

Word Order in Old Catalan

Afra Pujol i Campeny

Section of Theoretical and Applied Linguistics
Modern and Medieval Languages Faculty

Queens' College
University of Cambridge

September 2018

This dissertation is submitted for the degree of Doctor of
Philosophy

ABSTRACT

This thesis describes the word order of *El Llibre dels Fets del Rey en Jacme*, a 13th century Catalan chronicle widely known as *El Llibre dels Fets*. By exploring the position of the verb, the subject and the information structure of the Old Catalan of this text, it is concluded that Old Catalan was an SVO language that shared several features with its modern counterpart, such as the invariable location of informational focus to the right of the verb. For this purpose, a database of 2,000 syntactically parsed main and embedded clauses was produced, the *Llibre dels Fets del Rey en Jacme* (LFRJ) database.

Verb position is studied in terms of linear verb positions and their frequencies, in regard to predicate types, in relation to adverb position, clitic placement, and in relation to polarity in both main and embedded clauses, showing that in both main and embedded clauses, the verb is located in the inflectional layer unless a specific trigger is identified as a trigger of V-to-C movement, as is the case with the expression of emphatic polarity. Furthermore, Old Catalan is shown to have an active projection in the inflectional domain where subjects and other discourse-anchoring elements, such as locatives, can occur, like its modern counterpart. Therefore, unlike it has been claimed for other Old Romance languages, Old Catalan had a specialised verbal prefield. Subjects can also be left and right-dislocated, as well as focalised. In terms of Information Structure, it is shown that Old Catalan's Informational Focus is located postverbally, and not preverbally as in other Old Romance languages. Finally, a comparison between Modern Catalan and Old Catalan data shows that Modern Catalan oral texts share word order features with the data from *El Llibre dels Fets*, highlighting the importance of orality as a factor determining word order.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specified in the text.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution.

The word count of this dissertation is 79,991 words, including footnotes and references but excluding the bibliography.

Word count: 79,991 words

ACKNOWLEDGEMENTS

E, quan aquest libre és aital, que coses de menuderies no y deu hom metre...

El Llibre dels Fets, Fol. 116v, 14

I arrived at Cambridge in October 2010, and little did I suspect then that I would end up writing up these lines to acknowledge all those that have contributed in different capacities to me completing this doctoral degree.

First and foremost, I need to thank the Catalan school system and the excellent teachers that made me discover language, history, and the pleasure of learning beyond the textbook. In particular, Assumpció Rost and her excellent syntax classes, Maria Lluïsa Rich, for her inspiring Ancient Greek lessons, and Sefa, for her thorough teaching of the Latin language. I would also like to thank the University of Girona, who awarded me with a Botet i Sisó scholarship and allowed me to take my first steps in research by the hand of Avelina Suñer Gratacós.

My first three years in Cambridge were an incredible intellectual and personal adventure that took place in the welcoming and stimulating contexts of the Faculty of Classics and Homerton College. I went through my BA in Classics under the supervision of Dr Pippa Steele, my ever-encouraging director of studies, and the support of excellent friends who have remained by my side since: Jack and Jon.

When I decided to pursue further studies in Linguistics, I counted with the support of ‘la Caixa’, who awarded me with a full scholarship for my MPhil in Theoretical and Applied Linguistics, and the AHRC, who fully funded my PhD in Old Catalan Linguistics at Queens’ College and has funded several work trips that have been key to my professional development, for which I am most grateful. I have been fortunate to count with Dr Ioanna Sitaridou as a supervisor for both the MPhil’s long essay and the PhD, providing me with constant support, valuable advice, and great opportunities to grow as a researcher, a teacher and a person. Ευχαριστώ πολύ.

Working within the Department of Theoretical and Applied Linguistics has allowed me to exchange ideas and receive the input of great colleagues among which Dr Theresa Biberauer, Dr Adam Ledgeway, Prof Ian Roberts, Prof Wendy Bennett, Prof Mari Jones and Dr Michelle Sheehan. The Graduate Centre and the hard-working lot that has inhabited it throughout these four years has provided an endless source of amusement and encouragement during lunch-breaks and sneaky trips to the pub. Thank you Kim, Alice, Ele, Abhi, Drasko,

Oli, and of course, Ben and Jamie. An immense thanks to all those that have commented on my work in conferences, talks and via email: Dr George Walkden, Dr Christine Salvesen, Prof Lluïsa Hernanz, Dr Susan Fischer, and the various reviewers of my work.

As a Batista i Roca fellow at Fitzwilliam College, I have also enjoyed the invaluable guidance and support of Prof Dominic Keown and the College itself, a place that I am proud to call home. As a member of the Spanish and Portuguese Section, I have been lucky enough to have Coral Neale guide me around the Section's bureaucratic maze, with her endless optimism and great generosity. *Muchísimas gracias*, Coral.

A big thank you to all my friends. Those in Cambridge, those in Catalonia, and those elsewhere. To na Laura, for all those Barcelona nights. To Víctor Lillo, for making it to our annual meeting all these years. To Jon, for the surreal and sarcastic conversations. To the *CUCS*, for being Catalan with me in the country of tea. To Chrissie, for all the *gambes* nights, timely chorizo risottos and *Taste the Difference* sauvignons blancs. To Kim and Francesco, the best housemates one could dream of (as long as one switches the lights off). To Harold, for his native English judgements (but... does he actually speak English?). To Víctor Acedo-Matellán, for the numerous *Art after Dark* sessions, the laughs, the moussakas, and the dairy free experiments. This thesis would not be the same without him. To Jack (I don't think I have to say anything else, here). To Edu, for pretending that he actually enjoys providing me with grammatical judgements when he actually loathes them! To Artur, for making me feel like I can still speak Catalan and like I could finish this thesis. To Ángel, for sharing his PhD woes with me and listening to mine. *Per les nenes, per ser-hi sempre. Gràcies.*

I ara, pels de casa. An immense thank you to all my family, for their tireless support, endless patience and constant encouragement at every step I have taken to reach this point. To Francesc Xavier Casademunt i Arimany, for his generous work on sayings and his love for our language and our country, which has inspired me to take the path that has lead me here. To Narcís Pujol i Martí and Rosa Casademunt i Piñol, for more than I can say here. To Facundo Campeny i Carbó and Rosa Puig Puig, for the long afternoons in the orchard. To Jordi Pujol Casademunt, for encouraging me to live fearlessly. To my siblings, for making me laugh when it seems impossible. To Isabel Campeny i Puig, for that *Història de Catalunya Il·lustrada* book that the *tió* shat me in 1998 and *El Llibre de les Bèsties* that the Three Wise Men got me in 2002. For all the pigeons, the nights by the fireplace with Marrufa and the walks with Nina. For putting up with Tit and his dribbling. For being always there.

Et finalement, pour Aurél. Pour m’avoir accompagné à chaque pas de cette aventure, pendant qu’on enfilait la nôtre, avec des après-midis de *punting*, happy hours sur la terrasse, promenades par Paris et nuits de canicule qui semblaient n’avoir aucune fin. Et pour tout ce qui nous attend.

And one last ‘thank you’ to Jaume I, the author (or at least ideologue) of the source of most of my data: *El Llibre dels Fets*. Thanks to him for writing in plain prose (!) about his life and for reproducing dialogue fragments that will forever accompany me, with the noteworthy contributions of Hug Malavespa, Pere Grony and Ramon de Muntcada.

*“Rendets-vos, En Ramon, rendets-vos!” e dix En Ramon:
“¿a qui·m rendria, miynon podent, a qui·m rendria?”*

El Llibre dels Fets, Fol. 24r, l 25

TABLE OF CONTENTS

1	Introduction	1
2	Methodology (and the data)	4
2.1	The Text	4
2.2	Databases	6
2.2.1	The LFRJ database	6
2.2.2	Non-finite form database	9
2.2.3	<i>Sí</i> database	9
2.2.4	<i>Victus</i> database	9
2.2.5	<i>Sobretaula</i> database	9
2.2.6	<i>Greuges</i> data base	10
3	Verb position in Old Catalan	11
3.1	Verb Position in Main Clauses	15
3.1.1	Verb position frequencies in Old Catalan	15
3.1.2	Verb initial clauses	17
3.1.3	Verb third clauses	43
3.1.4	Adverb position	46
3.1.5	Clitic placement	48
3.1.6	Verb position and polarity	56
3.2	Verb position in embedded clauses	81
3.3	Summary	87
4	Subject Position in Old Catalan	89
4.1	Old Catalan: a null subject language	89
4.2	Subject distribution in Modern Catalan main clauses	96
4.2.1	Modern Catalan as a VOS Language	96
4.2.2	Modern Catalan as a SVO Language	101
4.2.3	Considerations about Subject Position in Modern Catalan	104
4.3	Subject distribution in Old Catalan main clauses	105
4.3.1	The Old Catalan preverbal field: specialised for hosting subjects?	122
4.3.2	Locatives and the information structure value SpecTP	130
4.3.3	Summary	147
4.4	Subject distribution in Old Catalan embedded clauses	147

4.5	Summary	156
5	Old Catalan information structure	157
5.1	Basic notions	157
5.2	The left periphery in the LFRJ database	161
5.2.1	Scene Setters	162
5.2.2	Hanging Topic Left Dislocation	170
5.2.3	Clitic Left Dislocated Topics	174
5.2.4	Foci	180
5.3	Interim summary	199
5.4	Phenomena related to information structure	200
5.4.1	Anteposition of non-finite forms	200
5.4.2	Recomplementation	210
5.5	Summary	220
6	Register variation and word order	221
6.1	<i>El Llibre dels Fets</i> : a representative text for Old Catalan syntax?	221
6.2	<i>El Llibre dels Fets</i> and register variation in Modern Catalan	224
6.3	<i>El Llibre dels Fets</i> and the legal register	230
6.4	Summary	235
7	Conclusion and future prospects	236
7.1	Summary of the thesis and its main findings	236
7.2	Future prospects	239
8	References	242

LIST OF ABBREVIATIONS

1/2/3PL	1 st , 2 nd or 3 rd person plural
1/2/3SG	1 st , 2 nd or 3 rd person singular
ABL	Ablative case
ACC	Accusative case
Adv	Adverb
CFoc	Contrastive Focus
CL	Clitic
CILD	Clitic Left Dislocation
COND	Conditional
DAT	Dative
DO	Direct Object
DOM	Differential Object Marking
EF	Edge Feature
EPPA	Emphatic Positive Polarity Particle
FF	Focus Fronting
FocP	Focus Phrase
HT	Hanging Topic
HTLD	Hanging Topic Left Dislocation
IFoc	Informational Focus
INF	Infinitive
IO	Indirect Object
Lex	Lexical subject
MCat	Modern Catalan
MFr	Modern French
MGal	Modern Galician
MIt	Modern Italian
MPt	Modern Portuguese
MSic	Modern Sicilian
MSp	Modern Spanish
NEG	Negation
NP	Nominal element
NPI	Negative Polarity Item

NSL	Null Subject Language
NSP	Null Subject Parameter
OCat	Old Catalan
OFr	Old French
OIt	Old Italian
OIt	Old Italian
OOc	Old Occitan
OPt	Old Portuguese
OSard	Old Sardinian
OSp	Old Spanish
OVen	Old Venetian
PPI	Positive Polarity Item
Pres	Present tense
Pron	Pronoun
PST	Past tense
PTCP	Past participle
SBJV	Subjunctive
TopP	Topic Phrase
TP	Tense Phrase
V1	Verb in clause initial position
V2	Verb in clause second position
V3	Verb in clause third or more position
VP	Verb phrase
XP	Non tagged constituent

1 INTRODUCTION

This thesis offers a synchronic description of Old Catalan word order based on 13th century data. With it, we provide a missing piece in the puzzle of Medieval Romance word order, which has largely focused on other varieties, most notably Old French (Adams, 1987; Labelle, 2007; Roberts, 1993; Salvesen, 2013; Sitaridou, 2004; Vance, 1993, 1997; Vance, Donaldson, & Steiner, 2010), Old Occitan (Donaldson, 2016; Jensen, 1994; Kunert, 2003; Lafont, 1967; Sitaridou, 2004; Steiner, 2014), Old Italian dialects (Benincà, 1984, 2004, 2006; Benincà & Poletto, 2004b; Ledgeway, 2005; Wolfe, 2014, 2015a), Old Spanish (Batllori, 1993, 2015, 2016; Batllori & Hernanz, 2013; Bossong, 2006; Bouzouita, 2014; Elvira, 2015; Fontana, 1997; Hernanz & Rigau, 2006; Kaiser, 2002; Poole, 2013; Sitaridou, 2009, 2011, 2012, 2019; Wolfe, 2015b) and Old Portuguese (Martins, 1994, 2005, 2011; Ribeiro, 1995; Rivero, 1993).

