clear.’⁴³

- (351) *quia [prorsus] [nec] [in sella] [ascendi] poterat...*
because absolutely NEG in saddle-ABL ascend-INF.PASS be.able-IPFV-3SG
‘since it (i.e. *the mountain*) could under no circumstance be ascended while in the saddle...’

On the strength of such evidence, we might conclude that the tendency for the IP to fall into a head-final, left-branching pattern is appreciably stronger in adverbial and complement clauses than in main clauses. This is in line with traditional accounts like Linde’s (1923), and corroborates the claim made by Adams that ‘[i]n Latin of all periods, including that of very late Antiquity, final position of the verb was appreciably more common in subordinate than in main clauses (Adams 1976a:93, fn.61)’. We might interpret this as evidence for the diachronic conservatism of embedded clauses mentioned in the introduction to this section; embedded clauses, and perhaps in particular adverbial and complement clauses, to some extent retain a complementation pattern which seems to be strongly on the decline in the more innovative root clauses.

This also gives us a very natural way of accounting for the very frequent OV strings. In embedded clauses, we do not expect topicalisation or any other movement operations to the left periphery to be generally available, since the complementiser by default lexicalises the lowest C-head, Fin⁰. At the same time, evidence both from V2 languages and non-V2 languages show that embedded left-peripheries are sometimes available in the presence of a complementiser, in particular in clauses which are complements of viaduct verbs and in certain *peripheral* adverbial clauses. We must therefore clearly envisage the possibility

⁴³Notice how the verb is not clause-final in (350) since the PP *in his locis* has been extraposed because of the relative clause. This principle is very strong in *Egeria* and is therefore a quite common source of right-displacement. It is unclear if it is a purely prosodic phenomenon, since the resulting constituent is always quite heavy, or if it is triggered by the syntax to provide adjacency between the relative pronoun and the correlate.

that a subset of embedded clauses might feature XP-movement to the left periphery, and possibly also V-to-C movement, if this option turns out to exist in the grammar of *Egeria*. However, from there and to Ledgeway's claim, that all non-V1 embedded clauses are cases of embedded V2 (Ledgeway 2017:198), is quite a long step. Ledgeway finds somewhat more V1 in embedded clauses than what was found in my annotation, but even by his own figures, such a claim would entail that approximately 60% of all embedded clauses feature embedded V2. Such a percentage is disproportionate to anything similar reported from either the modern V2 languages or the Old Romance languages for which a V2 hypothesis has been proposed (Wolfe 2015b). By comparison, recall that only around 12-13% of embedded clauses in our Old French texts were candidates for V-to-C movement and that even this figure was clearly too high, given the fact that many instances of apparent V2 should rather be analysed as Stylistic Fronting, which does not involve V-to-C movement at all. Even if we loosened the strong assumption that access to the embedded left periphery is universally defined by S-selectional features of the matrix verb or the possibility of parataxis ('high attachment') in certain adverbial clauses, the discrepancy here is simply unrealistically large. The hypothesis is also further weakened by the fact that the type of embedded clause seems irrelevant to the availability of $V \geq 2$ strings, and that many supposed cases of embedded V-to-C movement would have to involve multiple frontings, witness (352)–(355):

- (352) *Quarta feria autem et sexta feria, quoniam [ipsis diebus] [penitus]*
 fourth weekday PRT or sixth weekday-ABL since these days-ABL hardly
[nemo] ieiunat, in Syon proceditur. (41.1)
 nobody-NOM fasts in Syon proceed-3SG-PASS
 'On Wednesdays or Fridays, since hardly anybody fasts on these days, they go to Syon.'
- (353) *Illud autem [...] fit et ualde admirabile, ut [semper] [tam ymni quam]*
 This was also very admirable that always both hymns-NOM as
antiphonae et lectiones] [...] [tales pronuntiationes] habeant,
 antiphons-NOM and readings-NOM such pronunciations-ACC have
ut...
 that... (47.5)
 'And this was also very impressive, that both the hymns as well as the antiphons or the passages read aloud always had such a content that...'
- (354) *... ego desideraueram semper, ut, [ubicumque uenissemus], [semper]*
 ...I desire-PRF-3SG always same place-NOM from book-ABL
[ipse locus de libro] legeretur. (4.3)
 read-3SG-PASS-SUBJ
 '...I always wished that, wherever we came, the corresponding passage from the Bible should always be read.'
- (355) *quoniam [episcopus], (licet siriste nouerit), tamen [semper] [grece]*
 since bishop-NOM although Syriac knows always in-Greek
loquitur... (47.3)
 speaks...
 'since the bishop, although he knows Syriac, always speaks in Greek...'

Finally, it is worth calling attention to a series of interesting near-minimal pairs, observable in *Egeria* just like in Classical Latin; some relevant examples are given in (356)–(359) below. In some embedded clauses, notably, but not exclusively in adverbial *cum*-clauses, a constituent in fact appears to the left of the complementiser. By the logic of our theory, this constituent is unambiguously in the left-periphery. The phenomenon is particularly frequent when the left-peripheral element is itself a relativised phrase (356a)–(357a), yielding various kinds of embedded ‘pseudo-relatives’, but it is also quite often encountered with a regular XP with a topical (358a)–(359a) or (somewhat more rarely) a focal reading. In some cases, more than one constituent appears to the left of the complementiser.

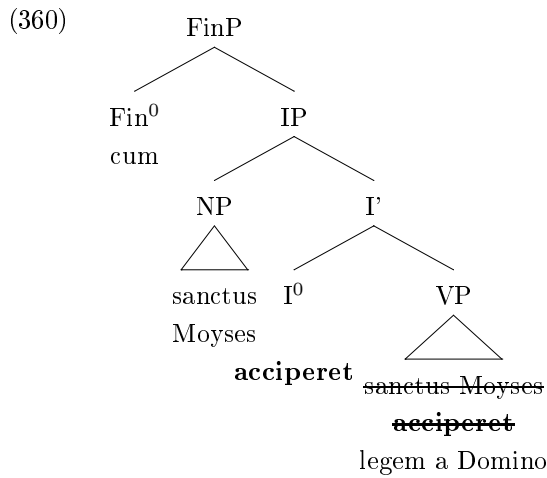
Clearly, it seems like complementisers in Latin do not really block movement to the left-periphery in all cases. But if this is the case, why should we assume that constituents that follow the complementiser are also left-peripheral? This would seem to force us to assume that one and the same complementiser, without any noticeable semantic difference, sometimes stays in a low left-peripheral position such as Fin^0 , from which it allows XPs to cross it, and that it sometimes is merged in or raised to a high left-peripheral position such as Force^0 , (redundantly) opening the left-periphery below it. On the contrary, if we adopt the natural null-hypothesis that the complementiser sits in Fin^0 , we can interpret the contrast between (a) and (b) in the following examples as involving left-peripheral and clause-internal constituents, respectively:

- (356) a. *[Ad quem puteum] cum uenissemus, facta est ab episcopo*
 To which well-ACC when come-PLPRF-SBJV-1PL made is by bishop-ABL
oratio... (21.1)
 prayer-NOM
 ‘When we came to that well, a prayer was said by the bishop ...’
- b. *Cum ergo uenissemus ad portam ipsam... (19.16)*
 when thus come-PLPRF-SBJV-1PL to gate-ACC same-ACC ...
 ‘when we came to the gate...’
- (357) a. *[Quod] cum dixisset, nos satis auidi*
 which-ACC when say-PLPRF-SBJV-3SG we very avid-NOM
optati sumus ire... (10.9)
 opt-PST-PTCP be-1PL go-INF
 ‘When we heard this, we very eagerly wished to go...’
- b. *cum [haec] [ad uestram affectionem] darem... (23.10)*
 when these-ACC to your-ACC affection-ACC give-IPFV-SBJV-1SG
 ‘When I send this (i.e. letter) to your affection...’
- (358) a. *... [sanctus Moyses] cum pasceret pecora*
 ... holy-NOM Moses-NOM when feed-IPFV-SUBJ-3SG cattle-ACC
soceri sui, iterum locutus est ei
 brother-in-law-GEN his-REFL-GEN again spoke-PST.PTCP is him-DAT
Deus de rubo in igne. (2.3)
 God-textitNOM from bush-ABL in fire-ABL
 ‘... when holy Moses was out feeding the herd of his brother-in-law, God spoke to him again from the burning bush.’

- b. ... *cum* [*sanctus Moyses*] *acciperet* *a Domino*
 ... when holy-NOM Moses-NOM receive-IPFV-SUBJ-3SG from lord-ABL
legem ad filios Israhel.
 law-ACC to sons-ACC Israel.
 ‘... when holy Moses received the law from the Lord for the children of Israel.’
- (359) a. [*hii fontes*] *ubi erupeierunt, ante sic fuerit campus intra*
 these sources where erupted before thus had-been field-NOM inside
ciuitatem... (19.4)
 city-ACC
 ‘where these sources erupted, earlier there had been an open ground inside the city...’
- b. *Sed postmodum quam* [*hii fontes*] [*in eo loco*]
 but later than these sources-NOM in that place-ABL
eruperunt ... (19.15)
 erupt-PLPRF-3PL
 ‘but after these sources had erupted in that place...’

On this interpretation, the (b) examples provide further evidence that there are preverbal, yet non-left-peripheral positions in embedded clauses.

It is clear that a head-final IP will result in a host of (linearly) preverbal positions, since in fact the entire clause will linearly precede the verb in such a configuration. However, example (358b) above is more informative than the other strings, since it clearly does not seem to involve a head-final IP, witness the presence of two postverbal constituents. We must therefore conclude that the IP is in fact head-initial here, and yet the subject precedes it. This is crucial, since the proper analysis of the inversion strings we observed in main clauses is still very much pending, and this analysis depends heavily on the position of the subject. Recall that the data on the subject position in main clauses was somewhat less than optimal, but that it overall tended to suggest an unmarked postverbal subject in a rather low position, presumably Spec-vP. However, this cannot be the position of the subject in (358b), where we might instead hypothesize something like the structure in (360). If the global input supports the hypothesis of such a high preverbal subject position, the V-to-C hypothesis would also be considerably strengthened.



5.8 Inversion and the position of the subject

We will now consider the position of the subject in embedded clauses, starting with adverbial and complement clauses. Table 5.9 provides information on how the three logical options *preverbal subject*, *postverbal subject* and *null subject* distribute in total and across different predicate classes.

The table reveals many clear asymmetries compared to main clauses. First, recall that the rate of inversion in main clauses was at 32.53%. In comparison, adverbial/complement clauses show considerably less overt inversion, although 13.94% postverbal subjects are far from a marginal phenomenon. Second, inversion does not seem to react strongly with the variable ‘predicate class’, except for a not surprising tendency for inversion with the copula. The relative preference for inversion with unaccusative predicates compared to transitives has almost vanished entirely and is no longer statistically significant. And finally, the amount of null-subjects in adverbial and complement clauses is much higher than in main clauses (60.00% vs. 35.32%). Both of the latter tendencies are presumably the result of a functional difference between main and embedded clauses, since the latter do not introduce new discourse referents as readily as main clauses.⁴⁴

⁴⁴As was noted earlier, 7.43% of all main clauses introduce new discourse referents by way of the subject position. In complement and adverbial clauses, this figure has sunk to 3.55%. As a consequence, the degree of inversion with transitives and unaccusatives is almost levelled out. This strengthens our hypothesis that the relatively greater tendency for inversion with unaccusatives over transitives in main clauses is at least largely due to a difference with respect to how these predicates are used in discourse.

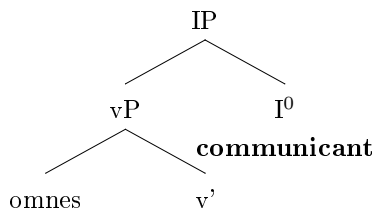
Table 5.9: The position of the subject (S) in complement and adverbial clauses in *Egeria*

	Transitive	Unaccusative	Copula	Functional	Total
Preverbal S	36 (22.64%)	59 (26.11%)	38 (34.55%)	9 (18.00%)	142 (26.06%)
Postverbal S	16 (10.06%)	29 (12.83%)	27 (24.55%)	4 (8.00%)	76 (13.94%)
Null S	107 (67.30%)	138 (61.06%)	45 (40.91%)	37 (74.00%)	327 (60.00%)
Total	159 (100.00%)	226 (100.00%)	110 (100.00%)	50 (100.00%)	545 (100.00%)

The apparently lower frequency of inversion compared to main clauses constitutes an important *explanandum*. Also, it is in fact a surprising finding for any hypothesis that considers the grammar of *Egeria* to be a staunchly head-initial VSO grammar with a low subject position, regardless of whether the verb is assumed to move to C^0 or to I^0 . The reason is simple: if the verb dominates the subject position in unmarked word order, the only way the subject can precede the verb is either by moving across it, for instance through topicalisation or focalisation into the left periphery, or by right-dislocation. At least topicalisation and focalisation are typically root phenomena which are much more restricted in embedded clauses, and hence there should be fewer contexts for the subject to move across the verb. A priori, we would rather expect inversion to increase in embedded clauses in a head-initial VSO grammar.

This suggests that something is wrong with the above hypothesis. Since it makes two different assumptions, there are also two possible explanations that immediately come to mind. The first one is that the rather tentative hypothesis of a low subject position in Spec-vP established in the section on main clauses was in fact misguided. If the subject position is in fact higher, such as Spec-IP, the inversion structures in main clauses could only have arisen through V-to-C movement. Since V-to-C movement is much more restricted in embedded clauses, this forces the verb to stay in I^0 , meaning the subject in Spec-IP will generally precede it, causing less inversion. Indeed, it seems tempting at first sight to draw the conclusion that, since preverbal subjects are almost twice as frequent as postverbal subjects, this might indicate a preferential subject position in Spec-IP or some other high projection. However, this conclusion does not follow at all, since subjects will also be preverbal in Spec-vP when combined with a verb in a head-final IP; cf. example (333), repeated for convenience here:

- (333) *sed omnes communicant*
 but all-NOM communicate-3PL
 ‘but all communicate...’



This suggests that the problem with the hypothesis of the staunchly head-initial VSO language is not necessarily the low subject position, but rather the assumption that the IP is always head-initial. In the previous section, we already witnessed very strong evidence suggesting that the IP is in fact head-final in many embedded clauses, as evidenced by the fact that one third of all embedded clauses feature *infinitive - functional verb* sequences; cf. table 5.8.

We must elaborate a bit on the inversion mechanism that is inherent to the head parameter of the IP. As just stated, a final IP will result in a preverbal subject in Spec-vP. If the IP on the other hand is head-initial, this will create inversion when the subject is in Spec-vP. In contrast, a subject in Spec-IP will remain preverbal regardless of the headedness of the IP. Disregarding for the moment other possibilities, such as embedded V-to-C movement or right-peripheral subjects, this schematically gives the following possibilities:

- (361) a. [Fin^0 [IP [vP **subject**] I' **verb**]] (Subject in Spec-vP, head-final IP)
 b. [Fin^0 [IP [I' **verb** [vP **subject**]]]] (Subject in Spec-vP, head-initial IP)
 c. [Fin^0 [IP **subject** [vP...] I' **verb**]] (Subject in Spec-IP, head-final IP)
 d. [Fin^0 [IP **subject** [I' **verb** [vP...]]]] (Subject in Spec-IP, head-initial IP)

Since three out of four possible combinations yield an order SV..., this state of affairs entails that we cannot really say much about the structural position of the subject from quantitative inversion facts like the ones presented in table 5.9. In particular, we can say nothing about how often (or if at all) the subject occupies Spec-IP (or another high projection) in embedded clauses, since a parse with the subject in Spec-vP could in fact underlie all the data on table 5.9. Furthermore, the fact that there is such significant inversion already strongly suggests that Spec-vP is a very frequent subject position, since the Spec-IP parse can only give inversion strings in conjunction with V-to-C movement or right-dislocated subjects.⁴⁵ When we consider that a sizeable majority of adverbial and complement clauses in fact lack an overt subject (60.00%), it is clear that the inversion facts point to a fundamental difference between the syntax of *Egeria* and the Old French texts reviewed in chapters 3 and 4. In the Old French texts, there was considerably less inversion even though null-subjects are very rare in embedded clauses in Old French (averaging 6.5 – 7.00% in *Tristan* and only 4.5% in *Eustace*).

The combinatorial possibilities in (361) offer us an interesting hypothesis regarding the asymmetry between main clauses and embedded clauses. If we assume that the subject position is, or at least can be, Spec-vP in both main and embedded clauses, then a stronger tendency for head-final IPs in embedded clauses, which we have already established, will suffice to produce less inversion in main clauses. Furthermore, it will provide a natural explanation of the very numerous cases where there are several constituents in front of the verb, more natural than assuming that these strings arise through embedded V-to-C movement coupled with multiple XP-fronting to the left periphery.

⁴⁵This is strictly speaking only true as long as we restrict the possible subject positions to Spec-vP and Spec-IP and the possible positions for the finite verb to I⁰ and C⁰. Once we start considering more positions for the subject and verb, other combinatorial possibilities arise. For instance, in Danckaert (Danckaert 2017b,a), there is both a functional projection FP and a even higher projection GP between CP and IP, while there is also a position for the subject, SubjP, between GP and FP. These projections arise, it seems to me, partly due to Danckaert's ambition to derive the Latin clause structure in compliance with the LCA and the FOFC. On the base-generation approach adopted here, I will not adopt more positions unless compelled by the evidence.

Table 5.10: The syntax of embedded inversion: some possible structures

	Head-initial IP	Head-final IP
S in Spec-vP	VS...	SV...
S in Spec-IP	SV...	SV...

In order to gain a better understanding of the position of the subject, we must consider some qualitative evidence. There are at least three different probes into the matter: the relative position of the subject and the verb with respect to adverbs, other verbal arguments, and infinitival constructions. Needless to say, long strings which spell out a combination of several of these projections are the most reliable evidence. Based on this kind of evidence, all adverbial and complement clauses which featured an overt subject were classified manually according to whether the strings favour a parse with the subject in Spec-vP or in Spec-IP. 11 strings were excluded in which the subject preceded the complementiser, revealing an unambiguously left-peripheral position.

5.8.1 The position of the subject

In the following discussion, I will use the term *unambiguous* several times, and it is therefore necessary to clarify exactly what is meant by it. The strings which are found in embedded clauses can under certain conditions be considered unambiguous *if and only if* one adopts the assumption that the word order alternations under consideration can only unfold clause-internally, that is without recourse to either the left or right peripheries. This might be considered a kind of default in embedded clauses, since the complementiser is assumed to sit in Fin⁰, either blocking of the peripheries or making movement to the peripheries visible (when a phrase is located to the left of the complementiser). Since this assumption is too strong in some cases, one must take care to consider the possibility of higher complementisers and embedded left peripheries in particular domains such as the complements of viaduct verbs as well as certain adverbial clauses that might permit *peripheral readings* (Haegeman 2007, 2010; see also section 2.3.5). Such potential cases will be pointed out as we consider the evidence. Also, right-dislocation for reasons of ‘heavyness’ is ostensibly possible in embedded clauses too, as we shall see. It might be the case that these constructions are created by right-adjunction to the IP or whatever is the highest projection of the clause. Such cases must also be singled out by qualitative consideration of every individual example.

Given these assumptions, we may conclude that, on a very strict interpretation of what counts as unambiguous evidence, no less than 82 strings were underdetermined with respect to the two alternative parses. For many of these strings, there is no obvious reason at all to prefer one parse above the other. This situation arises quite generally with very short strings, for instance any string *SV* (362) (cf. table 5.10), but also *SCV* strings where the constituent separating the subject and the verb is VP-internal, whether it is an internal argument (363) or the head of the VP (364), since these constituents would of course be preceded by the subject whether the latter is in Spec-vP or Spec-IP. This means that even some seemingly informative strings with many constituents and several verbs (365) cannot

help us distinguish between the two analyses:⁴⁶

- (362) ... *si* [*Deus noster Iesus*] *iusserit* (19.19)
 if god-NOM our Jesus-NOM order-PRF-SBJV-3SG
 ‘... if our Lord Jesus commands it...’
- (363) ... *ut* [*corpus*] [*subter altarium*] *iaceret* (16.6)
 so-that body-NOM under altar-ACC lie-IMPV-SBJV-3SG
 ‘... so that his body should rest under the altar.’
- (364) *sicut et* [*ceteri sancti episcopi uel sancti monachi*]
 as also other-NOM holy-NOM bishops-NOM or holy-NOM monks-NOM
 [*facere*] *dignabantur*... (20.13)
 do-INF deign-IMPV-3PL
 ‘as other holy bishops and monks deigned to do...’
- (365) *sicut* [*et illi sancti*] [*singula*] [*nobis*] [*ostendere*]
 as also those-NOM holy-NOM all-things-ACC us-DAT show-INF
dignabantur. (4.2)
 deign-IMPV-3PL
 ‘thus the holy men deigned to show us every single place.’

For quite a number of other strings, however, it might seem *a priori* like only one analysis is available, yet other considerations interfere which cast doubt on the necessity of the parse. For instance, relativisation in *Egeria* almost without exception right-dislocates the nominal to which the relative phrase is attached. As a consequence, an SVO string like (366), which would otherwise count as a clear Spec-IP parse, cannot be trusted beyond doubt; conversely, a VS string like (367), which would otherwise count as a clear Spec-vP parse, must also be discarded:

- (366) *Et at ubi* [*diaconus*] *perdixerit* *omnia*,
 and but when deacon-NOM proclaim-PRF-SBJV-3SG all-things-ACC
quae dicere habet... (24.6)
 which-ACC say-INF have-3SG
 ‘And when the deacon has said everything, that he has to say...’