While Old Catalan is understudied in comparison with other Old Romance varieties, this work fits into a long standing tradition that starts in the early 20th century with the project of the Alcover-Moll dictionary, by Antoni Maria Alcover i Sureda and, later, Francesc de Borja Moll, followed by Anfós Par's (1923) work explicitly on Old Catalan syntax, and the inestimable work of Pompeu Fabra, mostly on Modern Catalan (including the first Catalan grammar, published in 1918), ensued by the vast lexicographic works of Joan Corominas and Antoni Maria Badia i Margarit, among many others. In more recent times, studies on Catalan language have thrived in many fields, including the syntax of Old Catalan, which generally has focused on specific phenomena. To my knowledge, to date no other work has been done attempting to comprehensively describe Old Catalan word order after Par (1923).

The data for this thesis has been mainly extracted from *El Llibre dels Fets del Rey en Jacme*, also known as *El Llibre dels Fets*, a 13th century chronicle. The data is used to describe the position of the verb, the distribution of subjects and the Information Structure of 13th century Old Catalan, comparing them with the position of the verb, the distribution of subjects and the information structure of other Old Romance languages where appropriate.

For this work, we abide by the Cartographic Programme (Belletti, 2004; Cinque, 1999; Rizzi, 1997, among others). The Cartographic Programme proposes the existence of a universal hierarchy of functional projections, whose lexicalisation varies cross-linguistically and accounts for superficial differences across languages. Within the Cartographic Programme, movement is linked to discourse or criterial features, whose checking is associated with semantico-pragmatic notions such as Focus and Topic. Ample empirical

evidence from a variety of linguistic families has confirmed the syntactic reflexes of these categories. The basic structure of the left periphery, as described in the Cartographic Programme, consists of the following projections:

- (1) Force > Frame Field > Topic Field > Focus Field > FinP

The thesis is structured as follows:

In Chapter 2, the sources of data used in this thesis as well as the methodology used to gather and describe it are presented. The main source of data for this thesis is the LFRJ database, built on *El Llibre dels Fets*, but other databases have been built in order to complement it. They include databases built on Modern and Old Catalan texts.

Chapter 3 is concerned with verb position in Old Catalan. Verb position has attracted a lot of attention in the literature on Old Romance syntax since Adams (1987) and Benincà (1984) proposed that Old French and Old Italian respectively had a V2 grammar (a grammar that required for the verb and a constituent to raise to the left periphery). In this Chapter, we identify the position of the verb in unmarked main declarative clauses as being in the inflectional domain, below the left periphery. We also show that Old Catalan verb-initial clauses are truly verb initial, and pattern with Modern Catalan ones. Furthermore, we describe the distribution of clitics in main declarative clauses and show that it is not linked to verb position but to Information Structure. We also explore the interaction between emphatic polarity and verb position, identifying structures that involve the emphatic positive particle *sí* as cases in which the verb does raise to the left periphery in main clauses. On the basis of the evidence presented, it is concluded that Old Catalan did not have verb movement to the left periphery in unmarked main declarative clauses, and that its verb distribution patterns that of MCat, an SVO language.

In Chapter 4, the distribution of subjects in Old Catalan is explored. Firstly, we describe Old Catalan as a null subject language, just like its Modern counterpart. This is followed by a description of subject distribution in main and embedded clauses, where it is shown that the same positions are available to subjects in both syntactic environments, and that crucially, Old Catalan has a subject-dedicated position in the preverbal field, unlike Old Romance languages described as V2. The possibility of this projection being filled with elements other than nominal expressions, mainly locatives, is also explored. It is concluded that subject distribution in Old Catalan patterns with that of Modern Catalan, an SVO language.

In Chapter 5, we describe Old Catalan's left periphery and compare it with descriptions of the left periphery of other Old Romance languages. While it is found that the broad structure of the left periphery of Old Catalan abides by Cartographic descriptions of it, a significant difference is identified: Old Catalan does not have a preverbal projection that can host Informational Focus. Instead, Informational Focus is found postverbally. Furthermore, Focus Fronting, a phenomenon linked with polarity reversal shared with Modern Catalan, is explored at length. Finally, two phenomena linked with Informational Structure are explored in depth: anteposition of non-finite forms and recomplementation. The former consists of a description and analysis of the fronting of non-finite verbal forms of periphrastic tenses to the left periphery, and the latter, a discussion of clauses in which the complementiser *que* 'that' occurs twice.

In Chapter 6, the interaction between word order and register is explored. This is done by comparing data from the LFRJ database with that of two databases built on Modern Catalan texts, showing that genre and register are key factors in determining the frequency of certain word order patterns, with the notion of orality being central to it. This also holds when the LFRJ data is compared to a legal Old Catalan text, pointing at the importance of the study of syntactic variation linked to register in the study of the syntax of the Old Romance languages. With the acknowledgement of syntactic variation linked to register in mind, a reflection upon the notion of representativeness is brought forward, and its importance in the study of historical syntax is emphasised.

Finally, Chapter 7 provides a summary of the main findings of the thesis and points at questions for future research.

2 METHODOLOGY (AND THE DATA)

2.1 THE TEXT

The data used for this thesis are extracted from the *Llibre dels Fets del Rey En Jacme* (henceforth LFRJ), a 13th century Catalan chronicle. This text was deemed suitable for the study of word order based on three factors: (i) it is the first *crònica* ‘chronicle’, a genre that did not have a precedent in Catalan historiography and that broke with the established *annales* tradition;¹ (ii) the probable involvement of King James I in the production of the text, which brings us closer to the modern notion of authorship; and (iii) the first manuscript available of the text, Poblet’s Manuscript, has a very well documented historical context that suggests that it is the closest to the archetype. In what follows, each of these factors is explored independently.

LFRJ is the first of the four Great Catalan Chronicles. They are four historiographical works in prose that break with the preceding historiographical tradition. Their common denominator is their theme: they narrate and praise the feats of several Catalan kings who reigned during the 13th and 14th centuries in the Crown of Aragon: the LFRJ, written between 1244 and 1274; the *Crònica de Bernat Desclot* or *Llibre del rei en Pere d'Aragó e dels seus antecessors passats* ‘the Chronicle of Bernat Desclot’ or the ‘Book of the King Peter of Aragon and of his Predecessors’, written in 1288; *Crònica de Ramon Muntaner* ‘Chronicle of Ramon Muntaner’, written between 1325 and 1328; and *Crònica de Pere el Cerimoniós* ‘Chronicle of Peter II, the Ceremonious’, written between 1382-1383 and 1385.

Unlike their historiographical predecessor, the *Gesta Comitum Barchinonensium*, modled on Latin *annales* (factual record of events in the form of a list), chronicles presented a vivid narrative style. Several narrative devices, including dialogues, changes of register, personal comments, among others, were used to render chronicles amusing to the listener and credible. In the case of LFRJ, the first of the *Cròniques*, this change of register emanates from the King’s desire to legitimise the conquest of the Kingdom of València and the Balearic Islands, as well as his supremacy among the Catalan and Aragonese nobility.

Thus, the use of the *Crònica* genre in the case of LFRJ was not only a tool for listing the achievements of the King, as the *Gesta Comitum Barchinonensium*, but also a

¹ See Aurell (2005; 2008), and Martí i Castell (2002), for an overview of the evolution of Catalan historiographical literature.

propagandistic tool to portray King James I as a pious, hardworking, heroic and yet humane figure with whom the audience can empathise.

By being the first narrative work in prose to break with the Latin historiographic tradition and not having contemporary models in the Romance literature, LFRJ does neither have Latin influence in word order typical of legal texts,² nor the influence of the source language often found in translations. Therefore, it is a good candidate for the study of word order.

Secondly, a further argument in favour of the suitability of this text for the study of word order is the involvement of King James I in the production of LFRJ, as proposed by Bruguera (1998; 2012); and Ferrando i Francès (2012). Both authors offer compelling arguments in favour of the King's authorship of *El Llibre dels Fets*, among which: the oral nature of the text, the use of an intra-homodiegetic narrator (see also de Jong [2015, Chapter 2]) that identifies with King James I (the story is narrated in the first person plural, the majestic plural), the wealth of intimate details from the King's life,³ and finally, references to the composition of the text itself.⁴ The active involvement of King James I in the production of the text as argued by Bruguera (2012) and Ferrando i Francès (2012) gets us closer to the modern idea of authorship in spite of the fact that the text was certainly written down by a scribal body, the *Cancelleria* (Soldevila, 2007), and allows us to predict stability in word order across the text as it is one of the linguistic domains least affected by dialectal variation.⁵ Since the last part of the book could not count with the King's collaboration as it narrates his death, it has been excluded from the database.

Finally, the wealth of information that we have on Poblet's Manuscript, on which the edition of the text that I have used for the purpose of this thesis is based, allows us to establish

² See Castillo Lluch & López Izquierdo (2010) for a volume on the influence of Latin in OSp prose, and Aurell (2005, 2008) and Martí i Castell (2002) for more on the evolution of Catalan historiographic literature.

³ Examples of intimate details of the King's life include a reference to his first wedding night (Fol. 9v, l 9), mentions of visits to the Queen simply to 'make her happy' (Fol. 109, l 28), private conversations with the Queen (Fol. 116r, l 23) or mentions of the King crying of happiness (Fol. 120v, l 24), or sadness (Fol. 39v, l 14).

⁴ Read more on the composition of the book in Bruguera (2012); Casanova (2012); Ferrando i Francès (2012). Find more on the Medieval concept of authorship in Partridge & Kwakkel (2012).

⁵ There is controversy regarding the date of emergence of the two main Catalan dialectal areas, with some authors dating it as far back as the 14th century, just after the eastwards and southwards linguistic expansion that followed the conquest of the Kingdom of Majorca and the Kingdom of Valencia. Therefore, a relative level of dialectal/idolectal homogeneity can be assumed for the 13th century.

that the text is either a direct copy of the archetype or a contemporary copy of the archetype produced by the same scribal body in the *Cancellaria*. The manuscript dates of 1343 and was commissioned by the abbot of the monastery of Poblet (Catalonia), Ponç de Copons. For its production, he borrowed a copy of the text from Peter II, heir of James I. This fact is known to us thanks to a letter from Peter II to Ponç de Copons, asking for the book back (Ferrando i Francès 2012, p. 48). The book was copied by a monk, Celestí Destorrents, who signed and dated his work (see Fol. 201r in the manuscript).

The edition of the text used to construct the database is Bruguera (1991).

2.2 DATABASES

For the realisation of this thesis, three databases built on *El Llibre dels Fets*: the LFRJ database, the Non-finite form database and the *si* database; two databases were built on Modern Catalan texts: the *Victus* and the *Sobretaula* databases; and finally, one database was built from a legal Old Catalan text, the *Greuges* database. They are described in what follows.

2.2.1 The LFRJ database

The LFRJ database was conceived for the purpose of studying the word order of the LFRJ. It contains 1,000 main and 1,000 embedded clauses, drawn from three parts of the text (introduction, middle, and denouement, excluding the fragment that narrates the death of the King for authorship reasons, as explained in 2.1) in equal proportions so as to control for the intervention of different scribes as well as variation in the potential two periods of composition of the text (1244 and 1274).⁶

Main clauses exclusively contain declarative main clauses (interrogative, exclamative and imperative clauses have been excluded due to them being crosslinguistically linked to syntactic configurations linked to specific word order patterns). In the case of embedded clauses, only adverbial (causal, purpose, concessive, temporal, manner, comparative, consecutive and conditional) and complement clauses have been included. Relative clauses have been excluded given their cross-linguistically particular syntax, as the distribution of certain phenomena shows (see Chapter 5, section 5.5.4). This is also the case for non-finite

⁶ There is controversy around the double date hypothesis. See Riquer (1964) for arguments in favour of the double date of composition.

verbal clauses, which have not been included in the database as embedded clauses, even though they have been included as constituents of other clauses, main or embedded.

The following values have been controlled for each of the clauses, adapting the Hamburg methodology to this text:⁷

- i. Location in the text (*folium* and line).
- ii. Whether they appear in reported speech or in the narrative body.
- iii. Whether they display subject continuity with respect to the preceding clause (in the case of embedded clauses, if not coordinated, their main clause).
- iv. Their linear verb position (V1, V2 and V3<, understanding V3< as the verb occurring in the third linear position or later in the clause).
- v. Whether they are preceded by the coordinating conjunction *e* ‘and’.
- vi. For embedded clauses, their category (adverbial or completive, and within adverbials, their type) has also been controlled for.