⁴⁶Note that it does not matter, for the purpose of establishing the position of the subject, whether *dignor* – ‘to deign’ – in (364) and (365) is truly a monoclausal auxiliary (i–(ii) or (perhaps more likely) a control predicate selecting an infinitival clause (iii)–(iv). No matter their syntactic status, the infinitival constructions are incorporated into the core clause headed by their selecting verb. The uncertainties therefore all revolve around the status of the infinitival construction itself; for instance whether it is a VP or some larger constituent like a IP or CP. The string as a whole is syntactically ambiguous, but this does not affect the position of the preceding subject, which could be in Spec-vP or Spec-IP in any case. Omitting all constituents which are not spelled-out or lexicalised, this gives the following schematic representation:

- (i) [IP **illi sancti** [I' [vP/VP **singula nobis ostendere**] **dignabantur**]]
- (ii) [IP [I' [vP **illi sancti** [VP **singula nobis ostendere**]] **dignabantur**]]
- (iii) [IP **illi sancti** [I' [vP/VP [CP/IP/VP **singula nobis ostendere**]]] **dignabantur**]]
- (iv) [IP [I' [vP **illi sancti** [VP [CP/IP/VP **singula nobis ostendere**]]]] **dignabantur**]]

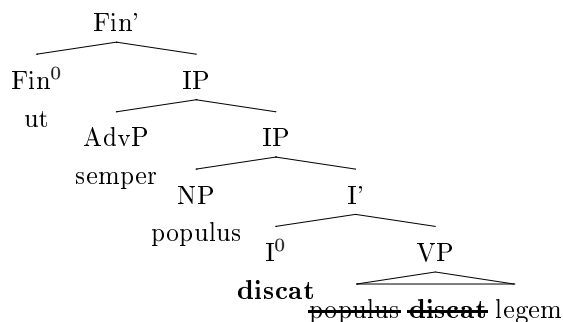
- (367) ... *ut impleantur ea, quae superius dicta sunt.* (46.4)
 so-that implement-SBJV-3PL those-things-NOM which-NOM above said
 were
 ‘... so that all the things should be implemented, which were mentioned above...’

For other strings, in turn, there is nothing that suggests extraposition of this kind, but the global evidence from the unambiguous strings reveal further syntactic positions which complicate the situation and again raise the question if there is really only one parse. The problem is related to the position of adverbs. As an illustration, consider the position of the adverb *semper* – ‘always’ – in (368). Its occurrence to the left of the subject suggests that the latter is in the low position in Spec-vP, although the rest of the string (object-verb) is compatible with a subject position in Spec-IP as well. However, in example (369), the same adverb is followed by a SVO string that cannot be generated with the subject in Spec-vP (assuming, as we do, that the finite verb always raises as high as I⁰), since in this case there is no reason to suspect that the object is extraposed. Example (369) must therefore be considered reliable evidence for the Spec-IP parse. This means that *semper* and presumably other adverbs as well may precede Spec-IP, which in turn has the repercussion that example (368) and corresponding strings cannot be taken as clear evidence for the Spec-vP parse after all.⁴⁷

- (368) *Illud [...] fit et ualde admirabile, ut [semper] [tam ymni quam antiphonae et lectiones...] [tales pronuntiationes] habeant,*
 That was also very admirable that always both hymns-NOM as
 antiphons-NOM and readings-NOM such announcements-ACC have-SBJV-3PL
ut... (47.5)
 that...

‘This [...] was also very impressive, that both the hymns and the antiphons as well as the readings[...] always have such contents, that...’

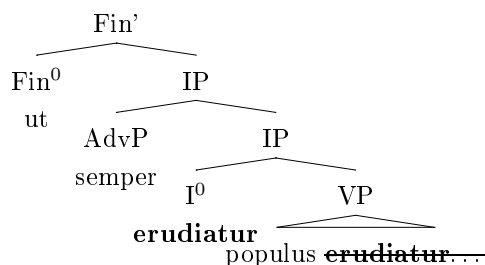
- (369) *Nam ut [semper] [populus] discat legem, et episcopus et presbyter predicant assidue.* (27.6)
 For so-that always people-NOM learn-SBJV-3SG law-ACC also bishop-NOM
 also priest-NOM preach-3PL assiduously
 ‘For in order that the people should always learn the law, both the bishop and the priest preach diligently.’



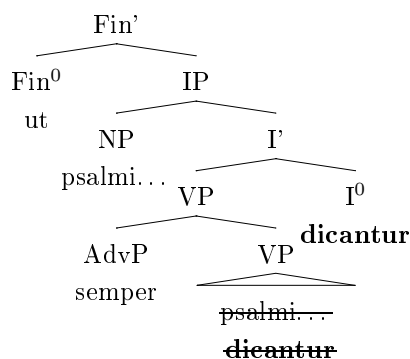
⁴⁷For convenience, adverbs are represented as Chomsky-adjoined to maximal projections rather than constituting separate projections along the clausal spine.

On the other hand, this does not allow us to draw the conclusion that *semper* in fact generally precedes Spec-IP and then use this as a firm criterion. Although other examples of this high adverbial attachment are found (example 370 is a case with a head-initial IP and a low subject), there are also strings where the same adverb is unambiguously below Spec-IP, witness (371) or the second clause of 372. In (371), the subject cannot be in Spec-vP, since *semper* intervenes between the subject and the verb, a setup which is impossible whether the IP is head-initial or head-final. If the subject is in Spec-IP, as we would have to conclude, the adverb must be below the (head-final) I⁰ head, since nothing can intervene between the specifier and the head:

- (370) *ut* [*semper*] *erudiatur* *populus* *in scripturis* *et in Dei dilectione* (25.1)
 so-that always instruct-SBJV-3SG-PASS people-NOM in scriptures-ABL and in god-GEN delight-ABL
 ‘so that the people may always learn about the Scripture and the love of God’



- (371) *ut* [*psalmi uel antiphonae apti*] [*semper*]
 so-that psalms-NOM or antiphons-NOM suitable-NOM always
dicantur (25.5)
 say-SBJV-3PL-PASS
 ‘so that suitable psalms or antiphons are always sung.’

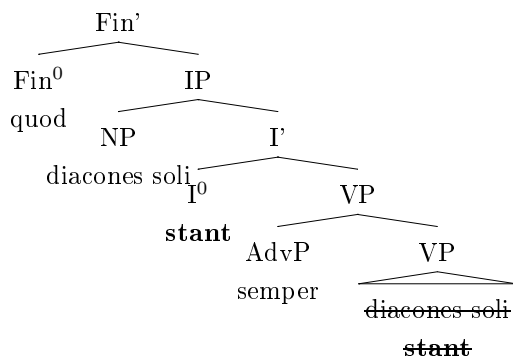


As for (372) – if it is indeed an embedded clause at all – the situation is slightly different, but the conclusion is the same; the subject *diacones* cannot be in Spec-vP, since this would entail that the IP must be head-final, which in turn makes the position of *semper* logically impossible, given that we consider all positions below the complementiser clause-internal and adverbs to always reside in left-leaning specifiers. If the subject is in Spec-IP,

the IP must be head-initial (otherwise the verb would be string-final), which in turn means the adverb must be below IP.

- (372) *nam omnis populus semper prasente episcopo*
 for all people-NOM always be-present-PRS.PTCP-ABL bishop-ABL
iubetur sedere, tantum quod [diacones soli] stant
 order-3SG-PASS sit-INF only that deacons-NOM alone-NOM stand-3PL
[semper].
 always

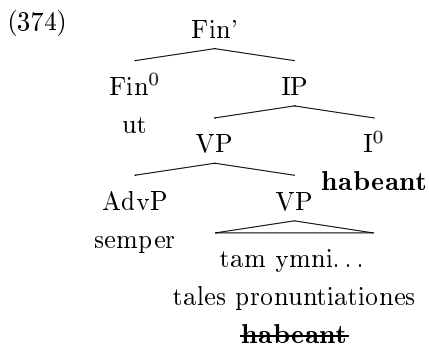
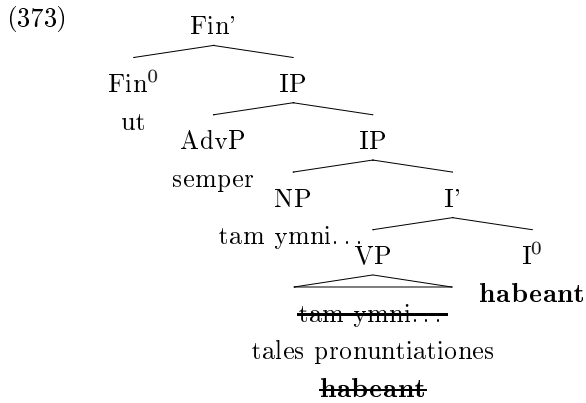
‘for all the people is always ordered to sit when the bishop is present, so that only the deacons are standing.’



It seems like the position of some adverbs relative to the IP can vary, then. It might be that this conclusion is a bit too convenient. If we take seriously the idea that the position of the adverbs are fixed unless they are moved to operator positions in the peripheries, which is a cornerstone of IP cartography as proposed by Cinque (1999), it cannot be the placement of the adverbs that varies, apart from perhaps in example (372), where the adverb could be interpreted as a narrow focus (although it would still be unclear how it could move to a peripheral position in an embedded clause). If that is the case, the variation would have to be interpreted differently from what is suggested in the trees above, meaning it is rather the verb or the subject (or possibly both) that have even more positions. Another possibility might be that the variation is only apparent and that *semper* can in fact be spelled out in two different adverbial positions with slightly different semantics and scope. Admittedly, it is hard to spot any such distinction in the examples above, although we will return to the issue shortly. It is also conceivable that the highest position, above IP, is in fact left-peripheral, and that some of these examples therefore involve high complementisers in Force⁰, rather than Fin⁰. In the latter case, movement of the adverb would be reconcilable with Cinque’s system.

For the moment, we will leave the question aside. What should be clear from this discussion is that a very substantial share of the strings are underdetermined with respect to the position of the subject. Returning to the initial example, it is clear that this uncertainty extends to case like (368) as well, since the position of the adverb is ambiguous, and hence also the position of the subject:⁴⁸

⁴⁸There is of course also a third parse, which is to keep the subject in Spec-vP and have the adverb attach high, above IP.



For the reasons just illustrated, 82 more strings were removed as undetermined with respect to the property we are investigating, namely the structural position of the subject. This leaves us with 125 adverbial and complement clauses for which it is possible to suggest with some (not absolute) certainty the position of the subject. The results are provided in table 5.11. Two clear conclusions can be drawn on the basis of these figures: (1) the preferential subject position in *Egeria* is indeed quite low, as was already suggested in main clauses; and (2) there is a higher subject position available, used in approximately one third of all adverbial and complement clauses, and which therefore cannot be considered marginal at all.

Table 5.11: The position of the subject in adverbial and complement clauses in *Egeria*

Spec-vP	Spec-IP	Total
80 (64.00%)	45 (36.00%)	125 (100.00%)

5.8.1.1 The low subject position: Spec-vP

The low subject position is in evidence in all inversion strings, since a parse with the subject in Spec-IP cannot give rise to inversion unless there is V-to-C movement, and we are basing this discussion on the assumption that complementisers and subjunctions in Latin quite generally lexicalise the lowest C-head, Fin^0 , as argued for at length in Danckaert (2012). As

table 5.9 above has shown, 13.94% of adverbial and complement clauses feature inversion, with no strong asymmetry with respect to the predicate class of the verb, apart for a slightly increased tendency for inversion with the copula; some examples are given in (375)–(379). Notice from (379) that there must clearly be positions available above the verb, since there are two preverbal constituents, but in this particular case we are in fact dealing with a complement clause under what is presumably a viaduct verb, *testor* – ‘to testify’ – and there is accordingly a very real possibility that one (or possibly both) of the preverbal constituents are left-peripheral.

- (375) *Locus etiam ostenditur ibi iuxta, ubi stetit sanctus*
 place-NOM also show-3SG-PASS there nearby where stand-PRF-3SG holy
Moyses, quando ei dixit Deus: (4.8)
 Moses-NOM when him-DAT say-PRF-3SG god-NOM
 ‘The place there nearby was also shown to us, where holy Moses stood when God said to him.’
- (376) ... *sed cum leget affectio uestra libros sanctos*
 but when read-3SG affection-NOM your books-ACC Moses-GEN
Moysi... (5.8)
 ‘... but when your Affection reads the books of Moses...’
- (377) ... *ut tamen dicebat sanctus episcopus. (20.3)*
 as say-IPFV-3SG holy bishop-NOM
 ‘... as the holy bishop told.’
- (378) ... *et dicuntur psalmi et antiphonae, donec*
 and say-3PL-PASS psalms-NOM and antiphons-NOM while
commonetur episcopus; (24.3)
 summon-3SG-PASS bishop-NOM
 ‘and psalms and antiphons are recited while the bishop is being summoned.’
- (379) *Nam uere scriptura hoc testatur, quoniam [ad*
 for truly scripture-NOM this-ACC testify-3SG that to
accipiendam sanctam Rebeccam] [huc] uenerit
 receiving-GDV-ACC holy Rebecca-ACC here come-PRF-3SG-SBJV
puer sancti Abraae... (20.10)
 boy-NOM holy-GEN Abraham
 ‘For the Scripture truly testifies that the servant of the holy Abraham came here to take the holy Rebecca...’

We have been assuming so far that this low subject position is Spec-vP, and there is nothing to suggest that this conclusion is wrong. There is even an apparent case of G-inversion which provides some evidence in favour of this parse, if we interpret (380) as a monoclausal domain with the subject ‘the bishop’ sandwiched between an auxiliary and the infinitival head of the VP. As already mentioned, this analysis is anything but clear, since the status of *coepi* as an auxiliary very much remains in doubt. However, the somewhat paradoxical combination of an inchoative verb with a telic infinitive at least suggests we should not discard the of a restructured, monoclausal domain. (381) is an example from the

relative clauses which features a much stronger candidate for auxiliary status, *solere* – ‘to be wont to’ – which is a restructuring verb in modern Italian and Spanish.

- (380) [*C'* *Cum autem* [*IP* [*I'* **coeperit** [*vP* *episcopus* [*v'* [*VP*
 When begin-*PRF-3SG-SBJV* bishop-*NOM* come-*INF*
uenire cum ymnis...]]]]]] (25.2)
 with hymns-*ABL*

‘When the bishops starts approaching with hymns...’

- (381) ... *in qua spelunca solebat Dominus docere*
 in which cave-*ABL* be-wont-*IPFV-SG* lord-*NOM* teach-*INF*
discipulos... (33.2)
 disciples-*ACC*

‘in that cave where the Lord used to teach the disciples...’

5.8.1.2 The higher subject position

The higher subject position manifests itself in various strings which cannot be generated by placing the subject in Spec-*vP*. This is the case in most *SVX* strings where the verb is followed by material which is neither extraposed nor parenthetical. Such an example was given above in (358b), repeated here for convenience as (382); (383) is another instance. The same in all likelihood applies to (384) as well, since *posse* – ‘to be able to – is a strong candidate for auxiliary status, as argued by Danckaert (2017), meaning the following infinitive is the head of the VP. Also, the high subject is in evidence in any string where elements intervening between the subject and the verb cannot be VP-internal, but rather must themselves belong to a position above the VP (385) :

- (382) ... *cum [sanctus Moyses] acciperet a Domino*
 ... when holy-*NOM* Moses-*NOM* receive-*IPFV-SUBJ-3SG* from lord-*ABL*
legem ad filios Israhel.
 law-*ACC* to sons-*ACC* Israel.

‘... when holy Moses received the law from the Lord for the children of Israel.’

- (383) *Domine Iesu, tu promiseras nobis, ne [aliquis*
 Lord-*VOC* Jesus you promise-*PLPRF-2SG* us-*DAT* that-not someone-*NOM*
hostium] ingrederetur ciuitatem istam... (19.9)
 enemies-*GEN* enter-*IPFV-3SG-SBJV* city-*ACC* this

‘Lord Jesus, you had promised us that no enemy would enter this city...’

- (384) ... *ut [nullus] non possit commoueri (47.2)*
 so-that nobody-*NOM* NEG be-able-*3SG-SBJV* move-*INF-PASS*

‘... so that nobody can fail to be touched...’

- (385) *nam dicent, eo quod [filii Israhel] [in honore ipsorum]*
 for say-*3PL* this that sons-*NOM* Israel in honour-*ABL* same-*GEN*
[eas] posuerint.
 them-*ACC* place-*PRF-3PL-SBJV*

‘for they say that the sons of Israel had placed them there in honour of them’ (i.e. *Moses and Aron*)

It is clear that there is a subject position above the I⁰-node. One example was found where it is clear that this position cannot be Spec-IP, since material intervenes between the subject and a verb which is unambiguously in I⁰ (386).⁴⁹ Example (387) seems like another instance, since the adverb or particle *iam* – ‘now’ – intervenes between the subject and the verb, but since this is another complement clause embedded under a viaduct verb, it might well be that the subject is topicalised – in fact that seems very likely here.⁵⁰

- (386) *ita tamen ut [pars eius maxima] [sedendo in asellis]*
 such that part-NOM it-GEN greatest-NOM sitting in saddles-ABL
possit subiri; (11.4)
 be.able-3SG-SBJV ascend-INF-PASS
 ‘so that the greater part of it (i.e. *the mountain*) could be ascended while sitting in the saddle...’

- (387) *Sed mihi credite, domine venerabiles, quia [columna ipsa]*
 but me-DAT believe-2PL-IMP ladies-VOC venerable, that column-NOM same
[iam] non paret (12.7)
 now NEG appear-3SG
 ‘But believe me, venerable ladies, (when I say) that the column itself is not visible now.’

It seems a bit drastic to draw conclusions based on a single example, so we will not argue here for an even higher subject position; as we shall see shortly (section 5.8.2), there is also another way to interpret examples like (386) which makes it possible to consider that the initial position is a non-argument position altogether.

In any case, it is not crucial exactly where the subject sits in these examples, the crucial point is that there is indeed such a position available. What does seem clear, however, is that there are further positions available above the high subject. We already have established that this is the case for the adverb *semper* – ‘always’ – but there are also other instances (388)–(389); in (389) there are even two constituents in front of the verb.

- (388) *id est ut [die dominica de pullo primo] legat*
 that is that day-ABL sunday from cock-ABL first read-3SG-SBJV
episcopus intra Anastase locum resurrectionis Domini
 bishop-NOM inside Anastasis-ABL passage-ACC resurrection-GEN lord-GEN
de euangelio (27.2)
 from gospel-ABL

⁴⁹In main clauses, there are many examples where there is material between a subject which must be at least as high as Spec-IP and the following verb; however, since it is impossible to decide if the subject itself has been topicalised to the left periphery in these strings, one cannot forcefully conclude that there is a clause-internal subject position higher than Spec-IP:

- (i) ... *[ex ea die] [hi fontes] [usque in hodie] permanent hic gratia Dei.*
 from that day-ABL these sources-NOM until in today remain-3PL here thank-ABL god-GEN
 (19.2)

‘from that day and until today, these wells remain here thanks to God.’

⁵⁰In general, there are some cases of complement clauses under viaduct verbs where topicalisation seems quite likely; (387) was one example, (385) another. If these truly feature high complementisers in Force⁰, this means that the Fin⁰-node must be available, and yet the verb seemingly resides in the head-final I⁰-node. This again suggests that the verb is not attracted to a higher position.

‘that is, that on Sunday from the first cockcrow, the bishop reads in the Anastasis the passage on the resurrection of the Lord from the gospel.’

- (389) *quando [de eo loco] [primitus] uidetur mons Dei*
 when from that place-ABL at-first see-3SG-PASS mountain-NOM god-GEN
 (1.2)

‘when from that place the mountain of God is seen for the first time.’

These examples clearly reveal an area between the Fin⁰ and the I⁰ nodes. In addition to the position of the high subject, there are at least two positions, as indicated by (389). At this point I should like to discuss some evidence suggesting that this area in fact includes landing places for clause-internal A’ movement, and that there are some factors which might indicate that the subject itself – at least sometimes – undergoes movement of this kind.

5.8.2 A scrambling/operator area above IP?

Recall from our discussion of the placement of *semper* that there seemed to exist two different positions available to this adverb; one below IP and another above. The cornerstone of the Principle of Transitivity in the IP-area is the immobility of the adverbs, which are assumed to be able to move only to operator positions in the peripheries of the clause (Cinque 1999). However, if we consider for a moment the possibility that *semper* is in fact movable clause-internally, we must assume that the higher position is the derived, post-movement position. It was already mentioned above that there is no obviously discernable semantic difference between the two positions, but one might perhaps argue that the higher position is associated with a slightly stronger *emphasis*. Observe again the ‘high attachment’ of example (370), repeated here with some more context as (390); as indicated by the translation, we might surmise some stronger emphasis on the adverb: ‘so that it should always be the case that...’:

- (390) *quae predicationes propterea semper dominicis diebus fiunt, ut*
 which sermons-NOM therefore always sunday days-ABL happen-3PL so-that
 [*semper*] *erudiatur populus in scripturis et in*
 always instruct-SBJV-3SG-PASS people-NOM in scriptures-ABL and in
Dei dilectione (25.1)
 god-GEN delight-ABL

‘these sermons are always held on Sundays in order that the people may always learn about the Scripture and the love of God’

This interpretation might seem somewhat speculative, but notice that adverbs and adverbial expressions are not the only categories we find in this area of the clause. In (391) the direct object *haec omnia* – ‘all of these things’ – is unambiguously moved to a position above a head-initial I⁰ node. This clearly smacks of topicalisation, yet the phrase appears below a complementiser in a temporal adverbial clause, arguably the most resilient domain of all towards any kind of embedded root phenomena (see chapter 2, section (2.3.5 – last line.). (392) is presumably another case, where the predicate complement of the subject *melior* – ‘better’ – is fronted to a position above the subject. In fact, it cannot be proven that the subject *ascensus* is in fact in the higher subject position, since a Spec-vP parse is also available here, but in either case the subject predicative is moved above the subject, making it reasonable to assume that we dealing with the same phenomenon as in (391). On

the other hand, it should be noted that (392) could also be interpreted as either a parenthetical main clause, or alternatively as a ‘because’-clause permitting ‘high attachment’, in which case the left-periphery might be available, opening for other parses.⁵¹

- (391) *Posteaquam ergo [haec omnia] retulit sanctus episcopus,*
 after-that these-things-ACC all relate-PRF-3SG holy bishop-NOM,
ait ad me: (19.16)
 say-3SG to me:

‘after the bishop had told me all these things, he said to...’

- (392) *Et quoniam nobis ita erat iter, ut prius montem Dei*
 And since us-DAT such was road-NOM that first mountain-ACC god-GEN
ascenderemus [...] quia [unde ueniebamus], [melior]
 ascend-IPFV-1PL-SBJV since where-from come-IPFV-1PL better-NOM
[ascensus] erat... (2.3)
 ascent-NOM was

‘And since our road was such, that we first had to climb the mountain of God, because from the side we were coming, the ascent was easier...’

The relevance of all this becomes clearer when we consider more closely the nature of the subjects which tend to be attracted to the higher position above I⁰; in fact, a surprising amount of them involve emphatic readings and are quantified:

- (393) *... ut [quamuis durissimus] possit moueri in*
 that even hardest-NOM be.able-3SG-SBJV move-INF-PASS in
lacrimis... (24.10)
 tears-ABL

‘so that even the toughest can be moved into tears...’

- (394) *ut [hora inquoante septima] [omnes] [in ecclesia]*
 so-that hour-ABL begin-PRS-PTCP-ABL seventh all-NOM in church-ABL
[parati] sint... (30.3)
 ready-NOM be-3PL-SBJV

‘so that everyone should be ready in church at the beginning of the seventh hour...’

⁵¹It could of course also be argued that *melior* is not a predicative complement at all, but just an attributive adjectival modifier of the subject: ‘there was better ascent’. While this is certainly possible, I find it less plausible. Another potential example is (i), where it seems like the predicative complement of the subject *omnibus altior* is fronted, but this is much more controversial, since the postverbal subject *ille medianus* might well be in extraposition, which in turn means that the IP is not necessarily head-initial:

- (i) *Illud sane satis admirabile est [...] ut cum [omnibus altior] sit ille*
 This truly very admirable is that while all-ABL higher-NOM be-3SG-SBJV this-NOM
medianus, qui specialis Syna dicitur, [...] tamen uideri non
 middle-one which -NOM specially Sinai say-3SG-PASS yet see-INF-PASS NEG
possit...
 be.able-3SG-SBJV

‘This was truly very impressive, that although the middle one, which in particular is called the Sinai, is higher than all the other, it still cannot be seen...’

Be that as it may, it is interesting to notice that the initial constituent is (doubly) quantified, parallel to the first constituent in (391).

- (395) *Ac sic est ut [in hisdem locis] [omnes fideles] sequantur*
 but such is that in these places-*ABL* all faithful-*NOM* follow-*3PL-SBJV*
Scripturas [...] quia [omnes] docentur per illos
 scriptures-*ACC* since everyone-*NOM* teach-*PASS-3PL* through these
*dies quadraginta*46.3)
 days-*ACC* forty
 ‘For it is so that in these regions, all may follow the Scriptures [...] because every-
 body is instructed throughout these forty days...’⁵²
- (396) *Vere enim ita misteria omnia absoluet, ut [nullus] non*
 truly thus mysteries-*ACC* all unravel-*3SG* that nobody-*NOM* NEG
possit commoueri... (47.2)
 be.able-*3SG-SBJV* move-*INF-PASS*
 ‘for he (i.e. *the bishop*) truly unravels all the mysteries in such a way, that nobody
 can fail to be moved...’
- (397) *statim post prandium ascenditur mons Oliueti, id*
 immediately after lunch-*ACC* ascend-*3SG-PASS* mountain-*NOM* of-Olives that
est in Eleona, unusquisque quomodo potest, ita ut [nullus christianus]
 is in Eleona, each-one how be.able-*3SG* such that no christian-*NOM*
remaneat in ciuitate... (43.4)
 remain-*3SG-SBJV* in city-*ABL*
 ‘right after lunch the people climb the mountain of Olives, that is the Eleona, each
 as he can, so that no Christian remains in the city...’
- (398) *Et sic unusquisque festinat reuerti in domum suam,*
 and such all-and-one-*NOM* hurry-*3SG* return-*INF* in home-*ACC* REFL.ADJ
ut manducet, quia [statim ut manducauerint], [omnes]
 so-that eat-*3SG-SBJV* because immediately when eat-*PRF-3PL-SBJV* all-*NOM*
uadent in Eleona... (35.2)
 go-*3PL* in Eleona
 ‘And thus every person hurries to return to his home to eat, for as soon as they have
 eaten, everyone goes to Eleona...’
- (399) *Illud etiam presbyter sanctus dixit nobis, eo quod [usque in*
 This also priest-*NOM* holy say-*PRF-3SG* us-*DAT* this that until in
hodierna die] [semper cata pascha], [quicumque
 of-today day-*ABL* always under Easter-*ABL* whoever-*NOM*
essent baptizandi in ipso uico] [...] [omnes]
 be-*IPFV-3PL-SBJV* baptise-*GDV-NOM-PL* in same village-*ABL* all-*NOM*
[in ipso fonte] baptizarentur (15-5)
 in same source-*ABL* baptize-*IPFV-3PL-SBJV-PASS*
 ‘And this the holy priest also told us, that every Easter until this very day, whoever
 were to be baptized in the village, all of them were baptized in that same spring.’

⁵²The complement clause in (395) may of course also involve a high complementiser in Force⁰, in which case there are other parses available; one could for instance suggest that the PP *in hisdem locis* is scene-setter (provided this projection is below ForceP, as in Old French) and that the subject *omnes fideles* has been topicalised.

- (400) *ecce etiam thiamataria inferuntur intro spelunca Anastasis, ut*
 behold also censers-NOM carry.in-3PL-PASS into cave-ABL Anastasis so-that
[tota basilica Anastasis] **repleatur** *odoribus.* (24.10)
 all basilica-NOM Anastasis fill-3SG-SBJV-PASS odours-ABL
 ‘for behold! Censers are brought into the cave of the Anastasis so that the whole of
 the basilica is filled with odours’.

What these examples serve to illustrate is that the higher subject position is not infrequently associated with some emphasis, and in particular quantified subjects tend to gravitate towards this higher position. Admittedly, this does not apply to all cases (cf. (382) or (385), but there does seem to be a tendency. For instance, there is not a single case of an unambiguously low *omnes* subject, while there are 7 unambiguously high cases and 4 which were underdetermined. Given that there are two thirds more low subjects in general, this is clearly relevant and suggests that there is a slight A’ or operator flavour to the higher subject position that is lacking from the lower one.

It might also be the case that we are not dealing with only two positions, but rather three. That is, it could well be that there is a subject position in for instance Spec-IP, which is not associated with any particular emphasis (cf. 382 or 385), and then an even higher position which is some kind of operator projection. The fact that we occasionally find non-subjects there, as in (391) – notice the quantifier – or possibly in (392), which is inherently quantified, provides some support to this hypothesis. Furthermore, this is theoretically more consistent than to suggest that there is a single position which is ‘associated’ with certain functions, a notion which is not so easy to formalise.

While there is not much clear evidence available that would allow us to decide the matter, there are some indications. It is worth recalling that the one instance where the subject was clearly higher than Spec-IP was in fact quantified (cf. 386).⁵³ Furthermore, there is a revealing case of hyperbaton where the subject is discontinuous, with the genitival modifier *fidelium* – ‘of the faithful’ – remaining in a position that cannot be lower than Spec-IP, and the head, interestingly enough an indefinite quantifier, is fronted to a higher position. In between there is an adverbial (*stans* is indeclinable and cannot really be considered a conjunct participle in *Egeria*) which signals the discontinuity:

- (401) *...mittet diaconus uocem et commonet, ut [unusquisque] [stans]*
 send-3SG deacon-NOM voice-ACC and remind-3SG that each-oneNOM standing
[fidelium] **inclinent** *capita sua...* (24.6)
 faithful-GEN.PL bow-3PL-SBJV heads-ACC REFL.ADJ
 ‘... the deacon raises the voice and reminds everyone to bow their heads...’

It seems preferable to assume that the emphatic or quantified subjects in fact move to an even higher position in the clause. Whether this is a case of scrambling or in fact a full

⁵³There is in fact another example where the subject is even higher, with no less than three (although *sibi* might be a clitic) constituents intervening between the subject and the head-initial I⁰ node (i). However, *quis* is a reduced form of *aliquis* – ‘someone’ – which tends to appear in this form adjacent to the subjunction *si* ever since Classical Latin, if not earlier. It might be that this is some kind of attraction or cliticization to the C/Fin⁰-node itself:

- (i) *si [quis] [subito] [iuxta] [sibi] uult facere domum...* (14.2)
 if someone-NOM suddenly nearby REFL-DAT wish-3SG make-INF house-ACC
 ‘... if someone all of a sudden wants to make a house for himself nearby...’

operator position, in which case it might be that even adverbs like *semper* could move here, is a topic that I will leave for the future. However, the consequence of this must be spelt out clearly: if what was termed ‘the higher subject position’ in fact includes both an unmarked A-position in Spec-IP and an even higher A’ position, the dominance of the unmarked, lower Spec-vP position is even more pronounced than what was suggested in table 5.11.⁵⁴ Given these observations, it seems altogether natural to assume that the unmarked position of the subject in the grammar of *Egeria* is in fact quite low. This means that there is no strong evidence for V-to-C movement in the text, since the inversion structures are accounted for by construing the verb in I⁰. Given the SSAP, this is the simpler and accordingly the preferred parse of the data in *Egeria*.

5.9 Summary

The analysis of embedded clauses has clear impacts on the hypothesis of V-to-C movement in general. We have seen ample evidence to support the claim that V-to-C is not generalized in the text. The question has therefore been to what extent there is any V-to-C movement at all in the grammar of *Egeria*. This complicated question could not be resolved in a satisfying way by only considering data from main clauses, since these did not allow us to establish the position of the subject with certainty. Much like in chapter 3 on Old French, the data from main clauses were undetermined with respect to a V-to-I or V-to-C parse.

However, unlike what was the case for Old French, the data from embedded clauses has shifted the balance in favour of the V-to-I parse; or rather, the embedded data has not provided any evidence that calls out for a V-to-C parse, since the preceding discussion of adverbial and complement clauses has clearly revealed a low position of the subject, which in turn means that a V-to-C parse is redundant. V-to-C is quite simply not needed to account for any word order facts. Children would seemingly gain nothing by pushing the verb from I⁰ to Fin⁰. All relevant facts are accounted for by a V-in-I⁰ parse with the subject in Spec-vP, including the widespread existence of inversion, the unmarked VSO order in wide focus clauses, and even occasional (although tentative) examples of G-inversion in embedded clauses.

The asymmetries which exist between main and embedded clauses are therefore not the result of the verb moving to different projections, but rather arise from the diachronically conservative nature of embedded clauses, which still display a robust level of head-finality in the IP, an option which seems to be strongly on the decline in main clauses.⁵⁵ This

⁵⁴Another interesting example is provided in (i). In this sentence, two temporal adverbial clauses with the exact same verb and verbal arguments are conjoined. The first features inversion and a low subject, but in the second, the subject ‘the Lord’ occupies at higher position :

- (i) *Quodam tempore, posteaquam scripserat Aggarus rex ad Dominum et*
 Some time-ABL after-that write-PLPRF-3SG Abgar king-NOM to lord-ACC and
[Dominus] rescripserat Aggaro... (19.8)
 lord-NOM rewrite-PLPRF-3SG Agbar-DAT
 ‘At some time, after King Agbar had written to the Lord and the Lord had written back to Abgar...’

It is not possible to say if the subject of the second clause is in Spec-IP or a higher position, but it is highly plausible to assume that there is a contrastive topic reading at hand, which might suggest the second alternative.

⁵⁵As already mentioned, the differences in the linear distribution of the verb is to a high degree completely banal, since average clause length (in terms of the number of constituents) is a major impacting factor, cf. the verb-early, verb-late paradox in table 5.5).

analysis correctly explains the relatively lower frequencies of inversion in embedded clauses compared to main clauses, whereas in fact the opposite result would be expected under an asymmetric V-to-C vs. V-to-I analysis, since the left periphery of the main clauses will often attract the subject to a preverbal position under topicalisation and (more rarely) focalisation. Furthermore, the V-to-I parse and the unstable head parameter of the IP provided a very natural explanation for the many cases of multiple preverbal constituents in embedded clauses.

I therefore only partially agree with former analyses of *Egeria*. I agree with Clackson and Horrocks (2007) and with Ledgeway (2017) that the text displays clear evidence for a syntax which must be described as innovative, and which seems to indicate an evolution towards a verb-initial grammar. Since this grammar (like presumably any verb-initial grammar) necessarily displays high levels of inversion, not restricted to specific predicates, it seems altogether natural to hypothesise a diachronic link between this grammar and the later, generalised inversion systems of medieval Romance. On the other hand, I must again emphasize that the syntax of *Egeria* and that of the Old Romance languages (to the extent that it makes sense to express such sweeping generalizations about this family) are still quite different; at least it is very different from the Old French texts that we examined in chapters 3 and 4. In particular, *Egeriae* does not feature a generalized, across-the-board V-to-C movement, a fact which is particularly clear in the (not as infrequent as sometimes claimed) cases where the prefield hosts a great variety of constituents which, apart from being too numerous, often correspond rather poorly with proposed left-peripheral roadmaps. Secondly, on the derivational assumptions adopted here, the finite verb demonstrably still resides in a head-final projection in many cases.⁵⁶

The more intriguing question is to what extent the text shows incipient signs of V-to-C movement. This matter is complex, but if we assume that children are conservative structure builders, there does at least not seem to be any *unambiguous* evidence for verb movement to such a high projection. The reason for this is, as always, closely connected with the global input, and in particular the lack of strong evidence for a consistently high subject position. This is the second major difference between *Egeria* and the Old Romance languages. The most natural parse of the data seems to involve a kind of ‘Celtic’⁵⁷ V-in-I setup with a subject that oscillates between a low, VP-internal subject position (as evidenced by several cases where IP adverbs intervene between the verb and the subject) and a higher, preverbal position. Presumably, the nature of this higher subject position and how it was analysed by children will have been crucial to the long-term evolution of the language. As long as this subject position is either felt to be secondary or is analysed as an A’-position related to specific information structural effects, the VSO nature of the language might be expected to be stable or even to solidify (once head-finality goes extinct).

In a VSO-language, the subject will still regularly precede the verb in discourse due to topicalisation or focalisation. We have ample and unambiguous evidence that topicalisation of the subject to the left periphery was common from an early age in Latin as in most languages (Danckaert 2012). However, it seems clear from *Egeria* that this is not the only way a subject might attain a preverbal position, since it is highly unlikely that the quite frequent sequences *complementiser-subject-verb-X* in embedded clauses all involve topical-

⁵⁶Perhaps the latter point is actually good news for anyone who wants to make sense of the Latin-Romance diachrony, since we presumably need head-finality for quite some time still, at least until this pattern has grammaticalised to give the new synchronic future and conditional tenses of Romance plus complex predicates like *certifier/certificare*, a process which seems far from complete in *Egeria*’s day.

⁵⁷I prefer to characterize the VSO-grammar of *Egeria* as ‘Celtic’ instead of the alternative ‘Semitic’ for obvious reasons; cf. the discussion of the Christian Latin ‘Sondersprache’ hypothesis in section 5.2.4.

isation to the left periphery. The question is therefore if this clause-internal position is another kind of A-bar position, such as the high scrambling projection identified by Danckaert (2017:25,203); see also Devine and Stephens (2006:28). The evidence from adverbial and complement clauses shows that many subjects do indeed have slightly different and more emphatic readings than the lower subjects, and for some reason, quantified subjects tend to prefer a higher position. At the same time, there are clear examples of completely normal, unmarked subjects which also appear at least as high as Spec-IP.

If we add to this an observation made in the discussion of main clauses, namely that subjects occasionally precede the subject in what may plausibly be analysed as wide focus SVX clauses, we may conclude that there seems to be evidence for saying that there is also a preverbal argument position for the subject as well. This makes it possible to discern a possible link between the grammar of *Egeria* and the Old Romance languages. Once this higher position is experienced as the default, unmarked position of the subject, we might expect that this will trigger two immediate changes in a VSO system of this kind: the emergence of an unmarked SVO order and the emergence of a grammar with V-to-C movement. In other words, there is a latent, but clear potential for V2 syntax in the grammar of *Egeria*.

It is therefore fitting to round off by returning to an inversion example from the main clauses; in (402) there is inversion with a quantified *omnes*-subject. We recall from our discussion in section 5.8.2 that these quite consistently target a high position in embedded clauses. If we can assume that the same applies to (402) as well, this suggests that the verb has in fact moved to a very high position. This analysis receives support from the fact that the inverted subject also exceptionally outscopes a sentential adverb, *similiter*, much against the grain of inversion structures in general, which tend to be of the non-contiguous kind (cf. section 5.5.1). This must be considered a plausible candidate for V-to-C movement, then.⁵⁸

- (402) *[Item hora sexta] [denuo] **descendent** omnes similiter ad*
 Likewise hour-ABL sixth again descend-3PL all-NOM similarly to
Anastasim. . . (24.3)
 Anastasis-ACC
 ‘Also at the sixth hour, everyone again goes down in similar fashion to the Anastasis. . .’

In chapter 6, I will discuss a possible diachronic evolution from the late Latin syntax exemplified by *Egeria* and towards the Old Romance languages in general and towards Old French in particular.

⁵⁸There are several other cases in main clauses where the verb precedes an *omnes*-subject, but these are unfortunately all clause-final, either modified by relative clauses and/or very plausibly in narrow focus.

Chapter 6

Bridges in time

6.1 Introduction

In the preceding chapters, a fairly detailed investigation was carried out into the syntax and word order patterns of Old French and Late Latin. In this final chapter, we will take the opportunity to look back and briefly summarize the most important conclusions reached during this investigation on both an empirical and theoretical level. This will be undertaken in section 6.2.

At the same time, an attempt will be made to go one step further by setting these conclusions into a wider context and by exploring their potential implications for the understanding of the general Latin-Romance diachrony as well as the particular diachronic path that leads to Old French. Concretely, we will try to understand what must have taken place in the long intervening period that separates Late Latin and the Old French prose texts of the 13th century. Needless to say, this discussion will necessarily be much more tentative, and partially outright speculative, in nature, since firm evidence for this period is quite simply lacking. Nonetheless, the discussion is important since it might serve to sharpen the analytical focus by narrowing down the set of possible scenarios, and crucially, to promote hypotheses that are sufficiently concrete to be corroborated or weakened by future research. In this respect, the conclusions reached in the previous chapters may serve as bounding conditions; we have the beginning and the end, now we must build the bridge that connects them. This is the topic of the final section 6.3.

6.2 Old French and Late Latin

In chapters 3 and 4, two French prose texts from the early thirteenth century were analysed. Numerous conclusions were reached on the basis of this investigation. Generally speaking, Old French was a staunchly head-initial SVO language with relatively rigid word order. In this respect, Old French already resembled modern French in many ways. At the same time, Old French syntax featured widespread inversion, and these inversion patterns could only be dealt with in a satisfactory way by adopting a V-to-C analysis. This was the only parse which was consistent with the global input in an economic and coherent way, since alternative solutions that have been proposed in the literature suffer from various drawbacks, most notably the failure to account for the asymmetries between main and embedded clauses.

These asymmetries are undeniable, but also highly predictable, and mirror the situation in the modern Germanic languages, in particular the Mainland Scandinavian languages, with astonishing precision, the sole difference being that Old French featured independent V-to-I movement in embedded clauses. The evidence showed that embedded V-to-C movement was available in Old French in complement clauses under viaduct verbs and in certain adverbial clauses like consecutive clauses. The texts did not provide the slightest evidence for a symmetric inversion system.

Furthermore, Old French had developed very strong constraints on the prefield; both by prohibiting it from being left radically empty, a property that was formalised in conventional fashion by adopting an EPP-feature on the Fin^0 -head, as well as by restricting the number of constituents in the prefield to exactly one. Due to these two latter properties, Old French, presumably as the only Old Romance language, should be characterised as a V2 language on any reasonable definition of that term. In particular, inversion in Old French was a completely syntactic phenomenon, triggered automatically by the fronting of any non-subject constituent to the prefield. Like in the modern Germanic V2 languages, information-structure played a crucial role in deciding what constituent should go to the prefield, but absolutely no role in triggering inversion per se, which was a grammaticalised property of the language, internalized during the acquisition process.

As all V2 languages, Old French permitted exceptions to the linear V2 rule. The important point is that these exceptions are generally just as predictable as embedded V2, in that they involve specific constructions or specific lexical items. Most of these were similar to those found in the modern Germanic V2 languages, much more so than generally admitted in the literature. As already mentioned, Old French permitted left-dislocated nominal phrases to precede the V2 construction, provided these were co-indexed with a resumptive pronoun inside the clause. The evidence did not permit us to conclude with certainty if these LDs included both Hanging Topics and CLDs. Other exceptions from linear V2 revolve around a small group of adverbial and interjection-like expressions.