Each clause has then been syntactically parsed constituent by constituent. Constituents have been tagged in terms of form and function, including the following categories:

⁷ The Hamburg methodology was developed for the Hamburg databases (Old Portuguese, Old French and Old Occitan) produced within the project ‘Historical Syntax of the Romance languages’, directed by Jürgen M. Meisel, at the Collaborative Research Center on Multilingualism at the University of Hamburg (2002–2004). As listed by Sitaridou (2012, p. 560), the Hamburg methodology abides by the following criteria: (a) the texts should be edited; (b) the edition of the text should be based on a single manuscript; (c) the manuscript should be clearly dated; (d) it is desirable that all texts have a clear geographical origin in order to control for dialectal variation; (e) the editor should respect orthographic peculiarities and conventions of the manuscript and clearly mark any interventions of his own. Where possible, I have adhered to it.

Form	Abbreviation	Functions	Abbreviation
Noun Phrases	NP	Subject Noun Phrase	SN
		Direct Object	DO
		Subject Complement	SC
		Vocative	Voc
Pronoun	Pron	Subject Pronoun	SP
Prepositional Phrases	PP	Indirect Object	IO
		Adjunct	PPAdj
		Prepositional Complement	PrepC
Adverb	Adv	Adjunct	AdvAdj
Adjective	Adj	Noun Complement	AdjP
Finite Adverbial Clauses	AdvC	Adjunct	AdvCAdj
Non-finite Adverbial Clauses	AdvC.non-fin	Adjunct	AdvC.non-fin.Adj
Completive Clauses	ComplC	Direct Object	DO.fin
		Subject Completive Clause	SUB.fin
Other	XP		

Table 1 – Constituent labels used in the LFRJ database

Each verb has been tagged according to the following categories: transitive, unaccusative, unergative, *verba dicendi*, reflexive, existential, and passive.

Unless otherwise indicated, this is the main source of data for this thesis.

2.2.2 Non-finite form database

This database contains a compilation of 110 clauses from the LFRJ containing anteposition of non-finite forms (infinitives and participles), excluding cases of mesoclisism (INF – CL – AUX_{HAVE.FUT/COND}). It was elaborated during my MPhil in Theoretical and Applied Linguistics (2013/2014), for the purpose of writing a 4,000-word essay on non-finite form anteposition in Old Catalan. Formally, it follows the same principles as the LFRJ database.

2.2.3 *Sí* database

The *sí* database was produced after I realised that the distribution of *sí* ‘indeed, yes’ in Old Catalan differed significantly from the distribution of homophonous and cognate lexical items in other Romance varieties, requiring further investigation. Within the LFRJ database there were only two instances of *sí*, too few to produce a solid account of its distribution in Old Catalan. Therefore, I compiled all clauses containing *sí* from *El Llibre dels Fets*, creating this database, which has not been syntactically parsed.

2.2.4 *Victus* database

In order to control for word order variation and its interaction with genre and channel, two MCat databases were created. They were created following the same methodology used for the LFRJ database, detailed above.

The *Victus* database consists of a database with 136 main clauses and 27 embedded clauses, extracted from the first 18 pages of the Spanish-written historical novel *Victus*, by Albert Sánchez Piñol, translated by Xavier Pàmies into Catalan. This work was chosen on the grounds of its belonging to the same genre as *El Llibre dels Fets*, and both have an intrahomodiagetic narrator that identifies with the main character of the story. Nevertheless, it has to be pointed out that, while the Modern Catalan translation is highly idiomatic (as stated by the author, who is himself a native Catalan speaker) and does not contain, *a priori*, interferences from Modern Spanish, it is nevertheless a translation.

2.2.5 *Sobretaula* database

The *Sobretaula* (henceforth *Sobretaula*) database was produced with the aim of having an oral text as a standard of comparison of the LFRJ database. It consists of 208 main clauses and 84 embedded Modern Catalan clauses, extracted from a family conversation recorded during dinner time, drawn from the *Corpus Oral de Conversa Col·loquial (COC)*. *Conversa09. Sobretaula al menjador*, ‘Oral Corpus of Colloquial Conversation (OCC).

Conversation 09. Chat after the meal in the dining room’, compiled by the Grup d’Estudis de Variació of Universitat de Barcelona.

2.2.6 *Greuges* database

The *Greuges* database was produced with the aim of having an Old Catalan text from another register and close to vernacular language, so as to compare it against data from the LFRJ database and explore whether register has an impact on word order variation. The chosen text were the *Greuges de Guitard Isarn, Senyor de Caboet* ‘the Grievences of Guitard Isarn, Lord of Caboet’, a written complaint of Guitard Isarn against Mir Arnall, the son of one of his vassals. The edited version used for building this database was Russell-Gebbett (1965, text 14, p. 74). Like in the case of the *Victus* and *Sobretaula* databases, the *Greuges* database has been syntactically parsed as described above.

3 VERB POSITION IN OLD CATALAN

Within Romance syntax, the position of the verb in the Medieval or Old⁸ varieties has been the object of an ongoing controversy pivoting around the following question: did the Old Romance languages go through a V2 stage that rendered them similar to Modern Germanic varieties? This claim has been supported by the correlation of certain syntactic phenomena in Germanic languages and Old Romance ones by different authors with data from several varieties (Benincà & Poletto [2004]; Benincà [2006]) for OIt and Medieval Romance, Adams [1987], Roberts [1993], Vance [1997]; Vance, Donaldson & Steiner [2010] for OFr, Fontana [1993] for OSp). In the most recent literature, V2 is not understood as a linear constraint by which the verb would systematically need to occur in the second position of a clause, but rather as a structural requirement for the verb to raise to a position in the left periphery of the clause (Holmberg, 2015; Wolfe, 2014, 2015a, 2015c).

For the purpose of this work, we follow the standard assumption that a clause consists of three domains: the thematic domain, headed by the predicate and where thematic roles are assigned and arguments are base-generated, VP; the functional domain where functional heads related to tense, aspect and mood are located, TP; and the left periphery of the clause, related to clause typing. The internal structure of each of these domains has been studied and described as containing different functional projections. For the structure of the left periphery, we follow Rizzi's (1997) cartographic programme and the work that has followed. We heavily rely on Benincà (2004, 2006) description of the Medieval Romance left periphery. TP and VP are discussed at greater length in Chapter 4, but we assume them both to contain projections associated to subjects: SpecTP and SpecvP.⁹

One of the challenges that V2 analyses of the Old Romance languages are faced with is the high frequency of V1 and V3 orders (Bossong, 2006; Kaiser, 2002; Sitaridou, 2012). Nevertheless, research advocating for the V2 hypothesis, relying mainly on Italo-Romance

⁸ Following the literature on Old Catalan linguistics, we refer to the variety in which *El Llibre dels Fets* was written as 'Old', and not Medieval. Abbreviations used to refer to the different Old Romance varieties are the following: OCat for Old Catalan, OSp for Old Spanish, OOc for Old Occitan, OFr for Old French, OIt for Old Italian, OPt for Old Portuguese, MCat for Modern Catalan, MSp for Modern Spanish, MFr for Modern French, and MPt for Modern Portuguese. Other abbreviations can be found in the thesis' List of Abbreviations.

⁹ For reasons presented in Chapter 4, we assume the existence of a subject specific projection in SpecTP for OCat and MCat, against the proposals put forth in Alexiadou & Anagnostopoulou (1998). Other authors who also defend the existence of such a projection include Costa (2008, Chapter 2) for MPt, López (2009, p. 134ff) for MSp and Forcadell (2013) for MCat.

and Gallo-Romance varieties, has shown that those word order patterns can fit in a nuanced analysis of the nature of V2 in Old Romance (Benincà, 2006; Poletto, 2014; Wolfe, 2015a).

At the same time, other authors among which Martins (1994; 2011) for OPt; Sitaridou (2011; 2012; 2015; 2016; 2019) for OSp, Batllori [2009, 2015], Batllori & Hernanz [2008, 2011, 2013] for OSp and OCat, Fischer [2002] for OCat) propose that the Old Romance languages did not present a V2 grammar. Instead, these authors defend that the Old Romance languages behaved similarly to their modern SVO counterparts, in that they had an articulate left periphery where information structure was syntactically encoded, yielding a high percentage of linear V2 clauses and that phenomena that have parallels with Germanic languages can be explained by the transition from Latin SOV basic word order to Modern Romance SVO (Martins, 1994, 2002; Sitaridou, 2012, 2015). Some of these authors defend that, even though the Old Romance languages did not abide by a V2 constraint, they did exhibit V-to-C movement linked to the expression of assertiveness in unmarked main clauses (Rodríguez Molina, 2010, 2014 and Sitaridou, 2019 for OSp), but that overall, they presented the core syntactic features associated to (S)VO languages in unmarked main clauses, mainly (i) having prepositions; (ii) having postnominal genitives; (iii) and having auxiliary-lexical verb sequences, (iv) having sentence-initial complementisers; (v) having DO (Direct Object)/IO (Indirect Object) (Leonetti, 2017, p. 887; Poletto, 2014, p. 43); and (vi) having a preverbal subject-dedicated position.¹⁰

In Germanic linguistics, the V2 parameter was first explored by Den Besten (1983). A vast amount of literature followed. Recently, Holmberg (2015) defined V2 languages as follows:

- (1) ‘A language is called a verb-second (V2) language when the finite verb is obligatorily the second constituent, either specifically in main clauses or in all finite clauses.’

(Holmberg, 2015, p. 342)

As already pointed out by Den Besten, the raising of the verb to the CP layer is triggered by the fact that a head within this domain is related to finiteness.

¹⁰ Leonetti (2017) and Poletto (2014) use these features to describe two different concepts. On the one hand, Leonetti (2017) applies them to SVO languages, while, on the other hand, Poletto applies it to VO languages. The main difference is that one implies the existence of a preverbal subject designated position, while the other does not.

It can be understood from Holmberg's definition that there are two types of V2 languages: those that have V2 only in main clauses, as is the case in Modern German, and those that also have it in embedded clauses, like Yiddish or Modern Icelandic, referred to as asymmetric and symmetric V2 languages respectively. In any case, for the V2 requirement to be satisfied, a preverbal constituent needs to move to the left periphery. Constituents directly generated in the left periphery do not count towards V2. Therefore, he concludes that the two defining features of V2 are:

(2)

- a. A functional head in the left periphery attracts the finite verb.
- b. This functional head has an EPP feature that triggers the re-merging of a constituent to its specifier position.

(Holmberg, 2015, p. 375)

Wolfe (2015a) puts forward a V2 analysis of several Old Romance varieties (OSic, OVen, OOc, OFr, OSp and OSard), departing from the bi-partite structure of the V2 parameter, by which a head within the left periphery, endowed with unvalued phi features, and an Edge Feature (EF) attracts the clause's finite verb and an XP to it. This head acts as probe and attracts the verb, which carries valued phi features, to it. Subsequently, the EF associated to this head triggers the movement of a phrasal category which need not be the subject, to its specifier. Wolfe divides the Old Romance languages in two groups: those that have the V2 constraint located in FinP, like OSp and OFr, and those that have it in ForceP, like OSic and OVen, the highest and lowest heads of the left periphery as understood by the Cartographic Programme. Languages with the V2 constraint located in FinP allow for base-generated elements in the left periphery to precede the verb and the fronted XP (they have to be generated, given that once a constituent moves to the left periphery, no more constituents may move to it, a phenomenon that is known as the bottleneck effect). Therefore, in these languages, V3 and V4 clauses can be found without incurring ungrammaticality. In those languages where the V2 constraint is located in ForceP, since there are no projections available above it to host base-generated constituents, V3 and V4 orders are not expected, and thus, so-called *relaxed V2* languages are accounted for.¹¹ The key fact is that any XP can be

¹¹ Nevertheless, see Hsu (2017) for a feature-scattering account of relaxed V2 systems.

moved to the specifier of the head endowed with the V2 constraint, no matter its featural make-up or its informational value.¹²

In recent literature on the Old Romance languages, the following defining features of their V2 grammar have been put forward (Poletto, 2014; Wolfe, 2015a): (i) V-to-C movement, (ii) Germanic-style subject-verb inversion, (iii) a non-specialised pre-field that can notably host non-CILD DOs, (iv) asymmetric pro-drop between main and embedded clauses. As exposed in Pujol i Campeny (2017, 2019), the OCat of the LFRJ database did not show any of these features:

- (i) Lack of V-to-C movement: in OCat, the verb in unmarked main declarative clauses systematically occurs above low TP adverbs and below high TP adverbs (Cinque, 2004), indicating that the verb is located within this field, and not above it.
- (ii) Lack of unequivocal cases of Germanic-style subject-verb inversion (no cases of subjects occurring between auxiliaries and lexical verbs, or between inflected verbs and their complements).
- (iii) The preverbal field is specialised for hosting subjects. Postverbal subjects always bear a particular informational interpretation and are found with predicates receivingthetic readings.
- (iv) Null subjects present the same distribution in main and embedded clauses.