The most salient difference with respect to Germanic, however, is found in the behaviour of initial subordinate clauses, which almost invariably fail to trigger inversion. The V2 rule in Old French is ‘blind’ to the initial clause, which is left-dislocated rather than integrated into the prefield. In this respect, Old French is a slightly weaker V2 language than the Germanic languages. Moreover, the *Vie de Saint Eustace* showed some signs of instability in occasionally allowing various adverbial expression to precede the V2 construction, yielding V3 orders which are very rare in *Tristan*. It was suggested that these strings arise because the initial adverbs fail to trigger verb movement, but that this pattern should not be considered a part of the Old French V2 system, but rather an early sign of its decline.

It was also demonstrated that the restricted nature of the prefield in Old French cannot be captured theoretically by positing verb movement to a high left peripheral position like Force^0 , as suggested by Wolfe (2015). This analysis incorrectly predicts that initial scene-setters should be exempt from the V2 constraint, since it should be possible to merge scene-setters directly in the highest layer of the left periphery above ForceP once the V2 constraint has been satisfied by movement. This was not generally the case in Old French, since scene-setters trigger inversion just as much as any other constituent. To the extent that they sometimes fail to trigger inversion, it was argued, as already mentioned, that this betrays the first signs of erosion of the V2 grammar. Also, some of the productive V3 patterns in Old French seem to feature CLDs, and these left-dislocation constructions, along with initial subordinate clauses, are equally available in embedded clauses, showing that they belong below ForceP in the left periphery, not above it, as the Force-V2 analysis claims.

Furthermore, the Force-V2 analysis also predicts the existence of embedded ‘relaxed V2’ sequences under viaduct verbs, another prediction which is not borne out. In all of these aspects, the situation in Old French corresponds exactly to the situation in all standard varieties of Modern Germanic, a fact that suggests that the V-to-Force⁰ analysis may not work for V2 languages in general. What distinguishes Old French from the other Old Romance languages, or in more general terms, what distinguishes V2 systems from V-to-C movement systems without restrictions on the prefield, is therefore not to be sought in a distinction between Fin⁰ and Force⁰ as the locus of verb movement.

This does not come as a surprise in a non-nativist approach that assumes children to be conservative structure builders, since there quite simply is no evidence for a hierarchy of projections in the left-periphery in a V2 language. It would seem like the phonological and information-structural cues do not lead children to expand structure, but rather to syncretize features into a single position, thereby creating a single multi-functional projection. In this respect, the traditional V-to-C analysis actually seems to fare better than articulated left-peripheries in accounting for V2 systems, although it is arguably possible to have the best of both worlds by adopting the syncretic approach advocated by Hsu (2017). This was also the approach adopted in this thesis.

6.2.1 Late Latin

More than 800 years separate the Old French texts that were examined from the Latin witness, the *Itinerarium Egeriae*. It should therefore come as no surprise that the two stages of the language display very different properties. In general terms, the Latin itinerary gives witness to a historical stage that was still characterised by considerable word order freedom and where constructions like the AcI, the Ablative Absolute and discontinuous structures (hyperbata) are still quite frequent. In this respect, the language of the text is still much closer to Classical Latin than to Romance. At the same time, the language shows signs of an evolution towards the Romance situation. In main clauses there is a strong tendency towards head-initial structures. In adverbial and complement clauses, on the other hand, it was argued that the verb-final complementation pattern is still quite prevalent, interpreted as a sign of diachronic conservatism.

Interestingly, the text evinces a host of inversion strings, particularly in main clauses. In a novel analysis of the syntax of *Egeria*, Ledgeway (2017) interpreted this as evidence for a VSO-grammar, derived asymmetrically by V-to-C (more specifically, V-to-Fin⁰) movement in main clauses and V-to-I in embedded clauses. On Ledgeway’s definition, which is shared by many Romanists, this already makes the language of *Egeria* a V2 grammar. In this thesis, a sharp distinction is drawn between V-to-C and V2, and the text was therefore not considered a candidate for V2 status.

However, apart from this definitional question, the empirical evidence for V-to-C movement in *Egeria* was considered to be rather weak. The problem is in a sense related to the first claim made by Ledgeway, namely that Late Latin was a VSO language, a claim which is corroborated by our own analysis. In a VSO-language, the unmarked word order is already an inversion structure (as defined in this thesis), and the child acquiring the language will have to decide exactly how much to expand the clausal structure to accommodate the initial position of the verb. This, in turn, is crucially dependent on the position of the subject. Given certain theoretical assumptions regarding the organization of the clause, there are at least two different kinds of evidence that might lead the child to postulate V-to-C movement. First, G-inversion structures, where the subject intervenes between the finite auxiliary and

non-finite main verbs, in other word strings of the kind: *Auxiliary-Subject-Main Verb*. Second, the subject could outscope sentential adverbs in the IP-area, which would yield strings of the type: *Verb-Subject-Adv1-Adv2...-VP*. None of these strings were found, and the rather sparse evidence from adverbial positions rather suggested a low position for the subject.

A comparison between main and embedded clauses revealed some notable asymmetries in that the latter contained more V1 strings, as was also pointed out by Ledgeway. This finding could *a priori* be interpreted in favour of the asymmetric analysis proposed by Ledgeway. However, the embedded clauses also featured considerably less inversion than main clauses, and this is in fact completely the opposite of what is expected under the asymmetric analysis, since the possibilities for the subject to move to the left of the verb should be very limited in embedded clauses. It was argued that it does not seem plausible to derive all non-V1 strings in embedded clauses by embedded V-to-C movement, since this would entail that Latin is completely insensitive to the conditions on embedded root-phenomena proposed by Hooper and Thompson (1973), conditions which have been shown to be quite consistent cross-linguistically and which were also shown to hold in Old French. Therefore, it was argued that the asymmetries with respect to inversion arise as a result of the stronger propensity for the IP to be head-final in embedded clauses. In sum, the evidence points to a low subject position in Spec-vP and a grammar with a V-in-I setup, with little evidence for V-to-C. At the same time, it is clearly possible for the subject to precede the verb in I⁰, and the availability of this position in embedded clauses suggests that it cannot just be a topic position or any other A-bar position in the left periphery. *Egeria* therefore seems to oscillate between both VSO, SVO and SOV patterns.

6.3 From Latin to Romance... and Old French

In this section, I will address the complicated issue of the diachronic evolution that leads from Late Latin to Old Romance. The task is to construct a bridge that might plausibly lead from the grammar of *Egeria* and to the inversion systems of the Old Romance languages. In so doing, we have no choice but to go beyond the evidence itself and speculate. This speculation does not amount to pure conjecture, however, since we have some rather clear premises provided by the results obtained in the previous chapters. Naturally, these premises are not self-evidently correct, and we must admit that there is more uncertainty connected to the Late Latin situation than the Old French one, but provided that we accept these premises, we have sufficiently clear boundary conditions to be able to rule out some scenarios as less likely than others.

It might be objected at this point that we are not truly in the dark with respect to the entire period between 400 and 1250, and that what we really need to connect by speculation is not the grammar of *Egeria* and the Old French prose texts of the 13th century, but rather the former and the first Old French documents of the 11th and 12th centuries. This argument, though valid, is not necessarily true. The point is that the 10th, 11th and greater part of the 12th centuries (almost) only provide us with texts in verse, and these cannot be uncritically trusted as evidence of the situation in the spoken language. While this evidence should clearly not be disregarded, it cannot be granted more than ancillary status, at best. I will return to this point.

The historical evolution from Latin to Old French is just one side branch of a more general evolution from Latin to Romance. As is well-known by now, this particular branch Old French shows some distinctive features that set it apart from the other Old Romance

languages. It has been suggested that Old French belonged to the group of ‘strict V2’ languages (Benincà 1983), and that it was presumably the strictest of them all (although Wolfe (2015) finds a more or less equally strict pattern in Old Spanish.) On the definition of V2 used in this dissertation, the special status of Old French is emphasized even more clearly, since it is assumed that Old French was the only Romance V2 language in the medieval period. We therefore need to raise the question why this should be the case.

Let us start by summarizing concretely what the two ends look like, so as to get a clearer picture of what must be reconstructed. In table 6.1 is a summary of the principal differences between Old French and Late Latin as they were analysed in the preceding chapters. If we consider the column for Old French, we can say that the first four properties have been retained and still form the backbone of modern French syntax, while the two latter properties are specific to Old French and together constitute the V2 rule of the language.

Table 6.1: Some differences between Late Latin and Old French

Property	Late Latin	Old French
General word order:	Relatively free	Relatively fixed
Unmarked word order:	VSO/SVO	SVO
Subject position	Spec-vP/Spec-IP	Spec-IP
Head parameter:	Unstable/Head-initial	Head-initial
Inversion:	Generalized V-to-I	Generalized V-to-C
Prefield:	No restrictions	Highly restricted

Some diachronic analyses of the evolution from (Classical) Latin to modern Romance have focused particularly on the shift in unmarked word from a (presumed) SOV to a SVO pattern. This is the focus of several contributions of Bauer, who analyses this as a gradual resetting of the head parameter from a head-final to a head-initial pattern, thereby uniting property two and three in the table above (Bauer 1995, 2009). This analysis makes a lot of sense when comparing the situation in Classical Latin and Modern French, but it does not quite capture the intermediate stage represented by Old French. To stay in the metaphor of bridge-building, this account misses something because it starts building the bridge at the wrong place; if we start building from the SOV order of Classical Latin, and use the head parameter as our primary analytic tool, we span the bridge too high, over Old French, and never arrive there.

It seems plausible that the secret behind the Old Romance situation in general is connected to the VSO grammar of Late Latin. This is the core of the argument of Ledgeway (2017). However, Ledgeway attempts to connect Late Latin and Old Romance very directly by arguing that the two stages already featured much of the same syntax. In particular, Late Latin had already developed consistently head-final structures plus a generalized pattern of V-to-C movement in main clauses. In chapter 5, it was argued that the evidence for V-to-C movement in *Egeria* was not as strong as suggested by Ledgeway, but let us shift the focus here and rather ask how the resulting bridge would look, if this analysis were taken to be correct.

It seems like this analysis suffers from the opposite problem of the one proposed by Bauer,

namely the fact that it has lost the SVO order along the way. In a staunchly verb-initial stage with the subject in Spec-vP, there is no longer a preverbal argument position for the subject in either main or embedded clauses. If this is the case, then the Old Romance languages must simply have reinvented the SVO order in some period between the fourth century and the emergence of the first Romance sources. We might ask at this point how this process could have come about if the language had already developed V-to-C movement. There is no preverbal argument position in the left periphery. We would have to hypothesize that the frequent topicalisations of the subject were at some point reanalysed as a left-peripheral argument position, triggering a change to unmarked SVO order that would subsequently trickle down into the embedded clauses (although the structural position would in fact not be the same, since it would be Spec-CP in main clauses and Spec-IP in embedded clauses). This is not impossible, but it seems to me to get the things in the wrong order. For one thing, it was argued in the chapters on Old French that the subject position is not in Spec-CP, but rather in Spec-IP. Although evidence for this admittedly is tentative, it might be interpreted as evidence that the subject position was ‘invented’ in a VSO grammar with the verb in I⁰.

It therefore makes sense to assume that the development of the higher subject position predates the development of V-to-C movement, and that the former process fuels the second. This is not only more consistent with the empirical facts of Late Latin, as they were interpreted in chapter 5, but also provides for a better diachronic bridge into the Old Romance situation. We already have all the tools we need to create this bridge. Crucially, we do not have to conjure up a preverbal argument position in the undocumented period, as it is manifestly already there in the Late Latin data. Not all preverbal subjects in Late Latin are topics, this much seems clear from the analysis of *Egeria*. Furthermore, not all preverbal subjects are in Spec-vP, either. There is unambiguous evidence for a high subject position which is simply a less frequent alternative to the lower position in Spec-vP. Looking back at table 6.1, we can therefore hypothesize that the relevant reanalysis, V-to-C movement, is the result of the combination of two independent processes which are historically undisputable, namely the development of a staunchly head-initial verbal projection (both at the VP and IP levels) and the emergence of an unmarked SVO order. Notice, however, that the emergence of this unmarked word order does not come about exclusively through the resetting of the head parameter, since we cannot ignore the fact that Late Latin bears unmistakable signs of a VSO phase. If we trust the source we have been reviewing, we are left with no choice but to take this as the point of departure. And in this scenario, if there had not existed a structurally high, preverbal subject position, but only a consistently low, Spec-vP subject, then the gradual decline of head-final orders would in fact have had the effect of consolidating the VSO order, presumably turning Old Romance into a stable Celtic phase.

Concretely, then, the bridge from Late Latin into Romance consists of a first stage, *which is the only one on record* and which is characterised by considerable word order freedom, unstable head-parameters, widespread inversion, and a competition between a high and a low subject position. The analysis of this stage was given in chapter 5 and can be summarized as a VSO stage with the verb in I⁰. The subsequent evolution of the language tends towards gradually more head-initial structures and less word order freedom, but these properties at first do not lead to any significant reanalysis. However, the higher subject position must have come to assert itself over time, until it is perceived as the default subject position by children acquiring the language. This constitutes the second stage, which will have been characterised by a gradual decrease of inversion strings in embedded clauses as they are

replaced by strict SVO order. Finally, as this stage matures, the old V-to-I analysis will no longer capture the inversion strings in main clauses as economically as before, and a V-to-C analysis is invented to account for the exact same inversion strings, a reanalysis which constitutes the third stage and leads to the situation in Old Romance. Schematically, then:

Figure 6.1: From Latin to Romance in three stages

VSO with inversion	VSO with inversion	SVO with inversion
Unstable Head Parameter	Head-initial IP/VP	Head initial IP/VP
Subject in Spec-vP	Subject in Spec-vP/IP	Subject in Spec-IP
Symmetric V-to-I	Symmetric V-to-I	Asymmetric V-to-C

Although it is convenient to talk about stages, and this also permits us to present the evolution in schematized form, it should be emphasized that it is realistic to assume considerable overlap between these stages. Figure 6.1 is therefore meant to express the logic of the evolution more than its temporal dimension. The competition between head-initial and head-final in the IP, and the competition between Spec-vP and Spec-IP as subject positions must have run in parallel. Since both of these processes fuel the process of reanalysis towards V-to-C movement,¹ it is reasonable to assume that even the rise of V-to-C movement will have been gradual. Although it might seem theoretically cleaner to operate with abrupt and discrete reanalysis which subsequently spreads in a population, there is good reason to believe that this is not actually the way languages evolve. Change is fundamentally gradual, as old and new patterns co-exist as different options in individual grammars. As for V-to-C movement, I take it to be trivial that if a language L at a given time T_1 does not feature V-to-C movement at all, and at a historically subsequent time T_3 features generalized V-to-C movement, then there will have existed a transitional period of time T_2 where V-to-C movement found place occasionally, but not always.

The changes just discussed lead to Old Romance in a broad sense of the word. However, they do not lead to Old French. If we consider table 6.1, the properties on the left (Late Latin) have been replaced by the properties on the right. There is one exception, however, namely the final property of the table, the make-up of the prefield. As we recall, inversion strings in Late Latin involved strings with one, two, three or even more constituents to the left of the verb. We have no reason to assume that the rise of a V-to-C grammar will in and of itself lead to a more restricted prefield. The bridge into Old French therefore consists of a final, fourth stage which is lacking in the other Romance languages, namely the evolution into a V2 language.

¹In principle, one might imagine that the drift towards a head-initial IP and a high subject in Spec-IP would simply feed the emergence of SVO order and the disappearance of inversion. If one just plots these two changes onto the syntactic tree suggested for Late Latin, that is indeed the result; and if the Old Romance languages had been like their modern descendants, this would not have been an implausible story. However, the Old Romance situation suggests that this was not the outcome of the process; rather, the surface inversion strings persisted throughout the complex vicissitudes of the other changes and were eventually reinterpreted as V-to-C.

6.3.1 The particular status of Old French

Since Old French has this particular status within the Old Romance family, we must now address the inevitable question of why and how this situation arose. But first, let us just briefly recapitulate exactly in what sense Old French is different, since it is not sufficiently clear to just state that it was a V2 language. What sets Old French apart from the other Romance languages of the period is the fact that the former had developed very strong restrictions on the prefield, generally accepting only one constituent in front of the verb. In chapter it was argued that the prefield in Old French – and this may indeed extend to V2 languages in general – was multi-functional, consisting of a single, syncretised position able to host constituents with a variety of different information-structural features. What has changed is therefore the prefield itself.

Accounting for this final stage is not easy, but two different explanations come to mind. The first is a well-known, external explanation, namely the *Germanic influence* on the French language. The second is an internal explanation, and is therefore related to other grammatical properties that to some extent seem to set Old French apart, namely the highly rigid word order of the language.² These two factors may of course also be interrelated.

The hypothesis that the syntax and word order of Old French are influenced by the Frankish superstrate is time-honoured (Meillet 1931:37, von Wartburg 1958:128, Harris 1984:193, Thomason and Kaufman 1988:53, Posner 1996:53, Vincent 2000:62, Mutz 2009:61). It is also not in any sense implausible from a historical perspective, since the Germanic presence in Gaul was considerable and prolonged. Even so, the consensus is that the Franks were never more than a dominant minority even in the areas where their relative share of the population was most elevated. Based on a study of the density of Germanic place names, von Wartburg concluded that the areas most affected by Frankish settlement lay north of the Seine and (in particular) Somme rivers (von Wartburg 1939:104-110), a conclusion which finds support in archeological evidence (Petri 1973:123), although the latter is reportedly very hard to interpret (for a discussion, see James 1988:109-117). As for concrete numbers, these vary greatly both in relative and absolute terms; according to Lodge, von Wartburg's estimate of a total between 15% and 26% of the population contrasts with low estimates of around 3% (Lodge 2001:62). These figures are anyway not much more than guesswork.

It is simply futile to spend more time on this discussion. However, there is one thing that we may claim with a reasonable degree of certainty, namely that whatever the syntactic changes in French brought about through Germanic influence, *these changes must have occurred relatively early*. Although Germanic loanwords may have sifted into Old French throughout the Middle Ages, there is a time-window for more fundamental, structural impact, which presumably does not extend much beyond the Merovingian period.³ During this period, we may assume that bilingualism was at least not negligible in Northern Gaul.

²The rigid character of Old French word order is of course to be understood in structural, not in linear terms. The apparent freedom of the initial position is a direct result of the V2 constraint, itself a structural constant. Verb-second (and Stylistic Fronting) aside, the word order freedom of the old language with respect to Modern French is mostly a question of some highly limited scrambling or scrambling-like movement in the IP-VP area, in most cases the apparently optional short movement of the direct object to a position above the VP, cf. section 3.5.1.

³This is not to say that the Germanic impact on Gallo-Roman Latin will have been strongest in the very earliest years of Frankish settlement. On the contrary, it seems likely that the initial stage was characterised by some level of segregation between the ethnic groups, and that contact and networks of exchange needed some time to mature. Bilingualism will therefore inevitably have had a rising trend for some time as well. Inter-marriage will of course have been a powerful motor of bilingualism, but not much is known about the extent of ethnic inter-marriage in post-Roman Gaul (see Mathisen 1993:134-136 for a short discussion).

There are some indications that the seventh century constitutes a transitional period in this respect; James reports that the word *Frank* in the sixth century still refers to the ethnic group of the German-speaking community, while it takes on the meaning of *inhabitant from Northern Gaul* by the early eighth century (James 1982:32). Furthermore, the custom of separate burials for Franks and Gallo-Romans ceases in the seventh century (Lodge 2001:64). If we interpret this as meaning that the Germanic tongue ceases to be spoken in France in the early Carolingian period, excepting the royal family, the court and perhaps its most immediate entourage, there is *prima facie* an insurmountable mismatch between this scenario and the hypothesis that V2 was introduced into Old French by the Franks. If the V2 system of the early 13th century that we reviewed in chapters 3 and 4 is a relatively recent innovation of the preceding century or less, then we may simply rule out Germanic influence as a direct causal trigger.

This brings us back to the point already touched upon above, namely the reliability of the evidence before the 13th century and in particular before the 12th century. This period, often referred to as Early Old French, provides us with a significant amount of texts in verse, but virtually no evidence which could be described as prose. It is conspicuous that the emergence of prose texts in the late 12th/early 13th century coincides so neatly with several proposed changes in the evolution of the French language, such as rigidification of the V2 pattern in main clauses, the aversion against true V1 clauses (Skårup 1975:291), the evolution from a symmetrical V2 system to an unsymmetrical one (Hirschbühler and Junker 1988; Côté 1995), the first occurrences of embedded pronominal inversion (Zaring 2017), the rise of obligatory preverbal expletives, a change from word-based to phrase-based stress assignment (Marchello-Nizia 1995), changes in the Tobler-Mussafia Law (Labelle and Hirschbühler 2005; Zimmermann and Kaiser 2010), etc. This profound caesura in the evolution of the language raises some suspicion and indicates that there might exist a problem of *submerged Old French* as well, as it does not seem inconceivable that the situation in the 13th century is not so much one of whole-scale innovation or system change as one of continuity. Although the French language will naturally have changed between the 8th and early 13th centuries, we might raise the question how strong the evidence really is for saying that the V2 system of the later Old French period is of recent making rather than a centuries-old retention.

Admittedly, this is nothing more than guesswork, so we will not pursue the matter further. Let us rather turn to a more tangible question and say something about how plausible this Germanic contact explanation of V2 is, *if taken to be historically real*.

6.3.2 Extension and reanalysis

Harris and Campbell argue that language contact should not be understood as a mechanism of change in itself, but rather as a situation which facilitates language change through other mechanisms such as *reanalysis*, *extension* or *borrowing* (Harris and Campbell 1995:50-52). The latter term is presumably the one which is most intuitively associated with the notion of language contact or contact-induced change; when two languages are in contact, language A can borrow a grammatical property from language B. But while borrowing is a major factor in lexical transfer between languages, as was also the case between Old Franconian and Gallo-Roman, the idea that a word order constraint like V2 is borrowed seems a bit simplistic. It is therefore a clear advantage for the account proposed here that we do not have to rely on borrowing to construct it, but rather on extension followed by reanalysis.

Concretely, if the Old Franconian language of the Franks was indeed a V2 language, in itself by no means a trivial assumption, the bilingualism of the early Merovingian period will

have involved two languages with rather similar word order properties. The late Latin/early Romance tongue of the native Gallo-Romans will already have featured widespread inversion. It does not seem implausible that the differences between the prefields were sufficiently subtle to create interference effects in the acquisition process of the second and subsequent generations, particularly since the grammatical constraint that existed in the Frankish language did not violate anything in the Latin language. The growing obsolescence of head-finality will also have removed the verb-late strings from the Latin language, further reducing the differences in word order. Linear V2, with and without inversion, will have been a prominent word order even in the Latin language. This scenario is consistent with the claim that cross-linguistic transfer in bilingual acquisition is triggered by overlap in the input structures (Hulk and Müller 2000:229).

Another claim made in the literature on bilingual first language acquisition is that small children in bilingual communities might spend some time figuring out that they are in fact acquiring two different languages (Volterra and Taeschner 1978; Genesee 1989 for an overview). We might hypothesise that the Frankish and the Latin language will have constituted a single global input to the infant child, who set out to assign a grammar to this input in accordance with the *String-Structure-Assignment-Hypothesis*. In this single-input or ‘monolingual’ stage, the V2 rule is in fact the lowest common denominator, the only solution consistent with the global input. This is of course on its own an unrealistically simple explanation, but observe that even as the child grew conscious of the difference between the two languages, it would still have been possible to acquire a correct use of the prefield in both languages by generalizing the V2 constraint, whereas it would of course not be possible to go in the opposite direction and generalize the unrestricted nature of the Latin prefield. And finally, if this generalization did not happen with the first bilingual generation(s), even more moderate interference effects in the actual use of the prefield, that is to say an increased tendency to prefer linear V2, will in turn have reduced the difference between the two prefields, with concomitant effects for the acquisition process of the following generation. This corresponds to the notion of *incrementation* in the model of language change advocated by Labov (2007), namely that the change is brought about stagewise in a collective enterprise of several generations, as each generation goes one step further than the previous one.

We do not need to rely exclusively on the acquisition process as the sole locus of language change, an idea which is regarded as too simplistic by many today (Aitchinson 2001; Sankoff 2004); see also Stanford (2015) for discussion. It is even more simplistic to place all of the explanatory burden of the transfer of the V2 constraint on bilingual first language acquisition. While this scenario will surely not have been uncommon, it is perhaps more likely that the canonical bilingualism of Gaul will have involved Gallo-Romans who acquired Old Franconian as a second language, and vice versa. In such L2 acquisition, it is useful to follow Van Coetsem and distinguish between two types of contact-induced change or *transfer*, namely *borrowing* and *imposition*. *Borrowing* takes place when the agents of change are dominant in the recipient language. In our case, this would be when bilingual Gallo-Romans took from their Franconian L2 the habit of using the native prefield in a more Germanic way. This does not mean that they had to borrow any grammatical constraint, it could have been just a slight influence in actual usage, an interference effect. *Imposition* is when the agents of change are dominant in the source language. In our scenario, this would be the case when Franks transferred their Germanic L1 use of the prefield onto their Gallo-Roman L2 tongue. This seems like a very plausible scenario indeed. Not only did the Franks on our assumptions have a V2 rule in their native language, we also know that

imposition of native patterns in general is well-documented (Lucas 2015). We even have highly relevant examples pertaining to the acquisition of the prefield; recall from chapter 2 the studies of Bohnacker (2010) and Bohnacker and Rosén (2008) on L2 acquisition of the prefield in German and Swedish. Both studies document clear imposition effects: the Swedes imposed their *Rheme Later* preference onto German, generally avoiding new information in the prefield, while the Germans acquiring Swedish tended to ‘overuse’ the prefield according to their native habits. The grammatical constraint (V2) was respected, but the actual use of the prefield is highly subtle, and therefore very prone to interference effects.

We can imagine that such effects will have been widespread during the period of bilingualism in Gaul, and that Franks in particular drove the restriction of the prefield forward by imposition of linear V2 order. This would have led to gradual frequency changes in the language of the adult population. This process is part of the *extension* mechanism of Harris and Campbell (1995), where an already existing pattern (linear V2) is generalized at the expense of competitors. The spreading of the change, the *diffusion* in the sense of Labov’s model (Labov 2007), is of course driven by the adult population rather than the children.

This scenario has the major advantage of providing a principled account for what might otherwise appear like ‘uncaused drifts in usage frequencies that occur prior to and independently of grammar change’, in the words of Kroch (2005:2). Also, it is not dependent on borrowing or radical reanalysis of highly discrete properties. If the Gallo-Roman idiom of the Merovingian period had been like Modern French, a staunch SVO language with extremely limited inversion possibilities, then the V2 property of the invading Franks could not have penetrated so easily into this system, since there is little chance that children could mix up such saliently different properties during the bilingual acquisition process. In other words, reanalysis would presumably not take place, and extension would have nothing to operate on, since it is not possible to extend a non-existing pattern (inversion under transitives). The properties of the prefield, on the other hand, are much more insidious, and lend themselves easily to extension, thereby paving the way for future reanalysis.

As this process proceeds, the evidence for distinct projections in the prefield is gradually weakened, and sooner or later – but certainly not as late as the 12th century⁴ – this incremental process will have aligned the two prefields to the extent that grammatical constraints are formulated during acquisition; an EPP-feature on the Fin⁰ head in response to the (virtual) absence of V1 orders, a syncretised, multi-functional projection in response to the absence of $V \geq 3$ orders. This final step is of course an instance of *reanalysis*, and we might assume that this change can only take place during transmission/acquisition. At some point, grandfather’s use of *Topic-focus-verb*-sequences started sounding awfully old-fashioned to the young in Northern France, although the same structures were very much alive elsewhere in România.

It seems like this is exactly the right kind of scenario for the change we need to create the final section of our bridge. Naturally, more research is clearly needed on this topic, and the hypothesis should be evaluated against a more articulate framework of linguistic change, itself embedded in a more detailed model of the particular sociolinguistic context of bilingual Gaul. This is a topic for future research; but I believe the hypothesis expressed here is a highly concrete, and I believe, reasonably plausible explanation for the exceptionality of Old French and for the rise of the V2 system. Somewhat paradoxically, then, if there was

⁴It is perhaps possible to maintain the idea that the process of change described in this section, namely the gradual shrinking of the prefield brought about by Germanic influence, subsisted as a vector on the evolution of the language even after the period of bilingualism had waned, and that it only led to reanalysis at a much later date. This would be a nice example of Sapirian *drift*, a latent force in diachrony.

any Germanic contribution to the rise of the V2 system in Old French, it will not have been *Germanic inversion*, which was an internal development, but rather the restrictions on the prefield.

Bibliography

- Abels, Klaus (2003). “Successive cyclicity, anti-locality, and adposition stranding”. PhD thesis. University of Connecticut.
- (2012). “The Italian Left Periphery: A view from locality”. In: *Linguistic Inquiry* 43.2, pp. 229–254.
- (2018). “Beyond phrase-phrase interactions in the left periphery”. Talk presented at the Workshop on Word Order in the Left Periphery. University of Oslo, October 17-18.
- Abels, Klaus and Ad Neeleman (2012). “Linear Asymmetries and the LCA”. In: *Syntax* 15.1.
- Aboh, Enoch O. (2004). *The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa*. New York: Oxford University Press.
- (2010). “Information Structuring Begins with the Numeration”. In: *Iberia: An International Journal of Theoretical Linguistics* 2.1, pp. 12–42.
- Adams, James N (1976a). “A Typological Approach to Latin Word Order”. In: *Indogermanische Forschungen* 81. Reprinted, pp. 70–99.
- (1976b). *The Text and Language of a Vulgar Latin Chronicle (Anonymus Valesianus II)*. Bulletin Supplement No. 36. University of London: Institute of Classical Studies.
- (1977). *The Vulgar Latin of the Letters of Claudius Terentianus*. Manchester: Manchester University Press.
- (2007). *The Regional Diversification of Latin*. New York: Cambridge University Press.
- (2011). “Late Latin”. In: *A Companion to the Latin Language*. Ed. by James Clackson. Chichester: Wiley-Blackwell, pp. 257–283.
- (2013). *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Adams, James N and Nigel Vincent, eds. (2016). *Early and Late Latin: Continuity or Change?* With the assistance of Valerie Knight (ed.) Cambridge: Cambridge University Press.
- Adams, Marianne (1987a). “From Old French to the Theory of Pro-Drop”. In: *Natural Language and Linguistic Theory* 5, pp. 1–32.
- (1987b). “Old French, Null Subjects, and Verb Second Phenomena”. PhD thesis. UCLA.
- (1989). “Verb second effects in Medieval French”. In: *Studies in Romance Linguistics: Selected Papers from the Seventeenth Linguistic Symposium on Romance Languages*. Ed. by Karl Kirschner and Janet Ann DeCesaris. Current Issues in Linguistic Theory 60. Amsterdam: John Benjamins.
- Adger, David (2003). *Core Syntax: a Minimalist Approach*. Oxford: Oxford University Press.
- Aitchinson, Jean (2001). *Language change: progress or decay?* 3rd ed. Cambridge: Cambridge University Press.

- Alexiadou, Artemis (2006). “Left-dislocation (including CLLD)”. In: *The Blackwell Companion to Syntax*. Ed. by Martin Everaert and Henk van Riemsdijk. Oxford: Blackwell, pp. 668–699.
- Alexiadou, Artemis and Elena Anagnostopoulou (1998). “Parametrizing AGR: Word Order, V-movement and EPP checking”. In: *Natural Language and Linguistic Theory* 16 (3), pp. 491–539.
- Ambar, Manuela (1992). *Para uma Sintaxe da Inversão Sujeito-Verbo em Português*. Lisbon: Colibri.
- Anderson, James M. and Bernard Rochet (1979). *Historical Romance Morphology*. Ed. by Ann Arbor. Published for the University of Calgary by University Microfilms International.
- Anderson, Stephen R. (1993). “Wackernagel’s Revenge: Clitics, Morphology, and the syntax of Second Position”. In: *Language* 69, pp. 68–98.
- (2000). “Towards an optimal account of second position phenomena”. In: *Optimality Theory: phonology, syntax and acquisition*. Ed. by Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer. Oxford: Oxford University Press, pp. 302–333.
- (2004). “Subject Clitics and Verb-second in Surmiran Rumantsch”. In: *MIT Working Papers in Linguistics* 47, pp. 1–22.
- Andersson, Lars-Gunnar (1975). “Form and function of subordinate clauses”. PhD thesis. University of Gothenburg.
- Andrade, Aroldo de (2018). “Aboutness Topics in Old and Middle French: A corpus-based study on the fate of V2”. In: *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 63.2, pp. 194–220.
- Angantýsson, Ásgrímur (2008). “Sentential Negation and verb placement in embedded clauses in Icelandic”. Talk presented at the NORMS Workshop on Negation at the University of Oslo, March 11-12.
- (2011). “The Syntax of Embedded Clauses in Icelandic and Related Languages”. PhD thesis. University of Iceland.
- Antomo, Mailin and Markus Steinbach (2010). “Desintegration und Interpretation: Weil-V2-Sätze and Schnittstelle zwischen Syntax, Semantik und Pragmatik”. In: *Zeitschrift für Sprachwissenschaft* 29, pp. 1–37.
- Antonelli, André (2015). “Roots of V-to-C Movement in Romance: Investigating Late Latin Grammar”. Presented at the conference ‘Traces of History’ at the University of Oslo, March 9-10 2015.
- Arnbjörnsdóttir, B., H. Thráinsson, and I.E. Nowenstein (to appear). “V2 and V3 orders in North American Icelandic”. In: *Journal of Language Contact*.
- Axel, Katrin (2007). *Studies on Old High German Syntax*. Linguistik Aktuell/Linguistics Today 112. Amsterdam: John Benjamins.
- Baker, Mark (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- (2001). “The Natures of Nonconfigurality”. In: *The Handbook of Contemporary Syntactic Theory*. Ed. by Mark Baltin and Chris Collins. Malden, MA: Blackwell, pp. 407–438.
- Banniard, Michel (1992). *Viva Voce: Communication écrite et communication orale du IVE au IXe siècle en occident latin*. Paris: Institut des études augustiniennes.
- Barbance-Guillot, Céline, Serge Heiden, and Alexei Lavrentiev (2017). “Base de français médiéval : une base de référence de sources médiévales ouverte et libre au service la communauté scientifique”. In: *Diachroniques* 7, pp. 168–184.

- Barbosa, Pilar (1995). "Null Subjects". PhD thesis. Cambridge, MA: MIT.
- Bauer, Brigitte (1995). *The Emergence and Development of SVO Patterning in Latin and French: Diachronic and Psycholinguistic Perspectives*. New York: OUP.
- (2009). "Word Order". In: *New Perspectives on Historical Latin Syntax. The Syntax of the Sentence*. Ed. by Philip Baldi and Pierluigi Cuzzolin. Vol. 1. Trends in Linguistics. Berlin: Mouton de Gruyter, pp. 241–316.
- Bayer, Josef (1984). "COMP in Bavarian Syntax". In: *Linguistic Inquiry* 3, pp. 209–274.
- (2001). "Asymmetry in emphatic topicalisation". In: *Audiatur vox sapientiae*. Ed. by Caroline Féry and Wolfgang Sternefeld. *Studia grammatica* (52). Berlin: Akademie Verlag, pp. 15–47.
- (2008). "What is Verb Second?" Manuscript, University of Konstanz.
- Becker, Martin (2005). "Le corpus d'Amsterdam face à une vieille question: l'ancien français est-il une langue V2?" In: *Romanistische Korpuslinguistik II. Korpora und diachrone Sprachwissenschaft*. Ed. by Johannes Kabatek Claus D. Pusch and Wolfgang Raible. Tübingen: Gunter Narr Verlag, pp. 345–358.
- Behagel, Otto (1909). "Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern". In: *Indogermanische Forschungen*. *Indogermanische Forschungen* 25, pp. 110–142.
- Belletti, Adriana (2004a). "Aspects of the Low IP Area". In: *The Structure of CP and IP: The Cartography of Syntactic Structures*. Ed. by Luigi Rizzi. Vol. 2. Oxford and New York: Oxford University Press, pp. 16–51.
- (2004b). "Introduction". In: *Structures and Beyond: The Cartography of Syntactic Structures*. Ed. by Adriana Belletti. Vol. 3. Oxford and New York: Oxford University Press, pp. 1–14.
- (2009). *Structures and Strategies*. New York and Abingdon: Routledge.
- Benincà, Paola (1983). "Un' ipotesi sulla sintassi delle lingue Romanze antiche". In: *Quadri Patavini di Linguistica* 4, pp. 3–19.
- (1995). "Complement Clitics in Medieval Romance: The Tobler-Mussafia Law". In: *Clause Structure and Language Change*. Ed. by Ian Roberts and Adrian Battye. Oxford: Oxford University Press, pp. 325–344.
- (2004). "The Left Periphery of Medieval Romance". In: *Studi Linguistici e Filologici Online* 2, pp. 243–297.
- (2006). "A Detailed Map of the Left Periphery of Old Romance". In: *Crosslinguistic Research in Syntax and Semantics*. Ed. by Raffaella Zanuttini. Georgetown: Georgetown University Press, pp. 53–86.
- (2013). "Caratteristiche del V2 romanzo, lingue romanze antiche, ladino dolomitico e portoghese". In: *Introduzione alla linguistica del mocheno*. Ed. by Ermenegildo Bidese and Federica Cognola. Turin: Rosenberg & Sellier, pp. 65–84.
- Benincà, Paola and Cecilia Poletto (2004). "Topic, Focus, and V2: Defining the CP Sublayers". In: *The Structure of CP and IP*. Ed. by Luigi Rizzi. Oxford Studies in Comparative Syntax. Oxford and New York: OUP.
- Bentzen, Kristine (2007). "Order and Structure in Embedded Clauses in Northern Norwegian". PhD thesis. University of Tromsø.
- (2009). "Embedded root phenomena, assertion, presupposition, and main point of utterance". Talk given at the Workshop on Root Phenomena, ZAS Berlin.
- (2014). "Verb placement in clauses with initial adverbial maybe". In: *Nordic Atlas of Language Structures (NALS)* 1, pp. 225–239.

- Bergh, Lars (1952). “Quelques réflexions sur l’inversion après la conjonction et en ancien et moyen français”. In: *Mélanges de philologie romane offerts à M. Karl Michaëlsson par ses amis et ses élèves*. Göteborg: Bergendahl, pp. 43–55.
- Berman, Judith (2003). *The Clausal Syntax of German*. Stanford: CSLI Publications.
- den Besten, Hans (1983). “On the interaction of root transformations and lexical deletive rules”. In: *On the Formal Syntax of the Westgermania*. Ed. by Werner Abraham. Linguistik Aktuell 3. Amsterdam: John Benjamins, pp. 47–132.
- den Besten, Hans and Corretje Moed-van Walraven (1986). “The syntax of verbs in Yiddish”. In: *Verb second phenomena in Germanic languages*. Ed. by Hubert Haider and Martin Prinzhorn. Dordrecht: Foris, pp. 111–136.
- Bhatt, R.M. (1999). *Verb Movement and the Syntax of Kashmiri*. Dordrecht: Kluwer.
- Biberauer, Theresa (2002). “Verb second in Afrikaans: is this a unitary phenomenon?” In: *Stellenbosch Papers in Linguistics* 34, pp. 19–69.
- Biberauer, Theresa, Anders Holmberg, and Ian Roberts (2014). “A Syntactic Universal and its Consequences”. In: *Linguistic Inquiry* 45, pp. 169–225.
- Boeckx, Cedric (2014). “What Principles and Parameters got wrong”. In: *Linguistic Variation in the Minimalist Framework*. Ed. by Carme Picallo. Oxford: Oxford University Press, pp. 155–179.
- Boeckx, Cedric and Kleantes G. Grohmann (2005). “Left dislocation in Germanic”. In: *Focus on Germanic Typology*. Ed. by Werner Abraham. Berlin: Akademie Verlag, pp. 131–144.
- Bohnacker, Ute (2010). “The clause-initial position in L2 Swedish declaratives: word order variation and discourse pragmatics”. In: *The Nordic Languages and Second Language Acquisition Theory*. Ed. by Ute Bohnacker and Marit Westergaard. Special Issue of Nordic Journal of Linguistics 33 (2), pp. 105–143.
- Bohnacker, Ute and Christina Rosén (2008). “The clause-initial position in L2 German declaratives: transfer of information structure”. In: *Studies in Second Language Acquisition* 30.4, pp. 511–538.
- Borsley, Robert D, Maria-Luisa Rivero, and Janig Stephens (1996). “Long head movement in Breton”. In: *The Syntax of the Celtic Languages*. Ed. by Robert D Borsley and Ian Roberts. Cambridge: Cambridge University Press, pp. 53–74.
- Bošković, Željko (2018). “On Movement out of Moved Elements, Labels, and Phases”. In: *Linguistic Inquiry* 49.2, pp. 247–282.
- Brandtner, Ellen (2004). “Head-movement in minimalism, and V2 as FORCE-marking”. In: *The Syntax and Semantics of the Left Periphery*. Berlin and New York: Mouton de Gruyter, pp. 97–139.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler (2015). *Lexical-Functional Syntax*. Second edition. Wiley-Blackwell.
- Burguy, Georges F. (1869a). *Grammaire de la langue d’oïl, ou grammaire des dialectes français aux XIIe et XIII siècles*. 2nd ed. Vol. 2. Berlin: W. Weber.
- (1869b). *Grammaire de la langue d’oïl, ou grammaire des dialectes français aux XIIe et XXIII siècles*. 2nd ed. Vol. 1. Berlin: W. Weber.
- Buridant, Claude (2000). *Grammaire nouvelle de l’ancien français*. Paris: Sedes.
- Büring, Daniel and Katharina Hartmann (2001). “The syntax and semantics of focus-sensitive particles in German”. In: *Natural Language and Linguistic Theory* 19.2, pp. 229–281.
- Burton, Philip (2011). “Christian Latin”. In: *A Companion to the Latin Language*. Ed. by James Clackson. Chichester: Wiley-Blackwell, pp. 485–501.