In this Chapter, we focus on the description of the position of the verb in main and embedded clauses, point (i), showing that OCat did not have V-to-C linked to the fulfilment of the V2 constraint, nor did it have independent V-to-C movement for the encoding of assertive force. Points (ii-iv) are addressed in Chapter 4.

The Chapter is structured as follows: within section 3.1 Main Clauses, the frequency of linear verb positions in OCat main clauses is presented in 3.1.1; followed by an analysis of V1 clauses in 3.1.2; a description of V3 clauses in 3.1.3; a description of verb position in relation to adverbs in 3.1.4; a description of the interaction of verb position and clitics in 3.1.5; and finally, several considerations on the interaction between verb position and polarity in OCat in section 3.1.6. In 3.2 Embedded Clauses, the frequency of linear verb orders is presented, followed by a description of verb position in relation to adverbs. Finally, in section

¹² Acedo-Matellán (personal correspondence) notes that in Modern German, a strict V2 language, the choice of fronted constituent can be connected to its informational value.

3.3, we summarise the findings and establish that OCat did not have V-to-C movement, and that, in terms of verb position, it was not a V2 language, but that instead, it operated like MCat.

3.1 VERB POSITION IN MAIN CLAUSES

3.1.1 Verb position frequencies in Old Catalan

The LFRJ database displays the following frequencies of verb position in main and embedded clauses (V1 being understood as verb initial, V2 as the verb occurring in the second linear position, and V3<, as the verb occurring on the third, fourth or fifth linear position in the clause):

	V1	V2	V3<
Main Clauses	47.0% (470/1000)	44.5% (439/1000)	8.5% (85/1000)
Embedded Clauses	46.7% (467/1000)	50.3% (503/1000)	3.0% (30/1000)

Table 1 - Frequency of linear verb orders in the LFRJ database

As it can be seen in Table 1, the most frequent order in main clauses is V1 (including those preceded by the coordinating conjunction *e*, whose true verb initial condition is demonstrated in section 3.1.2.1), followed by V2, and V3< clauses. In embedded clauses, V2 is the most frequent word order pattern, closely followed by V1, while V3< is, again, the least attested.

A wide range of constituents can occur preverbally, including lexical and pronominal subjects, adverbs and adverbial phrases, prepositional phrases and embedded clauses in V2 and V3< clauses. There does not seem to be a limitation on the number of constituents that can occur to the left of the verb, even though, as shown in the Table 1, V3< orders are the least attested. Examples (3-8) illustrate these word order patterns:¹³

¹³ Throughout the examples provided in this paper, the inflected verb appear in **bold**, while other elements relevant for the example are underlined. Cases of focalised constituents will be highlighted with SMALL CAPS. Square brackets ([]) are used to delimit constituents when relevant.

V1 clauses

- (3) **Retrau** mon seyor sent Jacme que fe sens obres
states my lord saint James that faith without deeds
mortaés.
dead is

My lord Saint James states that faith without deeds dies.

Fol. 1r, l 1

- (4) E **vol** tant dir en romans que (...)
And wants much say.INF in Romance that

And in Romance, it means more or less that (...).

Fol. 1v, l 20

V2 clauses

- (5) Ara₁ **comptarem** en qual manera nós fom engenrats
now tell.1PL.FUT in which manner we were.1PL conceived
e en qual manera fo lo nostre neximent.
and in which manner was.3SG the our birth

Now we will tell how we were conceived and how our birth took place.

Fol. 3v, l 2

- (6) [Nostre pare, lo rey En Pere]₁, **fo** lo pus
our father the king Sir Peter was.3SG the most
franch rey que anch fos en Espanya (...)
honest king that ever was.3SG in Spain

Our father, the king Sir Pere, was the most honest king that there ever was in Hispania (...).

Fol. 4r, l 7

V3 clauses

- (7) E [a aquel dia]₁ [la comtessa]₂ **vench**.
and in that day the countess came.3SG

And that day the countess came.

Fol. 23v, l 26

- (8) E [puys]₁ [En Guillem de Montpestler]₂, [estant ella viva]₃, **près**
and then Sir Guillem of Montpellier being she alive took.3SG
·i_a· altra dona, (...).
one other woman

And then Sir Guillem of Montpestler, being her alive, took another woman (...).

Fol. 3r, l 11

As illustrated in (3) and (4), V1 clauses can be absolutely verb initial or be preceded by the coordinating conjunction *e*. In V2 clauses, any constituent (DO NPs, subject NPs, subject pronouns, adverbs, prepositional phrases, embedded clauses, non-finite forms, adjective phrases) can occur preverbally. In V3< clauses, combinations of those constituents that can be preverbal in V2 clauses are found. The preverbal field of V2 clauses is extensively examined in Chapter 4, section 4.3.1, while the structure of the left periphery of V3 clauses is described in Chapter 5.

In this Chapter, we are concerned with the identification of the position of the verb of OCat main clauses by using data from the LFRJ database, concluding that in OCat, like in MCat, the verb does not raise to the left periphery, but that instead it raised only up to the inflectional layer. The pieces of evidence in favour of this analysis brought forward are: (i) the position of the verb relative to adverbs, in Section 3.1.4; (ii) the fact that OCat clitic placement can be explained through the interaction with information structure and is not connected to verb placement, explored in Section 3.1.5; (iii) the true verb initial nature of V1 clauses, preceded or not by the coordinating conjunction *e*; and finally, (iv) the fact that the interaction of verb position and PolP, a head in the left periphery, is limited to emphatic positive polarity contexts. Furthermore, in Chapter 4, evidence is brought in favour of OCat having a preverbal field specialised for hosting subjects.

3.1.2 Verb initial clauses

In this section, we examine V1 clauses from the LFRJ database and contrast them against analyses proposed for V1 clauses of Old Romance varieties considered to present V2 grammar (Benincà, 2004; Poletto, 2005, 2014; Wolfe, 2015a). Firstly, we explore *eV* clauses, the most numerous type of V1 clauses in section 3.1.2.1. This is followed by an examination of V1

clauses in association with two predicate types cross-linguistically prone to exhibit this word order: *verba dicendi* and unaccusative predicates in section 3.1.2.2. Finally, absolute verb initial clauses are considered in section 3.1.2.3. Section 3.1.2.4 offers a summary of the conclusions reached.

As it can be observed in Table 2, most clauses are preceded by the coordinating conjunction *e* ‘and’, regardless of the linear position of the verb:

V1	470	47.0%	<i>e</i> V clauses	443	94%
			V clauses	27	6%
V2	445	44.5%	<i>e</i> V clauses	375	84%
			V clauses	70	16%
V3<	85	8.5%	<i>e</i> V clauses	77	90.6%
			V clauses	8	9.5%
			Total <i>e</i> V clauses	903	90.3%
			Total V clauses	97	9.7%
Total	1000	100		1000	100

Table 2 – Linear verb position in main clauses from the LFRJ database

The data shown in Table 2 suggests that the presence of *e* is not linked to verb position. *E* precedes 97% of the total main clauses of the LFRJ database. V1 clauses do indeed present the highest frequency of being preceded by *e*, 94%, followed closely by V3< clauses, which are preceded by *e* in 90.6% of cases. V2 clauses are the least preceded by *e*, in 84% of cases, with a 10% difference with V1 clauses. Table 2 also shows that absolute V1 are rare, suggesting that they have a marked nature.

The role of *e* in OCat as well as the nature of V1 in the LFRJ database clauses and their relationship with predicate types and Information Structure is explored in the following subsections.

3.1.2.1 Verb initial clauses and the role of *e*

As shown in Table 2 above, within V1 clauses, *e*V sequences are much more frequent than absolute verb-initial clauses (although it must be said that *e* ‘and’ also precedes an overwhelming majority of V2 and V3 clauses). In the literature on Old Romance word order, *e*V clauses have captured the authors’ attention since they present a challenge for a V2 analysis

of the Old Romance languages. Well studied V2 languages such as Modern German or Swedish lack the possibility of producing clauses in which the verb is only preceded by a coordinating conjunction (Holmberg 2015, p. 375), since the coordinating conjunction is always clause-external and cannot satisfy the V2 requirement.

Two types of analyses have been proposed to account for the existence of *eV* clauses in V2 accounts of Old Romance syntax. On the one hand, we find those that postulate the presence of a null topic continuity element (Benincà, 2004, 2006; Benincà & Poletto, 2004a) to account for V1 structures without necessarily associating it with the presence of *e*. On the other hand, other analyses propose that the *e* element is not a coordinating conjunction, but a particle associated to the presence of a null topic continuity element in the left periphery. This has been proposed for OIt (Poletto, 2005, 2014), OFr (Vance, 1993, 1997), and comparatively, for OFr, OIt, OSp, and OOc in Wolfe (2015c). These analyses are described in what follows.

Benincà (2006) and Benincà & Poletto (2004b) propose the following structure for the Old Romance left periphery:

- (9) [Force C° [Rel*wh* C°] / {Frame [ScSett][HT] C°} {Topic [LD] [LI] C°} {Focus [I Focus] [II Focus] / [Inter*wh*] C°} [Fin C°

Benincà (2006, p. 13, example 18)

In Benincà's analysis, the left periphery of Old Romance languages is divided in three fields, each being able to host different projections: the Frame field, where Scene Setting elements and Hanging Topics (henceforth HT) are located; the Topic field, where Clitic Left Dislocated Topics (henceforth CILD) and Familiar Topics (henceforth FamTop) are located, and finally the Focus field, where different types of Foci are located (the left periphery of OCat is extensively discussed in Chapter 5).

To account for the fact that in Old Romance, clitic pronouns could not occur in clause initial position due to them needing a preceding phonological host (a phenomenon that is known as Tobler-Mussafia Law, henceforth TML)¹⁴, and that in verb-initial clauses they consistently appear in enclisis to the verb (V – CL vs. XP – CL – V), Benincà (2004, p. 288) proposes, for *eV* and *e*-less clauses equally, that when the verb moves to FocP in the CP layer to check finite features and no constituent is merged in SpecFocP, the verb raises to a higher projection in the Topic field. The specifier of this projection would be filled by a *pro*-like

¹⁴ See section 3.1.5 in this Chapter for more on TML and clitic placement in OCat.

element, the ‘default’ topic of the discourse, interpreted on the basis of the context. Since FocP would be empty, enclisis would follow, given that proclisis is borne when SpecFocP is filled (the distribution of OCat clitic pronouns is discussed at length in section 3.1.6 in this Chapter). According to this, Old Romance V1 clauses would have had the following structure, with clitics remaining in FocP:

- (10) [Force [TopP [SpecTopP *pro*_{Top} [Top’ V_i] [FocP [SpecFocP [Foc’ CL t_i] [FinP ...

Poletto (2005, 2014) builds on Benincà’s idea of a null *pro*-like element to explain the occurrence of *eV* clauses within OIt’s V2 grammar. Poletto identifies two homophonous elements that derive from Latin E(T) coordinating conjunction within OIt : the coordinating conjunction *e* ‘and’, and the ‘Continuation of the same Discourse Configuration’ (CDC) (Poletto, 2014, p. 24) *e(t)*, associated with the presence of a null Topic in the clause. The function of the CDC marker is to establish that the content that follows is to be added to the established discourse.

According to Poletto, the *e(t)* CDC marker differs from *e* coordinating conjunction semantically and syntactically: (i) it can precede both main and embedded clauses; (ii) it can occur between main and embedded clauses, giving the impression that they are coordinated; (iii) it licenses null topics that need not be the subject of the preceding clause; and (iv) since it can be followed by other topics, it is located in the highest topic position, HT within Benincà & Poletto (2004b) description of the Old Romance left periphery, and hence, it is followed by enclisis, since it co-occurs with verb movement to the Topic field, as described above. (11) shows how *e* CDC marker can ‘seemingly’ coordinate main and embedded clauses as well as the formal difference between the coordinating conjunction and the topic continuity marker in certain dialects. In (12), the coordinating conjunction and the CDC marker present the same form, but the coordinating conjunction and the topic continuity functions can be distinguished by analysing their syntactic context:

- (11) quando entrò nella chiesa, et uno **parlò** e **disse**
 when got.3SG in;the church and one spoke.3SG and said.3SG

When he entered the church, one of them spoke and said...