- Cabrillana, Concepción (1999). “Type of Text, Pragmatic Function, and Constituent Order: A Comparative Study of the *Mulomedicina Chironis* and the *Peregrinatio Aetheriae*”. In: *Latin vulgaire - latin tardif V. Actes du V^e colloque international sur le latin vulgaire et tardif, Heidelberg 5-8 septembre 1997*. Ed. by Hubert Petersmann and Rudolf Ketteman. Heidelberg: Winter.
- Cardinaletti, Anna and Ian Roberts (2002). “Clause Structure and X-Second”. In: *Functional Structure in DP and IP*. Ed. by Guglielmo Cinque. The Cartography of Syntactic Structures. Oxford: Oxford University Press, pp. 123–166.
- Cerquiglini, Bernard (1981). *La parole médiévale : discours, syntaxe, texte*. Paris: Les Editions de minuit.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Studies in Generative Grammar 9. Dordrecht: Foris.
- Chomsky, Noam (1970). “Remarks on Nominalization”. In: *Readings in English Transformational Grammar*. Ed. by Roderick A. Jacobs and Peter S. Rosenbaum. Waltham: Ginn, pp. 184–221.
- (1986). *Barriers*. Cambridge, MA: MIT Press.
- (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- (2000). “Minimalist Inquiries”. In: *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Ed. by Roger Martin, Michaels David, and Juan Uriagereka. Cambridge, MA: MIT Press, pp. 89–155.
- (2001). “Derivation by phase”. In: *Ken Hale: A life in language*. Ed. by Michael Kenstowicz. Cambridge, MA: MIT Press, pp. 1–52.
- (2012). “Foreword”. In: *Phases : Developing the Framework*. Ed. by Ángel J. Gallego. Studies in Generative Grammar 109. Berlin and Boston: De Gruyter Mouton, pp. 1–7.
- Chomsky, Noam and Howard Lasnik (1977). “Filters and Control”. In: *Linguistic Inquiry* 8.3, pp. 425–504.
- Cichosz, Anna (2017). “Inversion after clause-initial adverbs in Old English: The Special Status of θ_a , θ_{onne} , nu , and swa ”. In: *Journal of English Linguistics* 45.4, pp. 308–337.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford Studies in Comparative Syntax. Oxford: OUP.
- Cinque, Guglielmo and Luigi Rizzi (2009). “The Cartography of Syntactic Structures”. In: *The Oxford handbook of linguistic analysis*. Ed. by Bernd Heine and Heiki Harrog. Oxford: OUP, pp. 51–65.
- Clackson, James and Geoffrey Horrocks (2007). *The Blackwell History of the Latin Language*. Oxford: Blackwell.
- Clark, Herbert H. and Susan E. Haviland (1977). “Comprehension and the given-new contract”. In: *Discourse production and comprehension*. Ed. by Roy O Freedle. Norwood, New Jersey: Ablex Publishing Corporation, pp. 1–40.
- Cognola, Federica (2013). *Syntactic Variation and Verb Second: A German Dialect in Northern Italy*. Amsterdam: John Benjamins.
- Coleman, Robert (1987). “Vulgar Latin and the diversity of Christian Latin”. In: *Latin vulgaire - latin tardif*. Ed. by József Herman. Tübingen: Niemeyer, pp. 37–52.
- Côté, Marie-Hélène (1995). “Concurrence structurale, conditions d’appréhensibilité et changement syntaxique: la chute de la structure V2 en français”. In: *Revue canadienne de linguistique* 40, pp. 165–200.
- Cruschina, Silvio (2010). “Fronting as focalisation in Sicilian”. In: *Syntactic Variation. The Dialects of Italy*. Ed. by Roberta D’Alessandro, Adams Ledgeway, and Ian Roberts. Cambridge: Cambridge University Press. Chap. 247-260.

- Csipak, Eva (2018). “Adverbial clauses and V3”. Manuscript, University of Konstanz.
- Curtis, René L., ed. (1963). *Le Roman de Tristan en Prose. Tome 1*. München: Max Hueber Verlag.
- (1983). “Who wrote the Prose Tristan? A new look at an old problem”. In: *Neophilologus* 83, pp. 35–41.
- Cuzzolin, Pierluigi and Gerd Haverling (2009). “Syntax, sociolinguistics, and literary genres”. In: *New Perspectives on Historical Latin Syntax*. Vol. 1: Syntax of the Sentence. Berlin and New York: Mouton de Gruyter, pp. 19–63.
- Danckaert, Lieven (2012). *Latin Embedded Clauses: The Left Periphery*. Amsterdam: John Benjamins.
- (2017a). “Subject Placement in the History of Latin”. In: *Catalan Journal of Linguistics* 16, pp. 125–161.
- (2017b). *The Development of Latin Clause Structure: A Study of the Extended Verb Phrase*. Oxford University Press.
- Dardel, Robert de (1983). “Declaratives romanes à verbe initial”. In: *Vox Romanica* 42, pp. 1–37.
- (1989). “L’hypothèse d’une base OVS en protoroman”. In: *Probus* 1, pp. 121–143.
- (1996). *À la recherche du protoroman*. Tübingen: Niemeyer.
- Dardel, Robert de and Rinze A Haadsma (1976). “Le rejet du verbe dans les subordinées romanes”. In: *Vox Romanica* 35, pp. 24–39.
- De Cat, Cécile (2009). *French Dislocation: Interpretation, Syntax, Acquisition*. New York: Oxford University Press.
- deHaan, Germen and Fred Weerman (1986). “Finiteness and Verb Fronting in Frisian”. In: *Verb Second Phenomena in Germanic Languages*. Ed. by Hubert Haider and Martin Prinzhorn. Dordrecht: Foris Publications, pp. 77–111.
- Demske, Urike (2012). “Clause linking and information packaging”. Handout at the workshop ‘The Making of the Early Modern Period’, University of Agder in Kristiansand.
- Déprez, Viviane (1988). “Stylistic Inversion and verb movement”. In: *Proceedings of ESCOL 1988*, pp. 71–82.
- Devine, Andrew M. and Laurence D Stephens (2006). *Latin Word Order: Structured Meaning and Information*. Oxford: OUP.
- Dewey, Tonya K. (2006). “The Origins and Development of Germanic V2: Evidence from Alliterative Verse”. PhD thesis. Berkeley: University of California.
- Diderichsen, Paul (1966/1944). *Elementær Dansk Grammatik*. København: Gyldendahl.
- Diesing, Molly (1990). “Verb movement and the subject position in Yiddish”. In: *Natural Language and Linguistic Theory* 8.1, pp. 41–79.
- Diez, Friedrich (1877). *Grammatik der romanischen Sprachen. Drei Theile in einem Bande*. 4th ed. Bonn: Weber.
- Dixon, Robert (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Donaldson, Bryan (2012). “Initial subordinate clauses in Old French: Syntactic variation and the clausal left periphery”. In: *Lingua* 122, pp. 1021–1046.
- Drach, Eric (1963/1937). *Grundgedanken der deutschen Satzlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Dupuis, Fernande (1988). “Pro-Drop dans les subordinées an ancien français”. In: *Revue Québécoise de Linguistique Théorique et Appliquée* 7, pp. 41–62.
- (1989). “L’expression du sujet dans les subordinées en Ancien Français”. PhD thesis. University of Montreal.

- Eide, Kristine (2006). “Word Order Structures and Unaccusative Verbs in Classical and Modern Portuguese”. PhD thesis. University of Oslo.
- Eide, Kristine M (2011). “Norwegian (non-V2) declaratives, resumptive elements, and the Wackernagel position”. In: *Nordic Journal of Linguistics* 34.2, pp. 179–213.
- Einhorn, Elsabe C (1974). *Old French: A Concise Handbook*. Cambridge: Cambridge University Press.
- Ekerot, Lars-Johan (1988). *Så-konstruktionen i svenskan*. Lundastudier i nordisk språkvetsenskap. Lund: Lund University Press.
- Elcock, William D (1975). *The Romance Languages*. London: Faber.
- Elerick, Charles (1992). “Latin Word Order: Living on the Edge”. In: *The Classical World* 86, pp. 21–24.
- Elsig, Martin (2009). “Verb-second effects in Old French and Middle High German. Considerations from the Language Border”. MA thesis. Hamburg.
- (2012). “Subject-verb inversion in 13th century German and French”. In: *Multilingual individuals and multilingual societies*. Ed. by Kurt Braunmüller and Christoph Gabriel. Hamburg studies on multilingualism 13 13. Amsterdam: John Benjamins, pp. 223–239.
- Engdahl, Elisabeth and Enric Vallduví (1996). “Information packaging and grammar architecture: A constraint-based approach”. In: *Studies in HPSG. Edinburgh Working Papers in Cognitive Science*. Ed. by Claire Grover and Enric Vallduví. Vol. 12. Edinburgh: Centre for Cognitive Science, University of Edinburgh, pp. 1–31.
- Erdmann, Oskar (1886). *Grundzüge der deutschen Syntax nach ihrer geschichtlichen Entwicklung. Erste Abteilung*. Stuttgart: Cotta.
- Eythórsson, Thórhallur (1995). “Verbal syntax in the early Germanic languages”. PhD thesis. Cornell University.
- Faarlund, Jan Terje, Svein Lie, and Kjell Ivar Vannebo (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Fabricius-Hansen, Cathrine and Kåre Solfeld (1994). “Deutsche und norwegische Sachprosa im Vergleich: Ein Arbeitsbericht”. In: *Arbeitsberichte des Germanistischen Instituts der Universität Oslo* 6.
- Falk, Yehuda N. (2001). *Lexical-Function Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI.
- Fanselow, Gisbert (2002). “Quirky ‘subjects’ and other specifiers”. In: *More than words: a Festschrift for Dieter Wunderlich*. Ed. by Ingrid Kaufmann and Barbara Stiebels. Berlin: Akademie Verlag, pp. 227–250.
- Ferraresi, Gisella and Maria Goldbach (2002). “V2 Syntax and Topicalisation in Old French”. In: *Linguistische Berichte* 189, pp. 3–27.
- (2003). “Particles and sentence structure. A historical perspective”. In: *Syntactic Structures and Morphological Information*. Ed. by Uwe Junghanns and Luka Szucsich. Berlin: Mouton de Gruyter, pp. 101–127.
- Fiéis, Maria Alexandra (2003). “Ordem de palavras, transitividade e inacusatividade: Reflexão teórica et análise do Português dos séculos XIII a XVI”. PhD thesis. Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
- Fodor, Janet Dean (1998). “Unambiguous Triggers”. In: *Linguistic Inquiry* 29, pp. 1–36.
- Fontana, Josep M (1993). “Phrase Structure and the Syntax of Clitics in the History of Spanish”. PhD thesis. University of Pennsylvania.
- Foulet, Lucien (1923). *Petite syntaxe de l’ancien français*. 2nd ed. Paris: Librairie Ancienne Honoré Champion.

- Foulet, Lucien (1930). *Petite syntaxe de l'ancien français*. 3rd ed. Paris: Librairie Ancienne Honoré Champion.
- Franco, Irene (2017). "Stylistic fronting in Old Italian: a phase-based analysis". In: *Language: Historical Syntax* 93, pp. 114–151.
- Franzen, Torsten (1939). "Etude sur la syntax des pronoms personnels sujets en ancien français". PhD thesis. University of Uppsala.
- Frascarelli, Mara and Roland Hinterhölzl (2007). "Types of Topics in German and Italian". In: *On information structure, meaning and form: generalizations across languages*. Ed. by Kerstin Schwabe and Susanne Winkler. Amsterdam: John Benjamins, pp. 87–116.
- Frey, Werner (2004a). "Notes on the syntax and the pragmatics of German left-dislocation". In: *The syntax and semantics of the left periphery*. Ed. by Horst Lohnstein and Susanne Trissler. Berlin: Mouton de Gruyter, pp. 203–233.
- (2004b). "The grammar-pragmatics interface and the German prefield". In: *Sprache und Pragmatik* 52, pp. 1–39.
- (2006). "Contrast and movement to the German prefield". In: *The Architecture of Focus. Studies in Generative Grammar*. Ed. by Valéria Molnár and Susanne Winkler. Studies in Generative Grammar 82. Berlin and New York: Mouton de Gruyter, pp. 235–264.
- Frey, Werner and Federica Masiero (2018). "Desintegration versus Parordination bei obwohl- und weil-Konstruktionen". In: *ZAS Papers in Linguistics* 59, pp. 57–82.
- Freywald, Ulrike (2008). "Zur Syntax und Funktion von dass-Sätzen mit Verbzweitstellung". In: *Deutsche Sprache* 36.
- (2016). "Vorfeldbeschränkungen in adversativen während-Gefügen". Talk given at the 2th International Workshop on Verb-Second Order at the University of Wuppertal.
- Gallego, Ángel J. (2010). *Phase Theory*. Amsterdam; Philadelphia: John Benjamins.
- (2011). "Parameters". In: *The Oxford Handbook of Linguistic Minimalism*. Ed. by Cedric Boeckx. Oxford: Oxford University Press, pp. 523–551.
- ed. (2012). *Phases: Developing the Framework*. Berlin; Boston: De Gruyter Mouton.
- Gärtner, Hans-Martin (2001). "Are there V2 relative clauses in German?" In: *The Journal of Comparative Germanic Linguistics* 3.2, pp. 97–141.
- (2016). "A Note on the Rich Agreement Hypothesis and Varieties of Embedded V2". In: *Working Papers in Scandinavian Syntax* 96, pp. 1–13.
- Genesee, Fred (1989). "Early bilingual development: one language or two?" In: *Journal of Child Language* 16, pp. 161–179.
- Giorgi, Alessandra and Fabio Pianesi (1996). *Tense and Aspect: From semantics to morphosyntax*. New York: OUP.
- Givón, Talmy (1971). "Historical syntax and synchronic morphology: an archeologist's field trip". In: *Chicago Linguistic Society* 7, pp. 394–415.
- Green, John N. (2000). "Spanish". In: *The Romance Languages*. Ed. by Martin Harris and Nigel Vincent. 3rd ed. London and New York: Routledge. Chap. 3, pp. 79–130.
- Grewendorf, Günther (1993). *Aspekte der deutschen Syntax: eine Rektions-Bindungs-Analyse*. Tübingen: Günter Narr.
- (2008). "The clausal left periphery: Clitic left-dislocation in Italian and left-dislocation in German". In: *Dislocated elements in discourse*. Ed. by Benjamin Shaer, Philippa Cook, Werner Frey, and Claudia Maienborn. London: Routledge, pp. 49–94.
- (2009). "The left clausal periphery". In: *Dislocated elements in discourse*. Ed. by Benjamin Shaer, Philippa Cook, Werner Frey, and Claudia Maienborn. London: Routledge, pp. 49–94.

- Grewendorf, Günther and Cecilia Poletto (2011). "Hidden Verb Second. The Case of Cimbrian". In: *Studies on German Language Islands*. Ed. by Michael T. Putnam. Amsterdam: John Benjamins, pp. 301–346.
- Grohmann, K. G. (2003). *Prolific domains*. Amsterdam: John Benjamins.
- Günther, Susanne (2000). "wobei (.) es hat alles immer zwei seiten." Zur Verwendung von wobei im gesprochenen Deutsch". Manuscript, University of Konstanz. InLiSt 18.
- Haegeman, Liliane (1992). *Theory and Description in Generative Syntax. A Case Study in West Flemish*. Cambridge: Cambridge University Press.
- (1995). *The Syntax of Negation*. Cambridge: Oxford University Press.
- (1996). "Verb second, the split CP and null subjects in early Dutch dialects". In: *GenGenP* 4.2, pp. 133–175.
- (2007). "Operator movement and topicalisation in adverbial clauses". In: *Folia Linguistica* 41.3–4, pp. 279–325.
- (2010). "The internal syntax of adverbial clauses". In: *Lingua* 120, pp. 628–648.
- (2012). "Main clause phenomena: new horizons". In: *The syntax of MCP: deriving the truncation account*. Ed. by Lobke Aelbrecht, Liliane Haegeman, and Rachel Nye. Amsterdam: John Benjamins, pp. 113–134.
- Haegeman, Liliane and Ciro Greco (2016). "V>2". Paper presented at Rethinking Verb-Second, Cambridge, 22th-24th March 2016.
- Haider, Hubert (1986). "Verb Second Phenomena in Germanic Languages". In: ed. by Hubert Haider and Martin Prinzhorn. Dordrecht: Foris Publications. Chap. V-Second in German, pp. 49–75.
- (1993). *Deutsche Syntax Generativ. Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen: Günter Narr.
- (2010). *The Syntax of German*. Cambridge: Cambridge University Press.
- Hale, Ken (1983). "Warlpiri and the Grammar of Non-Configurational Languages". In: *Natural Language and Linguistic Theory* 1, pp. 5–47.
- (1989). "On Nonconfigurational Structures". In: *Configurationality: The Typology of Asymmetries*. Ed. by Marácz Laslo and Pieter Muysken. Dordrecht: Foris, pp. 293–300.
- Hänsch, Alexandra Y. (2014). "Germanic Properties in the Left Periphery of Old French: V-to-C-Movement, XP-Fronting, Stylistic Fronting and Verb-Initial Clauses". PhD thesis. University of Ottawa.
- Harris, Alice C. and Lyle Campbell (1995). *Historical syntax in cross-linguistic perspective*. Cambridge Studies in Linguistics 74. Cambridge and New York: Cambridge University Press.
- Harris, Martin (1977). "A Typological Approach to Word Order Change in French". In: *Romance Syntax: Synchronic and Diachronic Perspectives*. Ed. by Martin Harris. Salford: Salford University, pp. 33–53.
- (1978). *The Evolution of French Syntax: A Comparative Approach*. London: Longman.
- (1984). "On the Causes of Word Order Change". In: *Lingua* 63, pp. 175–204.
- (2000a). "French". In: *The Romance Languages*. Ed. by Martin Harris and Nigel Vincent. London and New York: Routledge, pp. 209–245.
- (2000b). "The Romance Languages". In: *The Romance Languages*. Ed. by Martin Harris and Nigel Vincent. 3rd ed. London and New York: Routledge, pp. 1–25.
- Haug, Dag T. T (2017). "Backward control in Ancient Greek and Latin participial conjuncts". In: *Natural Language and Linguistic Theory* 35. DOI 10.1007/s11049-016-9339-7, pp. 99–159.

- Haug, Dag T. T and Marius L Jøhndal (2008). “Creating a Parallel Treebank of the Old Indo-European Bible Translations”. In: *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (2008)*. Ed. by Caroline Sporleder and Kiril Ribarov, pp. 27–34.
- Head, Brian F. and Larissa Semënova-Head (2013). “Vestígios da Presença Sueva no Noroeste da Península Ibérica: na Etnologia, na Arqueologia e na Língua”. In: *Revista Diacrítica* 27 (2), pp. 257–277.
- Heraeus, Wilhelm, ed. (1908). *Siviae vel potius Aetheriae Peregrinatio*. Heidelberg: Carl Winter.
- Herman, J. (1963). “Sur la préhistoire du système roman des adverbes de temps”. In: *Beiträge zur Romanischen Philologie* 2.1, pp. 150–165.
- Herman, József (2000). *Vulgar Latin*. Pennsylvania: The Pennsylvania State University Press.
- Herzenberg, Marie Johanne Bordal (2015). *Third Person Reference in Late Latin. Demonstratives, Definite Articles, and Personal Pronouns in the Itinerarium Egeriae*. Tübingen: Mouton de Gruyter.
- Heycock, Caroline (2006). “Embedded root phenomena”. In: *The Blackwell Companion to Syntax*. Ed. by Martin Everaert and Henk van Riemsdijk. Oxford: Blackwell, pp. 174–209.
- Heycock, Caroline, Antonella Sorace, and Zakaris Svabo Hansen (2010). “V-to-I and V2 in subordinate clauses: an investigation of Faroese in relation to Icelandic and Danish”. In: *The Journal of Comparative Germanic Linguistics* 13.1, pp. 61–97.
- Hinojo, Gregorio (1986). “El orden de palabras en la Peregrinatio Aetheriae”. In: *Studia Zamorensia* 7, pp. 81–87.
- Hinterhölzl, Roland and Svetlana Petrova (2010). “From V1 to V2 in West Germanic”. In: *Lingua* 120, pp. 315–328.
- Hirschbühler, Paul (1990). “La légitimation de la construction V1 à sujet nul en subordonnée en prose et en vers en ancien français”. In: *Revue Québécoise de Linguistique* 19.1, pp. 33–55.
- Hirschbühler, Paul and Marie-Odile Junker (1988). “Remarques sur les sujets nuls en subordonnées en ancien et en moyen français”. In: *Revue Québécoise de Linguistique Théorique et Appliquée* 7.3, pp. 63–84.
- Hoekstra, Jarich (1993). *The Split CP hypothesis and the Frisian complementiser system*. Frisian Academy: Ms.
- Hofmann, Johannes B (1951). *Lateinische Umgangssprache*. 3rd ed. Heidelberg: Winter.
- Hofmann, Johannes B and Anton Szantyr (1965). *Lateinische Syntax und Stilistik*. München: Beck.
- Holmberg, Anders (2000). “Scandinavian stylistic fronting: how any category can become an expletive”. In: *Linguistic Inquiry* 31, pp. 445–483.
- (2006). “Stylistic Fronting”. In: *The Blackwell Companion to Syntax*. Ed. by Martin Everaert and Henk van Riemsdijk. Oxford: Blackwell, pp. 532–565.
- (2015). “Verb Second”. In: *Syntax - an International Handbook on Contemporary Syntactic Research*. Ed. by Tibor Kiss and Artemis Alexiadou. Berlin: Walter de Gruyter, pp. 342–382.
- Holmberg, Anders and Christer Platzck (1995). *Role of Inflection in Scandinavian Syntax*. Oxford: OUP.
- Hooper, Joan and Sandra Thompson (1973). “On the applicability of root transformations”. In: *Linguistic Inquiry* 4, pp. 465–497.

- Hornstein, Norbert, Jairo Nunes, and Kleantes G. Grohmann (2010). *Understanding Minimalism*. New York: CUP.
- Hrafnbjargarson, Gunnar Hrafn and Anna-Lena Wiklund (2009). “General Embedded V2: Icelandic A, B, C, etc.” In: *Working Papers in Scandinavian Syntax* 84, pp. 21–51.
- Hsu, Brian (2017). “Verb second and its deviations: An argument for feature scattering in the left periphery”. In: *Glossa: a journal of general linguistics* 2 (1).35, pp. 1–33.
- Hulk, Aafke and Ans van Kemenade (1995). “Verb Second, Prop-Drop, Functional Projections and Language Change”. In: *Clause Structure and Language Change*. Ed. by Adrian Battye and Ian Roberts. New York and Oxford: Oxford University Press, pp. 227–257.
- Hulk, Aafke and Natasha Müller (2000). “Bilingual first language acquisition at the interphase between syntax and pragmatics”. In: *Bilingualism: Language and Cognition* 3, pp. 227–244.
- Iatridou, Sabine and Anthony Kroch (1992). “The Licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon”. In: *Working Papers in Scandinavian Syntax* 50, pp. 1–25.
- Ingham, Richard (2005). “Adverbs and the Syntax of Subjects in Old French”. In: *Romania* 123, pp. 99–122.
- (2007). “À la recherche des origines de la syntaxe de ‘ja’ et de ‘onques’ antéposés en ancien français”. In: *Zeitschrift für romanische Philologie* 1.122, pp. 57–78.
- (2013). “A Derivational Approach to Negative Polarity Item Licensing in Old French”. In: *Research on Old French: The State of the Art*. Ed. by Deborah L. Arteaga. Dordrecht: Springer, pp. 261–281.
- (2014). “Old French negation, the Tobler/Mussafia Law, and V2”. In: *Lingua* 147, pp. 25–39.
- Jackendoff, Ray (1977). *X-bar-Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monographs 2. Cambridge, MA: M.
- James, Edward (1982). *The Origins of France, 500-1000*. London: MacMillan.
- (1988). *The Franks*. Oxford: Blackwell.
- Jensen, Frede (1974). *The Syntax of the Old French Subjunctive*. The Hague: De Gruyter Mouton.
- (1984). *The syntax of medieval Occitan*. Beihefte zur Zeitschrift für romanische Philologie 208. Tübingen: Niemeyer.
- Jonas, Dianne Elisabeth (1996). “Clause Structure and Verb Syntax in Scandinavian and English”. PhD thesis. University of Harvard.
- Jónsson, Jóhannes Gísli (1996). “Clausal architecture and case in Icelandic”. PhD thesis. University of Massachusetts at Amhurst.
- Josefsson, Gunlög (2003). “Non-target structures and non-target uses in Child and Adult Swedish”. In: *The Acquisition of Swedish Grammar*. Ed. by Gunlög Josefsson, Christer Platzack, and Gisela Håkansson. Amsterdam and Philadelphia: John Benjamins, pp. 155–195.
- Julien, Marit (2007). “Embedded V2 in Norwegian and Swedish”. In: *Working Papers in Scandinavian Syntax* 80, pp. 103–161.
- Embedded clauses with main clause word order in Mainland Scandinavian* (2009). URL: lingbuzz/000475.
- Julien, Marit (2015). “The force of V2 revisited”. In: *The Journal of Comparative Germanic Linguistics* 18, pp. 139–181.
- Kaiser, Georg A (2002). *Verbstellung und Verbstellungswandel in den romanischen Sprachen*. Linguistische Arbeiten 465. Tübingen: Max Niemeyer.

- Kaiser, Georg A (2009). “Losing the Null Subject. A Contrastive Study of (Brazilian) Portuguese and (Medieval) French”. In: *Proceedings of the Workshop "Null Subjects, Expletives and Locatives in Romance"*. Ed. by Georg A. Kaiser and Eva-Marie Remberger. Arbeitspapier 123. University of Konstanz: Fachbereich Sprachwissenschaft, pp. 131–156.
- Kaiser, Georg A. and Franziska Maria Hack (2013). “Language change in comparison: The (special) case of Raeto-Romance”. In: *Sprachwandervergleich – Comparing Diachronies*. Berlin: Mouton de Gruyter, pp. 73–98.
- Kaiser, Georg A and Michael Zimmermann (2011). “On the decrease in subject-verb inversion in French declaratives”. In: *The Development of Grammar. Language acquisition and diachronic change. In honour of Jürgen M. Meisel*. Ed. by Esther Rinke and Tanja Kupisch. Amsterdam and Philadelphia: John Benjamins, pp. 355–381.
- Kato, Mary Aizawa (1999). “Strong pronouns and Weak Pronominals in the Null Subject Parameter”. In: *Probus* 11, pp. 1–37.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kiparsky, Paul (1995). “Indo-European origins of Germanic syntax”. In: *Clause structure and language change*. Ed. by Adrian Battye and Ian Roberts. Oxford: Oxford University Press, pp. 140–169.
- Kiparsky, Paul and Carol Kiparsky (1970). “Fact”. In: *Progress in Linguistics*. Ed. by Manfred Bierwisch and Karl Erich Heidolph. The Hague: Mouton. Chap. Fact, pp. 143–173.
- Klævik-Pettersen, Espen (2018). *Replication Data for: "Inversion, V-to-C, and verb second: an investigation into the syntax and word order of Old French and Late Latin"*. URL: <https://doi.org/10.18710/MKJU89>.
- Koeneman, Olaf (2000). “The flexible nature of verb movement”. PhD thesis. University of Utrecht.
- de Kok, Ans (1985). *La place du pronom personal régime conjoint en français*. Amsterdam: Radopi.
- Koopman, Hilda (1984). *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- Koster, Jan (1975). “Dutch as an SOV-Language”. In: *Linguistic Analysis* 1, pp. 111–136.
- Krifka, Manfred (2014). “Embedding illocutionary acts”. In: *Recursion: complexity in cognition*. Ed. by Thomas Roeper and Margaret Speas. Dordrecht: Springer, pp. 59–87.
- (2017). “Biscuit Conditionals as Anticipatory Speech Acts”. Presentation at the workshop ‘Biscuit Conditionals’, University of Hamburg, October 20-21.
- Kroch, Anthony (1989). “Reflexes of grammar in patterns of child language”. In: *Language Variation and Change* 1, pp. 199–244.
- (2005). “Modeling language change and language acquisition”. Manuscript, University of Pennsylvania. Available from <http://www.ling.upenn.edu/kroch/papers/lsa-forum.pdf>. Accessed May 2018.
- Kühner, Raphael and Carl Stegmann (1955). *Ausführliche Grammatik der lateinischen Sprache*. 3rd ed. Vol. 2. Leverkusen: Gottschalksche Verlagsbuchhandlung.
- Labelle, Marie (2007). “Clausal Architecture in Early Old French”. In: *Lingua* 117.1, pp. 289–316.
- Labelle, Marie and Paul Hirschbühler (2005). “Changes in clausal organization and the position of clitics in Old French”. In: *Grammaticalization and Parametric Change*. Ed. by Monserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca. Oxford: Oxford University Press, pp. 60–71.

- Labelle, Marie and Paul Hirschbühler (2014). “Y avait-il antéposition stylistique en ancien français?” In: *Congrès Mondial de Linguistique Française*. 8. SHS Web of Conferences. DOI 10.1051/shconf/20140801129, pp. 277–296.
- (2018). “Topic and focus in Old French V1 and V2 structures”. In: *Canadian Journal of Linguistics* 63.2, pp. 264–287.
- Labov, William (2007). “Transmission and diffusion”. In: *Language* 83.2, pp. 344–387.
- Langus, Alan and Marina Nespors (2015). “On the nature of word order regularities”. In: *Representing Structure in Phonology and Syntax*. Ed. by Marc van Oostendorp and Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 141–166.
- Larsson, Ida and Janne B. Johannessen (2015). “Incomplete Acquisition and Verb Placement in Heritage Scandinavian”. In: *Moribund Germanic Heritage Languages in North America: Theoretical Perspectives and Empirical Findings*. Ed. by R.S. Page and M. T. Putnam. Leiden: Brill Academic Publishers, pp. 153–189.
- Le Coultre, Jules (1875). “De l’ordre des mots dans Chrestien de Troyes”. PhD thesis. University of Leipzig.
- Ledgeway, Adam (2005). “Moving through the left periphery: the dual complementiser system in the dialects of southern Italy”. In: *Transactions of the Philological Society* 103.3, pp. 339–396.
- (2008). “Satisfying V2 in Early Romance: Merge vs. Move”. In: *Journal of Linguistics* 44.2, pp. 437–470.
- (2012). *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford Studies In Diachronic and Historical Linguistics. Oxford and New York: OUP.
- (2017). “Late Latin Verb Second: The Sentential Word Order of the *Itinerarium Egeriae*”. In: *Catalan Journal of Linguistics*. Catalan Journal of Linguistics 16. Ed. by Jaume Mateau Fantanals and Renato Oniga, pp. 163–216.
- Lehmann, Winfred P (1972). “Converging Theories in Linguistics”. In: *Language* 48, pp. 266–275.
- Lemieux, Monique and Fernande Dupuis (1995). “Clause Structure and Language Change”. In: *The Locus of Verb Movement in Non-Asymmetric Verb-Second Languages: The Case of Middle French*. Ed. by Ian Roberts and Adrian Batty. Oxford: Oxford University Press, pp. 80–110.
- Lerch, Eugen (1934). *Historische Französische Syntax*. Leipzig: O.R. Reisland.
- Lightfoot, David (1989). “The Child’s Trigger Experience: Degree-0 Learnability”. In: *Behavioral and Brain Sciences* 12, pp. 321–375.
- (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: The MIT Press.
- (1999). *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Wiley-Blackwell.
- Linde, Paul (1923). “Stellung des Verbs in der lateinischen Prosa”. In: *Glotta* 12, pp. 153–178.
- Lippert, Jörg (1974). *Beiträge zur Technik und Syntax althochdeutscher Übersetzungen unter besonderer Berücksichtigung der Isidorgruppe und des althochdeutschen Tatian*. Munich: Fink.
- Lodge, R. Anthony (2001). *French: from dialect to standard*. London: Routledge.
- Löfstedt, Einar (1936 [1911]). *Philologischer Kommentar zur Peregrinatio Aetheriae: Untersuchungen zur Geschichte der lateinischen Sprache*. Uppsala: Almqvist & Wiksell.

- Loseth, Eilert (1891/1970). *Le Roman en prose de Tristan, le Roman de Palamede, et la compilation de Rusticien de Pise. Analyse critique d'après les manuscrits de Paris*. New York: Burt Franklin.
- Lucas, Christopher (2015). "Contact-induced language change". In: *The Routledge Handbook of Historical Linguistics*. Ed. by Claire Bowern and Bethwyn Evans. London and New York: Routle. Chap. 24, pp. 519–536.
- Maling, Joan (1990). "Inversion in embedded clauses in modern Icelandic". In: *Modern Icelandic Syntax*. Ed. by Joan Maling and Annie Zaenen. New York: Academic Press, pp. 71–91.
- Maraval, Pierre, ed. (1982). *Égérie: Journal de voyage (Itinéraire)*. Sources Chrétiennes. Paris: Les éditions du cerf.
- Marchello-Nizia, Christiane (1985). *Dire le vrai: L'adverbe "si" en français médiévale. Essai de linguistique historique*. Publications Romanes et Françaises 168. Geneva: Droz.
- (1995). *L'évolution du français: Ordres des mots, démonstratifs, accent tonique*. Armand Colin.
- (1998). "Dislocations en ancien français: thématization ou rhématisation?" In: *Cahiers de praxématique* 30, pp. 162–178.
- (1999). *Le français en diachronie: douze siècles d'évolution*. Paris and Gap: Ophrys.
- Mariás, Julián (1990). *Understanding Spain. Translated by Frances M. López-Morillas*. University of Michigan Press and University of Puerto Rico Press.
- Marouzeau, Jules (1922). *L'ordre des mots dans la phrase latine. Vol. 1: Les groupes nominaux*. Paris: Champion.
- Martins, Ana Maria (2000). "Polarity items in Romance: Underspecification and lexical change". In: *Diachronic Syntax: Models and Mechanisms*. Ed. by Susan Pintzuk, George Tsoulas, and Anthony Warner. Oxford: Oxford University Press, pp. 191–219.
- Marx, Friedrich (1909). "Die Beziehungen des Altlateins zum Spätlatein". In: *Neue Jahrbücher für das klassische Altertum, Geschichte und deutsche Literatur* 23, pp. 434–448.
- Mathieu, Eric (2006a). "Quirky Subjects in Old French". In: *Studia Linguistica* 60.3, pp. 282–312.
- (2006b). "Stylistic Fronting in Old French". In: *Probus* 18, pp. 219–266.
- (2009). "On the Germanic properties of Old French". In: *Historical Syntax and Linguistic Theory*. Ed. by Paolo Crisma and Giuseppe Longobardi. Oxford: OUP, pp. 344–357.
- (2013). "The left periphery in Old French". In: *Research on Old French: The State of the Art*. Ed. by Deborah L Arteaga. Dordrecht: Springer, pp. 327–350.
- Mathisen, Ralph Whitney (1993). *Roman Aristocrats in Barbarian Gaul. Strategies for Survival in an Age of Transition*. Austin: University of Texas Press.
- Meillet, Anthoine (1931). "Sur une période de bilinguisme en France". In: *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres, 75 année* 75.1, pp. 29–38.
- Ménard, Philippe, ed. (1987). *Le Roman de Tristan en Prose. Tome 1*. Genève: Librairie Droz.
- Meyer-Hermann, Reinhard (1988). "¿ Se debe la postposición del sujeto en español a una influencia árabe?" In: *Revista de filología española* 68, pp. 67–96.
- Mikkelsen, Line (2010). "On what comes first in a verb-second language". Unpublished manuscript, University of Berkeley.
- Miller, D. Gary (1975). "Indo-European: VSO, SOV, SVO, or all three?" In: *Lingua* 37, pp. 31–52.

- Mohr, Sabine (2009). “V2 as a single-edge phenomenon”. In: *Selected papers from the 2006 Cyprys syntaxfest*. Ed. by Kleanthes K. Grohmann and Phoevos Panagiotis. Newcastle-Upon-Tyne: Cambridge Scholars Publishing, pp. 141–158.
- Mohrmann, Christine (1958-1977). *Études sur le latin des chrétiens* (4 vols.) Roma: Storia y Letteratura.
- Müller, Gereon (2004). “Verb-Second as vP-first”. In: *Journal of Comparative Linguistics* 3.7, pp. 179–234.
- Müller, Stefan. (2018). *German clause structure: An analysis with special consideration of so-called multiple frontings*. With contributions by Felix Bildhauer and Philippa Cook. Berlin: Language Science Press.
- Murray, Jessie, ed. (1929). *La Vie de Saint Eustace*. Publié en ligne par l’ENS de Lyon dans la Base de français médiéval, <http://catalog.bfm-corpus.org/SEustPr1> dernière révision le 23-3-2016, Paris: Librairie Ancienne Honoré Champion.
- Mussafia, Alfredo (1898). “Enclisi o proclisi del pronome personale atono qual oggetto”. In: *Romania* 27.105, pp. 145–146.
- Mutz, Katrin (2009). “Über das Französische und seine Wechselwirkungen mit dem Deutschen”. In: *Unsere sprachlichen Nachbarn in Europa. Die Kontaktbeziehungen zwischen Deutsch und seinen Grenznachbarn*. Ed. by Christel Stolz. Bochum: Universitätsverlag Dr. N. Brockmeyer, pp. 47–67.
- Newmeyer, Friedrich J. (2005). *Possible and Probable Languages. A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.
- Nilsen, Øystein (2003). “Eliminating Positions: Syntax and semantics of sentential modification”. PhD thesis. University of Utrecht.
- Nissen, Harald (1943). “L’ordre des mots dans la chronique de Jean d’Outremeuse”. PhD thesis. University of Uppsala.
- Nordström, Jackie (2010). “The Swedish *så*-construction, a new point of departure”. In: *Working Papers in Scandinavian Syntax* 85, pp. 37–63.
- Oniga, Renato (2014). *Latin: A Linguistic Introduction*. Edited and translated by Norma Schifano. Oxford: Oxford University Press.
- Oniga, Renato and Carlo Cecchetto (2014). “Constituency as a Language Universal: the Case of Latin”. In: *Lingue Antiche e Moderne* 3, pp. 5–35.
- Önnerfors, Olaf (1997). *Verb-erst-Deklarativsätze. Grammatik und Pragmatik*. Lunder Germanische Forschungen 60. Stockholm: Almqvist & Wiksell International.
- Ott, Dennis (2014). “An Ellipsis Approach to Contrastive Left-Dislocation”. In: *Linguistic Inquiry* 45.2, pp. 269–303.
- (2018). “Stylistic Fronting as Remnant Topicalisation”. In: *Studia Linguistica* 72.1, pp. 1–38.
- Palmer, Leonard R. (1954/2001). *The Latin Language*. London: Bristol Classical Press.
- Panhuis, Dirk (1984). “Is Latin a SOV Language?” In: *Indogermanische Forschungen* 89, pp. 140–159.
- Paoli, Sandra (2007). “The Fine Structure of the Left Periphery: COMPs and Subjects: Evidence from Romance”. In: *Lingua* 117.6, pp. 1057–1079.
- Parkinson, Stephen (2000). “Portuguese”. In: *The Romance Languages*. Ed. by Martin Harris and Nigel Vincent. 3rd ed. London and New York: Routledge. Chap. 4, pp. 131–169.
- Pasch, Renate (1997). “Weil mit Hauptsatz - Kuckucksei im denn-Nest”. In: *Deutsche Sprache* 25, pp. 252–271.
- Penny, Ralph (2002). *A History of the Spanish Language*. 2nd ed. Cambridge: CUP.

- Pensalfini, Rob (2004). "Towards a Typology of Configurationality". In: *Natural Language and Linguistic Theory* 22.2, pp. 359–408.
- Petri, Franz (1973). *Siedlung, Sprache und Bevölkerungsstruktur*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Pinto, Carlos Felipe de C. (2011). "Ordem de palavras, Movimento do Verbo e Efeito V2 na História do Espanhol". PhD thesis. Unicamp.
- Pintzuk, Susan (1991). "Phrase structures in competition: variation and change in Old English clause structure". PhD thesis. University of Pennsylvania.
- (1995). "Variation and change in Old English clause structure". In: *Language Variation and Change* 7, pp. 229–260.
- Platzack, Christer (1985). "Narrative Inversion in Old Icelandic". In: *Íslenskt mál* 7, pp. 127–144.
- (1986). "Comp, Infl, and Germanic word order". In: *Topics in Scandinavian Syntax*. Ed. by Lars Hellan and Kristi Koch. Dordrecht: Reidel, pp. 365–391.
- (1995). "The loss of verb second in English and French". In: *Clause Structure and Language Change*. Ed. by Ian Roberts and Adrian Battye. New York: Oxford University Press, pp. 200–226.
- Poletto, Cecilia (2000). *The Higher Functional Field: Evidence from Northern Italian Dialects*. New York and Oxford: Oxford University Press.
- (2002). "The left-periphery of V2-Rhaetoromance dialects: a new view on V2 and V3". In: *Syntactic Microvariation*. Ed. by Sjeff Barbiers, Leonie Cornips, and Susanne van der Kleij. Meertens Institute.
- (2005). "'Si' and 'E' as CP Expletives in Old Italian". In: *Grammaticalization and Parametric Variation*. Ed. by Monserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca. Oxford: Oxford University Press, pp. 206–235.
- (2006). "Parallel phases: a study on the high and low left periphery of Old Italian". In: *Phases of Interpretation*. Ed. by Maria Frascarelli. Studies in Generative Grammar 91. Berlin: Mouton de Gruyter, pp. 261–294.
- (2014). *Word Order in Old Italian*. Oxford: Oxford University Press.
- Posner, Rebecca (1996). *The Romance Languages*. Cambridge: CUP.
- Prévost, Sophie (2001). *La postposition du sujet en français aux XVe et XVIe siècles*. Paris: CNRS Editions.
- Price, Granville (1966). "Contribution à l'étude de la syntaxe des pronoms personnel sujet en ancien français". In: *Romania* 87, pp. 476–504.
- (1973). "Sur le pronom personnel sujet postposé en ancien français". In: *Revue Romaine* 8, pp. 226–236.
- Prince, Ellen F (1981). "Toward a taxonomy of given-new information". In: *Radical pragmatics*. Ed. by Peter Cole. New York: Academic Press, pp. 223–255.
- Pulgram, Ernst (1958). *The Tongues of Italy: Prehistory and History*. New York: Greenwood.
- Radwan, Sonja (2011). "L'ordre des mots en ancien français. La syntaxe et la structure informationnelle de la zone préverbale". MA thesis. University of Oslo.
- Ramat, Anna Giacalone and Caterina Mauri (2011). "The grammaticalization of coordinating interclausal connectives". In: *The Oxford Handbook of Grammaticalization*. Ed. by Bernd Heine and Heiko Narrog. Oxford: Oxford University Press.
- Reinholtz, Charlotte (1989). "V-2 in Mainland Scandinavian: Finite Verb Movement to Agr". In: *WPSS* 44, pp. 101–117.
- Reis, Marga (1974). "Syntaktische Hauptsatz-Privilegien und das Problem der deutschen Wortstellung". In: *Zeitschrift für Germanische Linguistik* 2.3, pp. 299–327.