Nov, XXV, 189, apud Poletto (2014, p. 24, example [38])

- (12) e, inebriato il pane dell’ odore che n’ uscia, del
 and put the bread in;the smoke that from;it came out of;the

mangiare e quelli lo **mordea**, e cosò, il **consumò**
 food and he it.CL= bit.3SG and thus it.CL= finished.3SG
 di mangiare, ricevendo il fumo e mordendolo
 of eat.INF getting the smoke and biting=it.CL

And, the bread soaked with the smell that was coming out from it, from the food, and he would bit it [the bread], and in this way, he finished eating it [the bread], receiving the fumes [of the food] and biting on it.¹⁵

Nov. VIII, 147, apud Poletto (2014, p. 25)

In (11), the CDC *et* occurs between an embedded adverbial temporal clause and the main clause, two elements that cannot be coordinated. In turn, the main clause is followed by the coordinating conjunction *e*, which coordinates it with the following main clause. In (12), the coordinating conjunction and the CDC marker present the same form. In the first three instances, *e* is the CDC marker preceding verb initial main clauses, while in the last one, it coordinates two adjunct gerunds.

According to Poletto, *e(t)* can not only be a CDC, but also license null topics with a specific reference. While subjects are the most common coreferent for the null topic, it can also be coreferent with the object of the preceding clause. This is illustrated in (13), where the null subject of the last clause, which contains *ed*, is coreferent with the direct object of the preceding clause.

(13) E lo valletto **presentò** lo presente e tro[vò] Merlino
 and the valet showed.3SG the present and found.3SG Merlin
 a cenare: ed **era** in mezzo di Biagio e di Labegues
 to dine.INF and was in between of Biagio and of Labegues

And the valet showed the present and found Merlin while he was having dinner between Biagio and Labegues.

Merlino 41, apud Poletto (2014, p. 26)

OIt

¹⁵ This translation is my own (with the inestimable help of Kim Groothius and Libero Iaquinto). Poletto's (2014, p. 25) translation is reproduced here:

i. 'He put the bread close to the smell which came from the meat and then he ate the bread up'.

The distribution and uses of *e(t)* in OIt are not paralleled by MIIt, where *e*, merely a coordinating conjunction, does not have a topic marker entry in the lexicon. However, according to the author, some residual topic marker uses persist in the modern language in cases in which the CP layer is activated by the presence of an operator (mainly interrogative and exclamative clauses). In OIt, since the verb would have moved to the left periphery in main declarative clauses, *e* would have been licensed in a wider range of contexts. (14) and (15) illustrate such uses of *e* in MIIt in an interrogative and an exclamative clause respectively:

- (14) E adesso?
 and now
 Now what?

Apud Poletto (2014, p. 26)

- (15) E che vestito che ti sei comprato!
 and what dress that yourself.CL= are bought
 What a dress you've bought yourself!

Apud Poletto (2014, p. 26)

MIIt

In sum, by postulating the existence of null Topic elements, Poletto makes *eV* clauses conform with the V2 grammar that she proposes for OIt. The null Topic would check the EPP feature in FinP, where the verb raises in her view, and then move to the SpecHT, the projection targeted by the verb in verb initial clauses (Benincà, 2006).

Wolfe, (2015a, p. 25, 2015c), building on Benincà & Poletto (2004) and Benincà (2006b) refines the null Topic hypothesis by labelling the *pro*-like element for *eV* clauses as *pro_{top}*. *Pro_{top}* must be coreferent with a preceding nominal expression, establishing Topic Continuity, as already noted for *eV* clauses by Poletto (2014); Vance (1993, 1997). *Pro_{top}* is assumed to be a variant of *pro*, bearing phi-features that allow for it to satisfy the EPP feature on FinP. However, it also bears an underspecified [*uTop*] feature that causes it to raise to the Topic field in *eV* clauses, licensing V-to-Top movement, which would account for V-CL orders in *eV* clauses.

A piece of evidence against these analyses comes from the syntactic distribution of *e*: if *e* were a CDC marker, it would be expected for it to be found following adversative, correlative or disjunctive conjunctions, since the conjunction would be extra-clausal and *e*

would be located in the left periphery. Nevertheless, no data in this direction is provided by the advocates of this hypothesis.

These accounts have in common the presence of a null topic element and the raising of the verb to a position in the left periphery within the Topic field. While we resort to a null elements to account for instances of OCat V1 associated to specific predicate types, OCat *e*V clauses behave like MCat *e*V1 clauses, and there is no reason to postulate the existence of a null Topic element to explain their distribution, given that (i) in OCat the verb did not raise to the CP layer in declarative clauses and there was no EPP feature to be satisfied, as shown in the subsequent sections in this Chapter, and that (ii) the distribution of *e* in OCat differs substantially from the distribution of *e* in OIt, and can be readily translated into MCat.

As we saw in Table 3, *e*V clauses make up the 88.33% of verb-initial clauses.

Example (16) shows that OCat *e* is located outside the clause, above ForceP.

- (16) E dixem-li: ‘Donchs, nós farem així com fer
 and said.1PL=to;him.CL then we do.1PL.FUT as do.INF

 devem: nós **citarem** altra vegada, e **siran** ·iii·,
 must.1PL we cite.1PL.FUT again and be.3SG.FUT 3

 e, si vol fer dret, nós lo **prendrem**.
 and if wants do.INF right we it.CL= take.1PL.FUT

And we said to him: ‘Then, we shall do as we must: we will cite them again, and it will be the third time, and, if he wants to act righteously, we shall accept it’.

Fol. 21v, l 24

- (17) e [_{ForceP} [_{HT} si vol fer dret [_{Foc} nós [_{Fin} [_{TP} ho prendrem]]]]]

(16) contains an instance of reported speech that contains two *e*, both coordinating conjunctions that create a chain of three main clauses. The subject of the second coordinated clause is not coreferent with the first clause’s subject, but with an adjunct, *altra vegada* ‘one more time’. The third clause’s subject is not coreferent with any of the elements present in the preceding coordinated clauses, but with one of the three individuals participating in the conversation. One could thus suppose that in the latter case, *e* is associated with the presence of a generic ‘discourse topic continuity’ null element, which according to Poletto (2014) should be placed in the Topic field, in the same projection as HTs. Nevertheless, (17) already contains an *if*-clause, which, according to Munaro (2010) are located in the high-left

periphery, just below ForceP, in HangingTopicP (see Chapter 5 for more on the nature of the OCat left periphery). The *if*-clause is then followed by a focalised personal pronoun in FocP, followed by proclisis. Therefore, the position where the null topic marked by *e* would be expected to appear is already filled up with other material, yielding an unproblematic V3 clause. This suggests that *e* is located above ForceP. For the time being we consider its position to be extra-clausal, but as it is discussed below, other analyses are possible.

Furthermore, unlike in OIt, *e* is never found between embedded and main clauses: if it precedes a clause (see example (11)), it precedes all its constituents, including Frame Setters and HTs, as shown in (18) and (19):

- (18) E, quan la comtessa no havia a qui recórrer
 and when the countess not had.3SG to whom turn.INF
 posqués, sinó a nós, per ·ii· raons és venguda
 could.3SG.PST.SBJV but to us for 2 reasons is come
 denant vós.
 in front of you

And, when the countess did not have anyone to turn to but you, for two reasons she came in front of you.

Fol. 22r, l 6

- (19) E pus a él plach, bé **deu** plaure a nós
 and since to him pleases indeed must.3SG please.INF to us
 de tot en tot.
 of all in all

And since it pleases him [God], it must please us in every way.

Fol. 30r, l 14

In (18), *e* precedes the adverbial clause introduced by *quan*, located in FrameP, and containing information that sums up the countess' situation, exposed in the preceding discourse. This contrasts with the OIt example reproduced in (11), in that *e* does not occur between the embedded clause, located in FrameP, and the core of the main clause, but it occurs in a position external to the main clause or above ForceP, as it is suggested below. In (19), like in (18), *e* precedes an adverbial embedded clause located in FrameP, instead of

appearing between the embedded clause and the core of the clause, as expected if *e* were indeed a discourse topic continuity marker, as it is the case in (11). On the basis of this evidence, Poletto (2005, 2014)’s analysis cannot be applied to data from the LFRJ database.

The use of *e* to bind a string of sentences that share the same discourse Topic, however, is undeniable in examples like (20):

- (20) **E** [nostra mare], sempre que nós fom nats, **envià**·ns
 and our mother as soon as we were.1PL born.PPT sent.3SG=us.CL
 a Sancta Maria **e** **portaren**-nos en los braces; **e** **deÿen**
 to Saint Maria and took.3PL=us.CL in the arms and said.3PL
 matines en la església de Nostra Dona
 morning mass in the church of Our Lady

‘And our mother, as soon as we were born, sent us to Saint Mary’s, and they carried us in their arms, and they were singing the morning mass in the church of Our Lady’.

Fol. 3v, l 14

In (20), all V1 sentences are adding new information on the same discourse Topic: the birth of King James I. Therefore, while *e* cannot be linked to the presence of a null topic below ForceP and the raising of the verb to a Topic projection, as has been shown in (18) and (19), it is clear that it has a particular informational value linked to discourse topic continuity. This value of *e* transcends the boundaries of V2/SVO grammars, as demonstrated by the fact that, when translating (20) and (22) into MCat, a null-subject SVO language with the verb in TP, the result is completely grammatical, as it is in some uncontested V2 varieties:^{16 17}

- (21) **I** la meva mare, tan punt vaig néixer, em **va**
 and the my mother as point go.1SG be.born.INF me go.3SG
enviar a Santa Maria, **i** em **van** portar a coll,
 send.INF to Saint Mary and me go.3PL carry.INF to neck

¹⁶ Similar usages of the coordinating conjunction ‘and’ to the ones described here for *e* in OCat have been described by Diesing (1990, p. 56) for Yiddish, and Sigurdsson (1990, p. 45) for Old Icelandic, two doubtlessly V2 languages.

¹⁷ OCat *e*, [e], raises and becomes MCat *i*, [i].

i **deien** matines a l'església de Nostra Dona.
and said.3PL morning mass in the church of Our Lady

'And our mother, as soon as we were born, sent us to Saint Mary's, and they carried us in their arms, and they were saying morning mass in the church of Our Lady'.

MCat

Even though the grammaticality of (21) is robust, the fact that the referents of 'servants' and 'clergymen' are not active in the mind of modern speakers renders it bizarre. Therefore, so as to double check that this type of *eV* strings is allowed in the modern language, we provide a further example in (22), translated into MCat in (23):

(22) E **dixem** nós: “ Con la vila tenen éls?” e sempre lexam
and said we how the village have.3PL they and soon left.1PL
los cavals als escuders e **avalam**; e **prenguem**
the horses to;the squires and descended.1PL and took.1PL
nostres armes e **anam** -los combatre e
our weapons and went.1PL =them.CL fight.INF and
tolguem-los la vila.
took.1PL=them.CL the village

'And we **said**: “What? They have the village?” and quickly we **left** the horses to the squires and we **got off** [the horses], and we **took** our weapons and we went to face them and we **took** the village back'.

Fol. 15v, 11

(23) I nosaltres **vam dir**: “Com que han pres la ciutat?”
and we go.1PL say.INF how that have.3PL taken the city
i ràpidament **vam deixar** els cavalls als escuders,
and quickly go.1PL leave.INF the horses to;the squires
i **vam baixar**, i **vam prendre** les nostres armes,
and go.1PL get;off.INF and go.1PL take.INF the our weapons
i **vam anar** -los a combatre, i
and go.1PL go.INF =them.CL to combat.INF and

vam **prendre**'ls la vila.
 go.1PL take.INF=them.CL the village

‘And we **said**: “What? They have the village?” and quickly we **left** the horses to the squires and we **got off** [the horses], and we **took** our weapons and we went to face them and we **took** the village back’.

MCat

In examples (20-23), the distribution of *e/i* is parallel in OCat and MCat, with the grammaticality of MCat examples having been confirmed by several native speakers. Thus, I propose that in OCat, like in MCat, *e* can act as a discourse cohesion marker, instead of a marker of topic continuity. Cohesion markers contribute to binding the underlying structure of the ideas in a text (Schiffrin et al. 2001, p. 55), specifying the relationship between two segments of discourse, creating a sequence (Fraser, 1998). This is exactly what we find in OCat and MCat. *E* can contribute to the cohesion of a certain fragment of the text by syntactically coordinating main clauses, but its meaning is not always strictly additive.¹⁸

By considering *e* a discourse cohesion marker that links a clause with the common ground that is being updated by the clause’s new content, it could be argued that it is located within the Speech Act layer immediately dominating CP in the left periphery. Haegeman (2014) describes the existence of two speech act projections above ForceP: a higher one, which hosts performative discourse markers, and a lower one, which hosts grounding discourse particles. It is possible that *e/i*, when functioning as a discourse cohesion marker, occupies the latter projection. This would suggest the existence of two homophonous lexical items: *i/e* as a coordinating conjunction and *e/i* as a discourse grounding particle or discourse cohesion marker. On the other hand, if *e* were truly a coordinating conjunction located outside the clause, it could be hypothesised that instead of coordinating CPs, it coordinates clauses from their Speech Act layer, where a null discourse continuity operator could be postulated. In the interest of space, this line of inquiry has to be pursued in future work.

The repetitive use of a conjunction as a figure of speech is referred to as polysyndeton. In the literary register, it is often used deliberately to convey swiftness in the succession of events. However, in Medieval texts that does not seem to be the case. The abundance of *e* has

¹⁸ See Fraser (1998; 2009) for more on discourse connectors, and Cuenca (2008, sec. 31.2.2.1) for usages of *i* in MCat beyond coordination in a purely additive sense.

often been associated to a supposedly ‘paratactic’ stage of Old Romance languages (Martí i Castell, 2002). While I do not think that that is the case in the LFRJ, I consider that the pervasive use of this conjunction is linked to the oral nature attributed to them (not in vain, it has been argued that the text was dictated by the King himself by adducing its orality).¹⁹

In M_{Cat}, polysyndeton is associated with oral discourse and children speech production.²⁰ Within oral texts, a factor that contributes to a high frequency of ‘i’ is the linear narrative nature of the text,²¹ and the absence of secondary plot lines: all clauses are anchored on the same common ground to which new referents and new information are added for the narrative to advance. In Serra and Prunyoza (2008), the capacity of M_{Cat}’s *i* ‘and’ to concatenate independent clauses that do not seem coordinated in an additive sense is listed as a property of the coordinating conjunction. This function resonates with the discourse cohesion marker proposal that we have made for O_{Cat}’s *e*.

Summarising the findings so far, it can be stated that (i) in O_{Cat}, the coordinating conjunction *e* is an element located above CP not linked to rendering V1 clauses grammatical within a V2 grammar; that (ii) it can act as a cohesion marker, and not just a copulative coordinating conjunction with and additive meaning; that (iii) the use of *e* in linear narrations as a discourse cohesion marker is cross-linguistically attested and not dependant on the language’s grammar. With this evidence in hand, it can be established that *e*V clauses are truly verb initial, unmarked clauses.

3.1.2.2 *Verb initial clauses, predicate types and Information Structure*

In the previous section, it has been established that O_{Cat} *e*V clauses are true V1 clauses that do not involve the presence of a null topic operator, but the presence of a discourse continuity marker. In this section, we explore variables that have crosslinguistically been linked to VS

¹⁹ The oral nature of *El Llibre dels Feits* is uncontested, and is one of the key arguments used in favour of King James I authorship (Bruguera, 1991, 2012; Casanova, 2012; Ferrando Francès, 2012). Similar characterisations have been made for contemporary O_{Sp} texts (Fernández-Ordóñez, 2008).

²⁰ While we have no specific data on this matter, 4/5 speakers consulted for the grammaticality of (15) and (17) commented on it sounding “child-like” due to the high frequency of ‘i’. A quick browse through the *Corpus de Català Contemporani de la Universitat de Barcelona* confirmed that it is indeed the case in oral narrations directed to children: <http://hdl.handle.net/2445/11603>, where the story is linear and there is topic continuity.

²¹ Berman (2015, p. 461) identifies the use of clause initial *and* for the “temporal chaining of clauses in sequence” in narratives as the second stage of acquisition of the coordinating conjunction. This is found cross-linguistically in many languages (French, Hebrew, German and English are among the languages cited in Berman [2015] with different basic word order patterns).

orders: Information Structure (henceforth IS) and the argumental structure of certain predicate types.

Syntactically, the tendency for certain predicate types to occur in VS configurations has been explained by calling upon null elements: V1 orders with unaccusative predicates (and other intransitive predicates) have been related to the presence of a null locative (Borer [1980] for Hebrew, Devine & Stephens [2006] for Latin, Petrova & Hinterhölzl [2010] for Old High German, Belletti [1988, 1999] and Tortora [2001] for MIIt, Alexiadou et al. [2003] for unaccusative verbs in general), and null narrative operator has been postulated for *verba dicendi* (Rodríguez Molina, 2014; Willis, 1997; Wolfe, 2015c).

Table 3 shows the frequency of predicate types in V1 clauses from the LFRJ database:

Verb type	Total number of occurrences	%/ total V1
Transitive	179	38
<i>Verba dicendi</i>	117	25
Unaccusative	109	23
Copula	28	6
Unergative	22	4.7
Reflexive	9	2
Passive	6	1.3
Total	470	100

Table 3 – Occurrence of specific predicates in V1 clauses

Transitive verbs, not associated to V1 orders, present the highest frequency of V1 in the LFRJ database, followed by *verba dicendi* and unaccusative verbs. The fact that V1 clauses do not exclusively belong to *verba dicendi* or unaccusative predicates contrasts with the distribution of V1 clauses in uncontroversial V2 languages, as noted by several authors (Axel, 2007; Petrova & Hinterhölzl, 2010; Sigurdsson, 1990; Thräinsson, 2007).

In this section, we focus on two predicate types crosslinguistically prone to occur in VS sentences which have been claimed to fit within a V2 grammar in the Old Romance languages: *verba dicendi* and unaccusative predicates. It is shown that (i) the use of *verba dicendi* in the LFRJ database is parallel to the use of *verba dicendi* in MCat, and that (ii) the behaviour of unaccusative predicates is parallel to that of unaccusative predicates in the Modern Romance languages.

Verba dicendi

Verba dicendi, or ‘verbs of saying’ are transitive verbs that can take full CPs that constitute a separate proposition as their complement. In Table 3 we saw that 117 (25%) of V1 clauses contain a *verbum dicendi*. Table 4 shows that, while there is an overwhelming tendency for *verba dicendi* to occur in V1 clauses, they can also occur in V2 and V3 sentences.

<i>Verba dicendi</i> in...	Number of occurrences	%/total <i>verba dicendi</i>
V1	117	72.3
V2	37	22.9
V3	7	4.3
Total	161	100

Table 4 – Distribution of *verba dicendi* in the LFRJ database

Wolfe (2015b, p. 24), following Zwart (1997), proposes the existence of a null discourse operator, a variant of *pro*_{Top}, that would trigger the attested VS order by the same inversion mechanism that is discussed in section 3.1.2.1. Rodríguez Molina (2010) makes a similar claim for OSp.

In OCat while VS orders with *verba dicendi* are the most common, SV orders are also possible:

- (24) Context: the King and his loyal men are skirmishing against Sir Pero Ahonés and his men. After following him and his men for a while, Sir Pero Ahonés’ horse becomes tired. They climb up a hill to do a horse change. While they do it, Ahonés’ men start throwing stones downhill, so that the king and his men would not be able to climb up. The king knows an alternative way to access the hill, and he tells his men.

E Don Pero Ahonés mudà·s en aquel cavayl.
and Sir Pero Ahonés changed.3SG=himself.CL on that horse

E nós **dixem** a Don Assalit e a Don Domingo Lópeç de Pomar
and we said.1PL to Sir Assalit and to Sir Domingo Lópeç de Pomar

que per ·i_a· pujada que y havia podien venir
that for one slope that there.CL=had.3SG could.3PL come.INF

là on éls eren.

there where they were.3PL

And Sir Pero Ahonés changed horse. And we said to Sir Assalit and to Sir Domingo Lópeç de Pomar (...) that because there was only one sharp bend there, that they could go to where they were.

Fol. 15r, l 28

- (25) Context: the King is talking to several knights, among which, Guillem de Cardona.

E **dix** En Guillem de Cardona: “Séyer, fêts -me
 and said.3SG Sir Guillem of Cardona Lord make.2PL =me.CL
 guiar, e anar -me· n hé”. “E no y
 guide.INF and go.INF =REFL.1SG.CL=ADV.CL=have.1SG and not there.CL=
 fariets àls? - **dixem** nós. E él **dix** que no.
 do.2PL.COND other said.1PL we and he said.3SG that no

And Guillem de Cardona said: ‘Sir, get someone to guide me and I will be gone’.
 ‘And wouldn’t you like to do something else?’, we said. And he said that he would not.

Fol. 22r, l 25

- (26) Ab aytant **respòs** -los en Guillem de Montpestler
 with much answered.3SG =them.CL Sir Guillem de Montpellier
e son conseyl que d’altra manera no seria.
 and his council that of;other way not be.3SG.COND

And in the meanwhile, Sir Guillem of Montpellier and his council replied to them that it would not happen in any other way.

Fol. 2v, l 14

- (27) E **dix** ela que tot ço faria (...)
 and said.3SG she that all this do.3SG.COND

And she said that she would do all of this.

Fol. 34v, l 12

- (28) E En Pere Martel **dix**-los que·ls diria
 and Sir Pere Martel said.3SG=them.CL that=to;them.CL say.3SG.COND

noves (...)

news

And Sir Pere Martel told them that he would give them news.

Fol. 27r, v2

Examples (24) and (25) have several instantiations of *verba dicendi*, all with overt subjects that have already been introduced in the discourse and that are overtly expressed to signal topic shift. Nevertheless, even though the discourse value of all subjects is the same, their position with respect to the verb is not. In (24), the pronoun *nós* ‘we’, which shifts the topic back to King James I, occurs preverbally. However, in (25), we see that it is also possible for topic shifting subjects to occur postverbally. First, the topic shifts to *En Guillem de Cardona*, a postverbal subject. After Sir Guillem’s intervention, the King replies, and the postverbal pronoun *nós* ‘we’ indicates topic shift. Sir Guillem answers the King’s reply, and this time, the topic shifting pronoun *él*, ‘he’, occurs preverbally, as in (24). It is worth noting that the position of the reported complement clause does not interact with the position of the subject. This is apparent in (25), where the first clause containing a *verbum dicendi* presents the order VSO, while in the following sentence, the order is OVS. This is not the case with reported speech complement clauses, which consistently occur postverbally, in VSO or SVO clauses, as in (26) and (27). In (28), we can see that VSO orders do not necessarily require the presence of a preverbal constituent. No VOS clauses are found with *verba dicendi*.²²

SV(O) clauses are often bipartite at IS level, where the subject provides an aboutness topic and the element occurring in the rightmost position of the clause receives Information Focus (henceforth IFoc) and coincides with sentential stress (Zubizarreta, 1998). Clauses with bipartite IS are often referred to as *categorical*, while clauses in where no such division exists and the sentence is treated as a single informational unit grounded in the discourse with an all-focus reading are referred to as *thetic* and are associated to different word order patterns, among which, verb initial clauses (VS orders). Apart from IS, other factors are known to interact with word order at sentence level, notably the predicate’s argumental structure, its lexical aspect and the thematic prominence of arguments. While these factors will be shown to be especially relevant in relation to unaccusative predicates, Leonetti (2017, p. 896) proposes that VS orders with *verba dicendi*, also known as *quotative inversion*, should be interpreted as instances in

²² VOS orders are only found in collocations with light predicates and their direct objects. This word order pattern is discussed in length in Chapter 4.

which the subject is located postverbally to receive focal reading and to coincide with sentential stress.

As it can be seen in examples (25-27), there is no need for the DO to be fronted for inversion to take place, ruling out that the trigger for quotative inversion is not movement of the DO to the left periphery, as is the case with Focus Fronting (see Chapter 5, section 5.2.4). Instead, along the lines of Leonetti (2017, p. 896), we propose that quotative inversion endows the clause with athetic reading, whereby it is treated as a single informational unit that is added to the conversation. OVS is only possible when the DO is direct speech and is separated from the VS string by a prosodic pause in M_{Cat} (no O_{Cat} data is available, but since this holds crosslinguistically for languages presenting quotative inversion, we assume that that was also the case in O_{Cat}). VSO orders can either correspond to indirect speech, where the object takes the form of an embedded clause or to direct speech, where the object is quoted text. In the case of the former, the object occurs in situ and is prosodically independent (Steuk, 2016), while in the latter, there is a pause between the VS string and the object, also prosodically independent. Given the prosodic independence of DOs, the object does not compete with the subject for the prosodically stressed position within the core of the clause.

Therefore, the difference between SVO and VSO sentences with *verba dicendi* lies on the fact that the former are categorical, while the latter receive athetic reading. Thetic clauses need to be anchored in the discourse. In categorical clauses, it is the aboutness topic that anchors the clause, which tends to coincide with the subject. In thetic clauses, the clause is anchored in the discourse by a topic that corresponds to the spatio-temporal context in which the action takes place, the ‘here and now’, very much like in existential or unaccusative clauses (see Chapter 4, section 4.3.2 for existential predicates, the following section for stage topics and unaccusative predicates, as well as Bentley et al. [2015, p. 59] and Erteschik-Shir [1997, p. 26-29] for more on this concept). In clauses where a stage topic is present, the whole all-focus thetic clause is predicated of the stage topic (Erteschik-Shir, 1997, p. 27). In the case of *verba dicendi*, the stage topic can be identified within discourse (i.e. they are discourse-specified).