- Reis, Marga (2006). "Is German V-to-C Movement Really Semantically Motivated? Some Empirical Problems". In: *Theoretical Linguistics* 32.3, pp. 369–380.
- (2013). "'Weil-V2'-Sätze und (k)ein Ende? Anmerkungen zur Analyse von Antomo und Steinbach (2010)". In: *Zeitschrift für Sprachwissenschaft* 32.2, pp. 221–262.
- Reis, Marga and Angelika Wöllstein (2010). "Zur Grammatik (vor allem) konditionaler V1-Gefüge im Deutschen". In: *Zeitschrift für Sprachwissenschaft* 29, pp. 111–179.
- Repp, Sophie and Heiner Drenhaus (2011). "Working-memory effects of information-structure in German left dislocation (GLD)". Poster presented at Architectures and Mechanisms for Language Processing 17, University of Paris Descartes.
- Ribeiro, Ilza (1995). "Evidence for a Verb-Second Phase in Old Portuguese". In: *Clause Structure and Language Change*. Ed. by Adrian Battye and Ian Roberts. Oxford: OUP, pp. 110–139.
- Riemsdijk, Henk van (1997). "Left Dislocation". In: *Materials on left-dislocation*. Ed. by Elena Anagnostopoulou, Henk van Riemsdijk, and Frans Zwarts. Amsterdam: John Benjamins, pp. 1–10.
- Riemsdijk, Henk van and Edwin Williams (1986). *Introduction to the theory of grammar*. Cambridge, MA: MIT Press.
- Rinke, Esther (2003). "On the Licensing of Null Subjects in Old French". In: *Syntactic Structures and Morphological Information*. Ed. by Uwe Junghanns and Luka Szucsich. Berlin: Mouton de Gruyter, pp. 217–247.
- (2009). "Verb Placement in Old Portuguese". In: *Focus and Background in Romance Languages*. Ed. by Andreas Duffer and Daniel Jacobs. Studies in Language Companion Series 112. Amsterdam: John Benjamins, pp. 309–332.
- Rinke, Esther and Martin Elsig (2010). "Quantitative Evidence and Diachronic Syntax". In: *Lingua* 120.11, pp. 2557–2568.
- Rinke, Esther and Jürgen M. Meisel (2009). "Subject inversion in Old French: Syntax and information structure". In: *Proceedings of the Workshop "Null subjects, Expletives, and Locatives in Romance"*. Ed. by Georg A. Kaiser and Eva-Marie Remberger. Arbeitspapier 123, Universität Koblenz. Constance, pp. 93–130.
- Rizzi, Luigi (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- (1997). "The fine structure of the left periphery". In: *Elements of Grammar: Handbook of Generative Syntax*. Ed. by Liliane Haegeman. Kluwer International Handbooks of Linguistics 1. Dordrecht: Kluwer, pp. 281–337.
- (2001). "Locality and left periphery". Ms; University of Siena.
- (2004). "On the Cartography of Syntactic Structures". In: *The Structure of CP and IP: The Cartography of Syntactic Structures*. Ed. by Luigi Rizzi. Vol. 2. New York: Oxford University Press, pp. 3–15.
- (2007). "On some properties of Criterial Freezing". In: *CISCL Working Papers on Language and Cognition* 1, pp. 145–158.
- (2013). "Notes on cartography and further explanation". In: *International Journal of Latin and Romance Linguistics* 25.1, pp. 197–226.
- Roberts, Ian (1993). *Verbs and Diachronic Syntax: A Comparative History of English and French*. Studies in Natural Language and Linguistic Theory 28. Dordrecht: Kluwer.
- (2004). "The C-System in Brythonic Celtic Languages, V2, and the EPP". In: *The Structure of CP and IP*. Ed. by Luigi Rizzi. Oxford: OUP, pp. 297–328.
- (2011). "Head Movement and the Minimalist Program". In: *The Oxford Handbook of Linguistic Minimalism*. Ed. by Cedric Boeckx. Oxford: OUP, pp. 195–219.

- Roberts, Ian (2012). “Phases, Head Movement and Second-Position Effects”. In: *Phases: Developing the Framework*. Ed. by Ángel J. Gallego. Berlin; Boston: De Gruyter Mouton, pp. 385–440.
- Roberts, Ian and Anders Holmberg (2010). “Introduction: Parameters in minimalist theory”. In: *Parametric Variation. Null Subjects in Minimalist Theory*. Ed. by Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan. Cambridge: CUP, pp. 1–57.
- Rögnvaldsson, Eiríkur and Höskuldur Thráinsson (1990). “On Icelandic word order once more”. In: *Modern Icelandic Syntax*. Ed. by Joan Maling and Annie Zaenen. Syntax and Semantics 24. New York: Academic Press.
- Ross, John R. (1973). “The penthouse principle and the order of constituents”. In: *You Take the High Node and I’ll Take the Low Node*. Ed. by Claudi Corum, Cedrik Smith-Stark, and Weiser Ann. Chicago: Chicago Linguistic Society, pp. 397–422.
- Rowlett, Paul (2007). *The Syntax of French*. Cambridge and New York: Cambridge University Press.
- Salvesen, Christine Meklenborg (2011). “Stylistic Fronting and Remnant Movement in Old French”. In: *Romance Languages and Linguistic Theory 2009: Selected Papers from ‘Going Romance’ Nice 2009*. Ed. by Janine Berns, Haike Jakobs, and Tobias Scheer. Amsterdam: John Benjamins, pp. 323–342.
- (2013). “Topics and the Left Periphery: A Comparison of Old French and Modern Germanic”. In: *In Search of Universal Grammar: From Old Norse to Zoque*. Ed. by Terje Lohndal. Amsterdam: John Benjamins, pp. 131–172.
- Salvesen, Christine Meklenborg and Kristin Bech (2014). “Postverbal subjects in Old English and Old French”. In: *Corpus-based Studies in Contrastive Linguistics*. Ed. by Signe Oksefjeld Ebeling, Atle Grønn, Kjetil Rå Hauge, and Diana Santos. Oslo Studies in Language 6 (1), pp. 201–228.
- Salvesen, Christine Meklenborg and George Walkden (2017). “Diagnosing embedded V2 in Old English and Old French”. In: *Micro-change and macro-change in diachronic syntax*. Ed. by Eric Mathieu and Robert Truswell. Oxford: Oxford University Press, pp. 168–181.
- Salvi, Giampaolo (2000). “La formazione della sistema V2 delle lingue romanze antiche”. In: *Lingua e Stile* 35, pp. 665–92.
- (2004). *La formazione della struttura di frase romanza: ordine delle parole e clitici dal latino alle lingue romanze antiche*. Tübingen: Max Niemeyer.
- (2012). “On the Nature of the V2 System of Medieval Romance”. In: *The Cartography of Syntactic Structures. Volume 7: Functional Heads*. Ed. by Laura Brugè. Oxford and New York: Oxford University Press, pp. 103–111.
- Samo, Giuseppe (2018). “A Criterial Approach to the Cartography of V2”. PhD thesis. University of Geneva.
- Sankoff, Gillian (2004). “Adolescents, young adults and the critical period: two case studies from 7-Up”. In: *Sociolinguistic variation: critical reflections*. Ed. by Carmen Fought. Oxford: Oxford University Press, pp. 121–139.
- Santorini, Beatrice (1989). “The Generalization of the Verb-Second Constraint in the History of Yiddish”. PhD thesis. University of Pennsylvania.
- Schmid, Monika M. (2002). *First Language Attrition, Use and Maintenance: The case of German Jews in anglophone countries*. Amsterdam: John Benjamins.
- Schönenberger, Manuela (2006). “A glimpse of doubly-filled COMPs in Swiss German”. In: *Organizing Grammar: Studies in Honor of Henk van Riemsdijk*. Ed. by Hans Broekhuis et al. Berlin: Mouton de Gruyter, pp. 572–581.
- Schrijnen, Johannes (1932). *Charakteristik des altchristlichen Lateins*. Nijmegen.

- Schwartz, Bonnie and Stein Vikner (1989). “All verb-second clauses are CPs”. In: *Working Papers in Scandinavian Syntax* 43, pp. 27–49.
- (1996). “The verb always leaves IP in V2 clauses”. In: *Parameters and Functional Heads*. Ed. by Adriana Belletti and Luigi Rizzi. Oxford: Oxford University Press, pp. 11–62.
- Serbat, Guy (1996). *Grammaire fondamentale du latin*. Vol. VI: L’emploi des cas en Latin. Louvain-Paris: Peeters.
- Sigurðsson, Haldor A. (1989). “Verbal syntax and case in Icelandic”. PhD thesis. University of Lund.
- Simons, Mandy (2007). “Observations on embedding verbs, evidentiality and presupposition”. In: *Lingua* 117, pp. 1034–1056.
- Sitaridou, Ioanna (2011). “Word order and information structure in Old Spanish”. In: *Catalan Journal of Linguistics* 10, pp. 159–184.
- (2012). “A comparative study of word order in Old Romance”. In: *Folia Linguistica* 46.2, pp. 553–604.
- Skårup, Povl (1975). *Les premières zones de la proposition en ancien français: essai de syntaxe de position*. *Révue romane*. Numéro spécial. 6. Copenhagen: Akademisk forlag.
- Spevak, Olga (2005). “‘Itinerarium Egeriae’: l’ordre des constituants obligatoires”. In: *Mnemosyne* 58, pp. 235–261.
- (2010). *Constituent Order in Classical Latin Prose*. Amsterdam: John Benjamins.
- Stanford, James N (2015). “Language acquisition and language change”. In: *The Routledge Handbook of Historical Linguistics*. London and New York: Routledge.
- Starke, Michael (2001). “Move dissolves into Merge: a Theory of Locality”. PhD thesis. University of Geneva.
- Steiner, B. D. (2014). “The evolution of information structure and verb second in the history of French”. PhD thesis. Indiana University.
- Thiersch, Craig L. (1978). “Topics in German Syntax”. PhD thesis. MIT.
- Thomas, Charles (1981). *Christianity in Roman Britain to AD 500*. Berkeley and Los Angeles: University of California Press.
- Thomason, Sara G. and Terence Kaufman (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA: University of California Press.
- Thráinsson, Höskuldur (1979). *On complementation in Icelandic*. New York: Garland.
- (1986). “V1, V2, V3 in Icelandic”. In: *Verb Second Phenomena in Germanic Languages*. Ed. by Hubert Haider and Martin Prinzhorn. Dordrecht: Foris Publications, pp. 169–194.
- (2007). *The Syntax of Icelandic*. Cambridge: Cambridge University Press.
- Thrift, Erica (2001). “Topicalization and Topic Drop in Dutch”. In: *Calgary Working Papers in Linguistics* 23, <http://dx.doi.org/10.5072/PRISM/28966>.
- Thurneysen, Rudolf (1892). “Stellung des Verbums im Altfranzösischen”. In: *Zeitschrift für romanische Philologie* XVI, pp. 289–307.
- Tomaselli, Alessandra (1995). “Cases of verb third in Old High German”. In: *Clause Structure and Language Change*. Ed. by Adrian Battye and Ian Roberts. Oxford: Oxford University Press, pp. 345–369.
- Travis, Lisa (1984). “Parameters and effects of word order variation”. PhD thesis. MIT, Cambridge, Mass.
- Trotzke, Andreas and Stefano Quaglia (2016). “Particle topicalisation and German clause structure”. In: *The Journal of Comparative Germanic Linguistics* 19.2, pp. 109–141.
- Truckenbrodt, Hubert (2006). “On the semantic motivation of syntactic verb movement to C in German”. In: *Theoretical Linguistics* 32.3, pp. 257–306.

- Väänänen, Veikki (1987). *Le journal-épître d’Égérie. Etude linguistique*. Annales Academiae Scientiarum Fennicae. Suomalainen Tiedeakatemia.
- Valois, Daniel and Fernande Dupuis (1992). “On the status of (verbal) traces in French: the case of Stylistic Inversion”. In: *Romance languages and modern linguistic theory*. Ed. by Paul Hirschbühler and E.F.K. Koerner. Amsterdam: John Benjamins, pp. 325–338.
- van Kemenade, Ans (1987). *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.
- Vance, Barbara S (1989). “The evolution of pro-drop in Medieval French”. In: *Studies in Romance Linguistics*. Ed. by Carl Kirschner and Janet Ann DeCesaris. Amsterdam: John Benjamins, pp. 413–441.
- (1993). “Verb-first declaratives introduced by et and the position of pro in Old and Middle French”. In: *Lingua* 89, pp. 281–314.
- (1995). “On the decline of verb movement to COMP in Old and Middle French”. In: *Clause Structure and Language Change*. Ed. by Adrian Battye and Ian Roberts. Oxford: Oxford University Press, pp. 170–199.
- (1997). *Syntactic Change in Medieval French : Verb-Second and Null Subjects*. Studies in Natural Language and Linguistic Theory 41. Dordrecht: Kluwer.
- Vance, Barbara S, Bryan Donaldson, and B. Devon Steiner (2009). “V2 Loss in Old French and Old Occitan: The Role of Fronted Clauses”. In: *Romance Linguistics 2009. Selected Papers from the 39th Linguistic Symposium on Romance Languages (LSRL), Tuscon, Arizona*. Ed. by Colina Sonia, Antxon Olarrea, and Ana Maria Carvalho. Current Issues in Linguistic Theory 315. Amsterdam: John Benjamins, pp. 301–320.
- Vanelli, Laura (1998). “Ordine delle parole a articolazione pragmatica dell’italiano antico: La ‘prominenca’ pragmatica della prima posizione nella frase”. In: *Medioevo Romano* 2, pp. 229–246.
- Vanelli, Laura, Lorenzo Renzi, and Paola Benincà (1985). “Typologie des pronoms sujets dans les langues romanes”. In: *Actes du XVIIème Congrès Internation de Linguistique et Philologie Romanes (Aix-en Provence, 29 août - 3 septembre 1983)*. Vol. 3. Aix-en Provence: Université de Provence, pp. 163–176.
- Varvaro, Alberto (2013). “Latin and the making of the Romance languages”. In: *The Cambridge History of the Romance Languages*. Ed. by A. L. Martin Maiden John Charles Smith. New York: Cambridge University Press, pp. 6–56.
- Vennemann, Theo (1972). “Rule Inversion”. In: *Lingua* 29, pp. 209–242.
- Verstraete, Jean-Christophe (1999). “The distinction between epistemic and speech act conjunction”. In: *Belgian Essays on Language and Literature*, pp. 119–130.
- Vikner, Stein (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: OUP.
- (2001). “Verb movement variation in Germanic and Optimality Theory”. PhD thesis. Universität Tübingen.
- Villa-García, Julio (2012). “Recomplementation and locality of movement in Spanish”. In: *Probus* 24.2, pp. 257–314.
- Vincent, Nigel (1977). “Perceptual Factors and Word Order Change in Latin”. In: *Romance Syntax: Synchronic and Diachronic Perspectives*. Ed. by Martin Harris. Salford: University of Salford, pp. 54–68.
- (2000). “Latin”. In: *The Romance Languages*. Ed. by Martin Harris and Nigel Vincent. London and New York: Routledge, pp. 26–78.
- Volterra, Virginia and Traute Taeschner (1978). “The acquisition and development of language by bilingual children”. In: *Journal of Child Language* 5, pp. 311–326.

- Wackernagel, Jacob (1892). “Über ein Gesetz der indogermanischen Wortstellung”. In: *Indogermanische Forschungen* 1, pp. 333–436.
- Walkden, George (2014). *Syntactic reconstruction and Proto-Germanic*. Oxford Studies in Diachronic and Historical Linguistics 12. Oxford: Oxford University Press.
- (2016). “V2”. Manuscript. University of Manchester.
- (2017). “Language contact and V3 in Germanic varieties new and old”. In: *Journal of Comparative Germanic Linguistics* 20.1.
- Walkden, George and Hannah Booth (to appear). “Reassessing the historical evidence for embedded V2”. In: *Rethinking verb second*. Ed. by Theresa Biberauer, Sam Wolfe, and Rebecca Woods. Oxford University Press. Chap. Oxford.
- von Wartburg, Walther (1939). *Die Entstehung der romanischen Völker*. Halle-Saale: Niemeyer.
- (1958). *Evolution et structure de la langue française*. 5th ed. Bern: Francke.
- Watkins, Calvert (1964). “Preliminaries to the reconstruction of Indo-European sentence structure”. In: *Proceedings of the Ninth International Congress of Linguistics*. Ed. by Horace G. Lunt. The Hague, pp. 1035–1045.
- Wechsler, Stephen (1991). “Verb second and illocutionary force”. In: *Views on Phrase Structure*. Ed. by Katherine Leffel and Denis Bouchard. Dordrecht: Kluwer, pp. 177–191.
- Weil, Henri (1887). *The Order of Words in the Ancient Languages Compared with That of the Modern Languages*. Boston: Ginn & Co.
- Weiss, Helmut (2005). “Inflected Complementizers in Continental West Germanic Dialects”. In: *Zeitschrift für Dialektologie und Linguistik* Bd. 72, Heft 2, pp. 148–166.
- Weiss, Michael (2015). “The Routledge History of Historical Linguistics”. In: *The Comparative Method*. Ed. by Claire Bowerman and Bethwyn Evans. Abingdon and New York: Routledge. Chap. 4, pp. 127–145.
- Westergaard, Marit (2009). *The Acquisition of Word Order: Micro-Cues, Information Structure, and Economy*. Amsterdam: John Benjamins.
- Westergaard, Marit and Terje Lohndal (to appear). “Verb Second Word Order in Norwegian Heritage Language: Syntax and Pragmatics”. In: *Variable properties in language: Their nature and acquisition*. Ed. by David Lightfoot and Jon Havenhill. Washington DC: Georgetown University Press.
- Westman, Margarita (1974). *Bruksprosa*. Lund: Gleerups Liber Läromedel.
- Wiklund, Anna-Lena (2010). “In search of the force of dependent verb second”. In: *Nordic Journal of Linguistics* 33.1, pp. 81–91.
- Wiklund, Anna-Lena, Kristine Benzen, Gunnar Hrafn Hrafnbjargarson, and Thorbjörg Hróarsdóttir (2009). “On the distribution and illocution of V2 in Scandinavian that-clauses”. In: *Lingua* 119.2, pp. 1914–1938.
- Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Thorbjörg Hróarsdóttir, and Kristine Bentzen (2007). “Rethinking Scandinavian Verb Movement”. In: *Journal of Comparative Germanic Linguistics* 10, pp. 203–233.
- Williams, Edwin (1974). “Rule Ordering in Syntax”. PhD thesis. Cambridge, MA: MIT.
- Wolfe, Sam (2015a). “Medieval Sardinian: New Evidence for Syntactic Change from Latin to Romance”. In: *Historical Linguistics 2013. Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013*. Ed. by Dag Haug. Amsterdam: John Benjamins, pp. 303–324.
- (2015b). “Microvariation in Medieval Romance Syntax: A Comparative Study”. PhD thesis. University of Cambridge.
- (2015c). “The nature of Old Spanish verb second reconsidered”. In: *Lingua* 164, pp. 132–155.

- Wright, Roger (1982). *Late Latin and Early Romance in Spain and Carolingian France*. Liverpool: Cairns.
- ed. (1991). *Latin and the Romance Languages in the Early Middle Ages*. London and New York: Routledge.
- (2002). *A Sociophilological Study of Late Latin*. Turnhout: Brepols.
- (2013). “Periodization”. In: *The Cambridge History of the Romance Languages*. Ed. by Martin Maiden, John Charles Smith, and Adam Ledgeway. New York: Cambridge University Press, pp. 107–124.
- Wurmbrand, Susanne (2014). “The merge condition. A syntactic approach to selection”. In: *Minimalism and Beyond*. Ed. by Peter Kosta, Stephen L. Franks, Teodora Radeva-Bork, and Lilia Schürcks. Amsterdam: John Benjamins, pp. 139–177.
- Yang, Charles (2003). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.
- Zaenen, Annie (1997). “Contrastive dislocation in Dutch and Icelandic”. In: *Materials on left-dislocation*. Ed. by Elena Anagnostopoulou, Henk van Riemsdijk, and Frans Zwarts. Amsterdam: John Benjamins, pp. 119–148.
- Zagona, Karen (1982). “Government and proper government of verbal projections”. PhD thesis. University of Washington.
- Zaring, Laurie (2017). “The nature of V2 in Old French: Evidence from subject inversion in embedded clauses”. In: *Canadian Journal of Linguistics/Révue canadienne de linguistique* 63.2, pp. 288–308.
- Zeijlstra, Hedde (2004). “Sentential Negation and Negative Concord (unpublished Ph.D thesis)”. PhD thesis. University of Amsterdam.
- Zimmermann, Michael (2009). “On the evolution of expletive subject pronouns in Old French”. In: *Proceedings of the workshop ‘Null-subjects, expletives and locatives in Romance.’ Arbeitspapier 123*. Ed. by Georg A Kaiser and Eva-Marie Remberger. Universität Konstanz: Fachbereich Sprachwissenschaft, pp. 63–92.
- (2014). *Expletive and Referential Subject Pronouns in Medieval French*. Berlin: De Gruyter.
- Zimmermann, Michael and Georg A Kaiser (2010). “Much ado about nothing? On the categorial status of *et* and *ne* in Medieval French”. In: *Corpus 9 (La syntaxe de corpus/Syntax corpus)*, pp. 265–290.
- Zwart, Jan Wouter (1991). “Clitics in Dutch: Evidence for the position of INFL.” In: *Groninger Arbeiten zur germanistischen Linguistik* 33, pp. 71–92.
- (1993). “Dialektsyntax”. In: *Clues from dialect syntax: Complementizer agreement*. Ed. by Werner Abraham and Josef Bayer. Linguistische Berichte. Sonderheft 5. Opladen: Westdeutscher Verlag, pp. 246–270.
- (1997). *Morphosyntax of verb movement: a Minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer.
- (2011). *The Syntax of Dutch*. Cambridge: Cambridge University Press.
- Zwicky, Arnold (1977). *On Clitics*. Bloomington: Indiana University Linguistics Club.