Since quotative inversion in the LFRJ database is epiphenomenal (i.e. it is the result of athetic word order pattern rather than a phenomenon of its own right) and not linked to fronting, we take the verb in these clauses to be in TP, like in unmarked declarative clauses. The verb being in TP, we take the subject to be either in its base-generated position or in a position within the low left periphery (Belletti, 2001, 2004). This could be the case given that O_{Cat} did have an active low left periphery accessible to subjects, as it is discussed in Chapter 4, section 4.3.

In sum, these examples show that, while *verba dicendi* tend to present postverbal subjects (when overt), they can also have preverbal ones, with the same discourse value as postverbal ones, but the differences in their position correspond to different word order patterns associated with different sentence IS values.

Unaccusatives

Modern Romance unaccusative verbs have a tendency to present postverbal subjects (Hulk & Pollock, 2001; Pinto, 1994; Sheehan, 2010; Tortora, 2001). It is also the case of unaccusative verbs in OCat. This is shown in Table 5, where the proportion of unaccusative predicates across the LFRJ database with overt preverbal and postverbal subjects is shown:

		In LFRJ database	In V1 clauses
Non pro-drop clauses with unaccusative predicates	Preverbal subjects	43/133 (32.3%)	0/35
	Postverbal subjects	90/133 (67.7%)	35/35 (100%)
Total non pro-drop clauses with unaccusative verbs	-	133/243 (54.7%)	35/109 (32.1%)
Pro-drop clauses with unaccusative verbs	-	110/243 (45.3%)	74/109 (67.9.5%)
Total clauses with unaccusatives verbs	-	243	109/243 (44.8%)

Table 5 – Overt subject position in clauses with unaccusative predicates in the LFRJ database

While unaccusative clauses occur predominantly in V1 clauses (44.8% of overall occurrences of unaccusative predicates appear in V1 clauses), unaccusative predicates occurring in other word order configurations also tend to display postverbal subjects (67.7% of cases). Therefore, the VS configuration is not uniquely associated to the verb being in clause initial position. In Modern Romance languages the possibility for unaccusative verbs to present preverbal as well as postverbal subjects is referred to as free inversion (Hulk & Pollock, 2001). This phenomenon has been linked to information structure (Pinto, 1998; Leonetti 2017), as well

as to the presence of a null locative in SpecTP (Borer, 2009; Sifaki, 2003; Tortora, 2001). Here, I consider OCat data on the light of these analyses, and I establish that it is possible to account for the tendency of unaccusative predicates to present postverbal subjects by postulating the presence of a null locative, as it has been suggested for Modern Romance.

As we have mentioned above, beyond IS, the argumental structure of predicates, their lexical aspect and the thematic prominence of arguments interact with word order. As noted by Leonetti (2017, p. 893) unaccusative verbs favour thetic readings, partly due to their argument structure, partly due to their lexical aspect. In terms of argument structure, unaccusative verbs are monoargumental: the subject is their only argument, and it is internal. Monoargumental verbs, and especially unaccusative ones, favour VS orders. In regard to aspect, eventive predicates tend to favour the integration of arguments into the predicate, without partitions (Leonetti, 2017, p. 893), contrary to stative predicates, that tend to require a partition between old and new information. Most unaccusative verbs are eventive predicates. In terms of argumental prominence, prominent roles such as agent favour informational partition, while non-prominent roles such as theme or patient do not. This is the case of unaccusative verbs' subjects, which are assigned the theme role. This combination of factors makes unaccusative predicates prone to thetic VS orders with no informational partition, even though this does not preclude the possibility for them to occur in SV clauses with categorial reading. As mentioned above, clauses receiving an all-focus thetic reading need to be grounded in the discourse and possess a null stage-topic.

Tortora (2001) associates the alternation between preverbal and postverbal overt subjects in Italian unaccusative verbs to the presence or absence of a *pro*_{LOC} element (a locative equivalent of *pro* that would occupy the same position within SpecTP that can be identified with the stage-topic mentioned above) and would act as a stage-topic of sentences that receive an all-focus thetic reading (Pinto 1997). According to her, in clauses such as (29.a) in MCat, the presence of *pro*_{LOC} accounts for the postverbal subject. *pro*_{LOC}'s interpretation is deictic: it makes reference to the context in which the speaker utters the clause, to 'here'. The contrast between VS and SV orders and their relation to 'here' can be clearly seen in the MCat examples in (29). In (29.b), *Met* cannot arrive 'here', he has to arrive elsewhere, while in (29.c) shows that VS orders yield ungrammatical results when the place to which *Met* arrives is not coreferent with the context in which the clause is uttered.

- (29) a. **Arriba** en Met (aquí).
 arrives the Met (here)

Met arrives (somewhere).

- b. En Met **arriba** a l'Antàrtida.
the Met arrives at the;Antarctica.

Met arrives to Antarctica.

- c. ** Arriba en Met a l'Antàrtida.
arrives the Met at the;Antartica

Met arrives to Antartica.

MCat

A further consideration to be made about the distribution of subjects with unaccusative predicates, and especially those that involve motion, is that the grammaticality of VS orders improves if a temporal deictic adverb is added preverbally (in its base generated position, not in the left periphery):

- (30) Ara arriba en Met.
now arrives the Met

Met is arriving now (here)

- (31) En Met ara arriba [a l'Antàrtida].
the Met now arrives to the;Antarctica

Met is now arriving in Antarctica.

MCat

In (30), a VS clause, *Met* can only arrive 'here', where the speaker is located (it could be elsewhere, it could be a place referred to before in a narrative, for instance in live reporting). In contrast, in (31), an SVO clause with the adverb *ara* 'now' occurring in its based-generated position, *Met* can only arrive to a location other than the speaker's. This structure could be analysed as Focus Fronting (see section 5.2 in Chapter 5 for an extensive description of this phenomenon in OCat as well as MCat), by which *ara* 'now', would be contrasted with all the other possible moments when *Met* could arrive. Nevertheless, the intonation contour of (30) with *ara* 'now' being Focus Fronted contrasts heavily with that of (30) with *ara* in its based-generated projection, as a TP adverb. In the latter case, *ara* could be argued to anchor the clause in the discourse, in a way similar to *pro*_{LOC}. The possibility of having temporal as well as locative expressions as topic goes in line with Leonetti's (2017, p. 891) statement that in the

absence of a syntactic topic, temporal and locative expressions can play the same role in grounding the sentence into the context. This holds for unaccusative predicates that involve motion, but it does not seem to be the case for other unaccusative predicates such as *néixer* ‘to be born’ or *morir* ‘to die’, suggesting that the relation between the adverb and the predicate might also have a variable linked to aspect, specifically, perfectivity, as already pointed out in Sifaki & Tsoulas (2016) for Modern Greek. The relation of these adverbs and VS orders with unaccusative predicates are left for future research. Nevertheless, it is worth noting that cases such as (30) might provide further evidence for SpecTP having discourse grounding properties, as it is suggested from OCat evidence presented in what follows.

This being the case, we are presented with a further piece of evidence in favour of a discourse anchoring analysis of SpecTP, which is further explored in Chapter 4, section 4.3.2.

If we accept that the existence of *pro*_{LOC} is plausible for OCat unaccusative verbs and that Tortora’s analysis for VS orders with unaccusative predicates in MIt is valid for OCat, it could be hypothesised that in the absence of an overt expression of location, a null element occurs in SpecTP and prevents the raising of the subject to this position. The possibility of the null locative raising to a position in the left periphery, as it would be expected within a V2 analysis, is not considered due to the fact that, above, it has already been established that the behaviour of OCat V1 clauses is parallel to that of MCat, an SVO language. Further arguments for a non-V2 analysis of OCat are presented within this Chapter.

3.1.2.3 Absolute V1 clauses

Let us now focus on the remainder 6% of V1 clauses, not preceded by *e*. Wolfe (2015c) states that absolute V1 can be explained within a V2 grammar either by all-focus or by quotative inversion. Here, we show that *e*-less V1 clauses in the LFRJ database cannot be accounted for using only information structure or quotative inversion.

In the LFRJ database, as shown in Table 2, there are 27 cases of *e*-less verb initial clauses. Only 2/27 (7.4%) cases of *e*-less V1 clauses contain *verba dicendi* and could be explained by postulating the presence of a covert narrative operator. This, together with the fact that *verba dicendi* tend to present postverbal subjects in clauses with athetic reading (see section 3.1.2.2), leads us to dismiss the presence of a narrative operator to account for *e*-less V1 clauses.

Nevertheless, the scarcity of absolute verb initial clauses suggests that this word order pattern is marked, and thus, subject to distributional constraints that we describe in what follows. Table 6 summarises the structure and distribution of *e*-less V1 clauses:

Non pro-drop <i>e</i>-less V1 clauses	Word order pattern		Reported Speech	Narrative body
	VSO	2	1	-
	VOS	1	-	1
	VS	2	-	2
	V S XP	1	-	1
Total		6/27 (22.1%)	1/6 (16.7%)	5/6 (83.3%)
Pro-drop <i>e</i>-less V1 clauses	Word order pattern		Reported Speech	Narrative body
	V O	11	9	2
	V ADV O	2	1	1
	V PP	4	4	-
	V ADV	2	1	1
	V	2	1	1
Total		21/27 (77.7%)	16/21 (76%)	5/21 (24%)

Table 6 – Structure and distribution of *e*-less V1 clauses

Table 6 shows the different word order patterns attested in pro-drop and non pro-drop *e*-less V1 clauses. Non pro-drop clauses mostly display verb-subject adjacency (5/6), with only one case where the DO intervenes between the verb and the subject, reproduced in (32):

- (32) **ach** nom la ·i·a l' infant Don Fferrando (...)
 has name the one the= prince Sir Fernando
 One has the name of prince Fernando (...)

Fol. 9r, l 11

In (32), the light verb *haver* ‘to have’ incorporates the DO *nom* ‘noun’ is pseudo-incorporated to the verb forming a compound predicate that moves to TP, and thus, the subject, remaining in its base-generated position, cannot intervene between the verb and the

DO (Borik & Gehrke, 2016, p. 11; Forcadell, 2013, p. 57 for M_{Cat}), as it happens in transitive clauses with no incorporation:

- (33) “**Demanam**-vos de conseyl (...)”.
 ask.1PL= to;you.CL PART.CL= advice
 ‘We ask you for advice (...)’

Fol. 26v, l 26

- (34) **Retrau** mon seyor sent Jacme que fe sens obres
 states my lord saint James that faith without deeds
 morta és.
 dead is

My lord Saint James states that faith without deeds dies.

Fol. 1r, l 1

- (35) E, tenguda la festa, **demanà**· ns lo
 and held the party asked.3SG= us.CL the
dit rey de Castella de conseyl (...)
 aforementioned king of Castile of advice

And, after the party, the king of Castile asked us for advice (...).

Fol. 26v, l 25

In (33) we have an example of an *e*-less V1 clause where the object follows the verb, case marked for partitive by partitive marker *de*. Unfortunately, all instances of *e*-less VSO clauses in the LFRJ database present heavy objects, and therefore, it is expected for them to occur in sentence-last position, as in (34). However, VSO orders can be found elsewhere in the text, as in (35).

Leonetti (2017, p. 901-905) describes VSO as a word order pattern that precludes internal information partitions and gives rise tothetic readings. This is exactly what we find in O_{Cat}: *e*-less VSO clauses receive an all-focusthetic reading. This explains the distribution of this word order pattern, found in the opening line of the text (see (34) above) and in reported speech and in clauses in the narrative body with existential and unaccusative predicates. In reported speech, new content is introduced into the conversation’s context by the characters, as in (33). Existential and unaccusative predicates, as we have discussed above, favour VS

orders andthetic readings, so it is not unsurprising to find them in *e*-less V1 clauses, as in (36-38):

- (36) “Seyor, **prega-us** la comtessa que la escoltets (...)”
 Sir begs=you.CL the countess that her listen.2PL.SBJV
 Sir, the countess begs you that you listen to her (...).”

Fol. 24v, 1 21

- (37) E sí y havia bon pujador e qui fer-ho
 and yes there.CL had.3SG good climber and who do.INF-it.CL
 volgués. **Havia** ·i· escuder, lo nom del
 wanted.3SG.PST.SBJV had.3SG 1 squire the name of;the
 qual a nós no membre, (...)
 which to us not remember.INF

And there was indeed a good climber and someone who wanted to do it. There was one squire, whose name I do not remember (...).

Fol. 7v, 1 25

- (38) (...) e hac-se vestit ·i· gonió e ·i· capel
 and has=himself.CL dressed 1 chainmail and 1 hat
 de ferre en lo cap e l’ espaa en la
 of iron on the head and the sword in the
 mà; e veé que la batayla anava cessan, **moch**
 hand and saw.3SG that the battle went.3SG ceasing moved.3SG
 -se tant con los peus lo pogren levar,
 =himself.CL much as the feet him could.3PL bring.INF
 e començà a pujar (...).
 and started.3SG to climb.INF

And he dressed himself with a chain mail, and an iron hat on the head, and the sword in hand, and he saw that the battle was calming down, and he moved as much as his feet allowed him, and he started to climb.

Fol. 7v, 1 28

Therefore, *e*-less V1 clauses, which could also simply be described as VS(O) clauses) receive an all-focusthetic reading and are anchored in the ‘here and now’, an unspecified Stage Topic. Above, we have agreed with Tortora’s (2001) postulation of the existence of *proLOC*, which could be taken to be the syntactic realisation of the Stage Topic in VS clauses, whose referent depends either on the predicate of the clause (some unaccusative predicates deictically pointing at a location inherently), or on the relation of the clause with respect to the preceding discourse. In SV(O) clauses, it tends to be the subject that grounds the clause in the preceding discourse. We take *e*-less V(O) clauses with null subjects as instances of SVO clauses with *pro*, inherently referential and discourse grounding, in SpecTP.

Having analysed *e* ‘and’ as a discourse cohesion marker, it is unsurprising that VS(O) all-focusthetic clauses are not preceded by it, given that they built on a Stage Topic that is not coreferential with the preceding discourse. Instead, they establish a new background against which new information can be added. Nevertheless, it is worth pointing out that the presence of *e* ‘and’ is not needed to establish discourse topic continuity. This is apparent in (39), where an *e*-less V1 clause builds on the preceding discourse:

- (39) Dixem-li: “En Guillem de Cardona, vós no havets
said-PL=to;him Sir Guillem of Cardona you not have.2PL
aduyta aquí procuració neguna d’ En Guerau; l’ altre
brought here capacity any of;Sir Guerau the other
vós no volets respondre a dret. **volem** saber
you not want.2PL answer.INF righteously want.1PL know.INF
encara vós si volets respondre a la demanda
furthermore you if want.2PL answer.INF to the request
que· N Guillem Sasala vos fa”.
that Sir Guillem Sasala to;you.CL= does

We said to him: ‘Sir Guillem of Cardona, you have not brought any document here proving your capacity to act on behalf of Sir Guerau, the other one, you have not wanted to reply righteously. We still want to know if you want to respond to the request that Sir Guillem Sasala does to you’.

Fol. 21v, l 22

According to Leonetti (2017), the Modern Romance languages pattern in two groups: MSp, Modern Portuguese and Modern Romanian allow for the VSO orders, while MCat,

MFr, MIIt and MSard (the *central Romance languages*, as he refers to them), do not. OCat, therefore, patterns with *peripheral Romance languages* instead of doing so with its modern counterpart, in allowing for the production of VSO. This is the first of many instances that are explored in this thesis were OCat patterns with MSp (together with the expression of emphatic positive polarity and the availability of focus fronting). The fact that Catalan has lost the possibility of producing such orders in favour of other syntactic strategies to convey theticity (mainly SVO with transitive predicates, VS being reserved for certain unaccusative and unergative cases) shows its innovative nature when compared to other peninsular varieties and its alignment with central Romance varieties. The scarcity of VS in non-predicate determined contexts suggests that this structure was already weakened by the 13th century, when *El Llibre dels Fets* was produced, and that it was reserved for highly rhematic contexts, such as the opening line of the text or direct reported speech.

SVO is also found in all-focus clauses, as it is the case in MCat:

- (40) Ara **comptarem** en qual manera nós fom engenrats
now tell.1PL.FUT in which manner we were begotten
- e en qual manera fo lo nostre neximent. Primerament
and in which manner was.3SG the our birth firstly
- en qual manera fom engenrats nós. Nostre pare, lo
in which way were.1PL begotten we our father the
- rey En Pere, no **volia** veser nostra mare, la reyna.
king Sir Pere not wanted.3SG see.INF our mother the queen
- Now we will tell how we were begotten and how our birth took place. First, how we were begotten. Our father, King Peter, did not want to see our mother, the queen.

Fol. 3v, 12

In (40), the narrator announces a new section how he was conceived and his birth. The first sentence of the new section introduces two new characters, his parents. They are introduced in an SVO clause, where both characters are presented with specifying appositions. While one can assume that the referents of the parents of King James' I would be more or less active in the audience's/readership's common ground, so would Saint James' quote from the opening line of the text. This shows that SVO all-focus clauses were already

possible in OCat, and that they were competing with VSO for being associated with this informational make-up.

On a final note, it has been proposed in the literature that Old Ibero-Romance had a strong Polarity head that attracted the verb to the left periphery. Fischer (2002) uses this hypothesis to explain the oscillation between V-CL orders in OCat, and M Batllori & Hernanz (2008); Batllori & Hernanz (2013); Martins (2013); Rodríguez Molina, (2014) to account for the existence of verb-echo and do support in OSp. The interaction between verb position and polarity is explored in OCat in section 3.1.6 within this Chapter.

Wrapping up the data presented up to this point, it has been shown that *e*-less V1 clauses cannot be explained only by means of a narrative operator, since only two of the 37 cases found in the LFRJ database contain a *verbum dicendi*. Nor can they be explained by associating this word order pattern with all-focus, since in most instances, there is undoubted discourse topic continuity, as well as topic subjects (marking either topic continuity or topic shift). The fact that *e*-less V1 clauses are compatible with topic continuity shows that *e* is not a requirement to convey discourse topic continuity and reinforces the findings of section 3.1.3.1.

3.1.2.4 Summary

In this section we have explored OCat verb initial clauses, showing that they are unmarked clauses. It has been shown that the high frequency of *eV* in the LFRJ data does not respond to a mechanism to fulfil a V2 requirement, based on the fact that (i) *e* is located in an extra-clausal position above CP, and (ii) that it appears not only preceding V1 clauses, where it could be said to fulfil the V2 requirement, but also 88.33% of the total of clauses from the LFRJ. We have also explored the value of *e* and its potential entry in the lexicon as a discourse cohesion marker. The relation between V1 orders and certain predicate types has also been explored. In the case of *verba dicendi*, the presence of a narrative operator triggering subject-verb inversion has been dismissed in favour of athetic analysis of such closes, while in the case of unaccusative predicates, the presence of a null locative element in SpecTP has been linked to the high frequency of VS orders with this predicate type. Finally, it has been shown that while all-focus clauses may be absolute verb-initial clauses, they but need not to.

3.1.3 Verb third clauses

V3 clauses are the least frequent in the database, amounting to 8.5% of the total (85/1000). Nevertheless, their value lies in the fact that they can tell us about the information structure of the left periphery of the clause, and the several positions available to subjects (information structure and the possible combinations of constituents in the left periphery is explored in Chapter 5, while subject position is described in Chapter 4). Table 6 shows the different combinations of elements in the left periphery of V3 < in non-pro-drop and pro-drop clauses respectively.

Structure	XP1	XP2	XP3	XP4	V	XP1	Occurrences
	Adv	Adv				S _{LEX}	3
Non Pro-Drop Total: 292 66.5%	Adv	PP					3
	Adv	S _{LEX}					6
	Adv	S _{PRO}					16
	Adv	Adv	S _{PRO}				1
	Adv	PP	S _{LEX}				1
	Adv	PP	Adv				1
	Adv	S _{LEX}	Adv				1
	Adv	S _{PRO}	Adv				1
	Adv	DO	PP	S _{LEX}			1
	DO	S _{PRON}					1
	DO	S _{LEX}					1
	PP	S _{PRON}					8
	PP	S _{PRON}	Adv				1
	PP	S _{LEX}					3
	PP	S _{LEX}	Adv				1
	PP	Adv				S _{LEX}	3
	S _{LEX}	Adv					3
	S _{PRON}	Adv					7
	S _{PRON}	PP					3
	Pro-Drop 147 33.5%	ADV	PP	ADV			
ADV		PP					6
PP		ADV					5
Total							5

Table 6 – Preverbal constituents in V3 main clauses

As shown in Table 6, a wide range of constituents, parallel to those available preverbally in V2 clauses, can co-occur in preverbal position forming V3< clauses. This is illustrated in (41) and (42). In (41), a lexical subject and a temporal subordinate clause precede the verb. In (42), a prepositional phrase and a pronominal subject precede the verb.

- (41) E [·i· bisbe e ·ii· richs hòmens qui venien ab ella]₁, [quan
and 1 bishop and 2 rich men who came.3PL with her when

foren a Montpestler]₂, **saberen** que·1 rey Don Alfonso, nostre
were.3PL to Montpellier knew.3PL that;the king Sir Alfonso our

avi, havia presa la reyna Dona Sanxa, fiyla de
grandfather had.3SG taken the queen Lady Sanxa daughter of

l'emperador de Castella, per muyler.
the;emperor of Castile for wife

And one bishop and two rich men who came with her, when they arrived to Montpellier, they realised that king Alphonse, our grandfather, had taken as wife queen Sanxa, daughter of the emperor of Castile.

Fol. 2r, 18

- (42) [Sobre açò]₁, [él]₂ **féu** resposta al bisbe e
about this he made.3SG answer to.the bishop and

als nobles qui eren venguts.
to.the noble men who were.3PL come.PTCP.MSC.PL

And regarding this matter, he answered the bishop and the noblemen that had come.

Fol. 2v, 12

As shown in Table 6, V3 clauses have less postverbal subjects than V2 clauses: while in V2 clauses they amount to 25.6%, in V3 clauses their incidence is of 14.1%. Moreover, there is not a single case of a pronominal postverbal subject in V3 clauses: only lexical postverbal subjects are found, and they need not to be adjacent to the verb.

The purpose of this section was to describe the different possible configurations available to V3 clauses in OCat. OCat's left periphery is extensively described in Chapter 5.

3.1.4 Adverb Position

Cinque (1999) proposed a universal adverbial hierarchy, postulating the existence of fixed Tense, Mood and Aspect heads throughout the clause, each being able to host adverbs expressing related semantic notions. He distinguishes three types of adverbs according to their location within the clause: CP, TP and VP adverbs. Thus, if the verb raises from VP to TP, it appears above VP adverbs but below high TP ones. If it raises from TP to CP, then it appears above VP and TP adverbs, but below CP ones. The following examples contain adverbs that can be located within Cinque's hierarchy, helping us determine the position of the verb within the clause. MCat translations of the OCat examples are provided so that both stages of the language can be compared. In all examples, the verb appears in bold and the relevant adverb is underlined:

VP Adverbs

Voice

- (43) E no us **porem** servir bé.
and not to;you.CL= can.1PL serve.INF well

And we will not be able to render you good service.

Fol. 29v, l 4

- (44) I no us **podrem** servir bé.
and not to;you.CL= can.1PL serve.INF well

And we will not be able to render you good service.

MCat

TP adverbs

Tense – Future (i.e. then)

- (45) E puys **anam-nos-en** rebre l'altra
and then went.1PL-REFLX.1PL.CL=ADV.CL receive.INF the;other

partida de la host de Barcelona.

part of the army of Barcelona

And afterwards we went to receive the other part of Barcelona's army.

- (46) I llavors nos n' **anàrem** a rebre l'altre
 and then REFLX.1PL.CL=ADV.CL=went.1PL to receive.INF the;other
 part de l'exèrcit de Barcelona.
 part of the;army of Barcelona

And afterwards we went to receive the other part of Barcelona's army.

MCat

- (47) E puys **levà's** lo probost de Tarregona.
 and then stood=REFL.CLthe mayor of Tarragona

And then the mayor of Tarragona stood up.

Fol. 36v, 116

- (48) I llavors s' **alçà** el prebost de Tarragona.
 And then REFL.CL=stood up.3SG the mayor of Tarragona.

And then the mayor of Tarragona stood up.

MCat

Tense – Anterior (i.e. 'already')

- (49) E ja **vench** la ora del vespre.
 and already came.3SG the hour of;the evening

And it was already evening.

Fol. 33r, 19

- (50) I ja **era** el vespre.
 and already was.3SG the evening

And it was already evening.

MCat

Aspect – Proximative (i.e. 'soon')

- (51) E sempre **faem** armar bé ·l· cavalers.
 and soon made.1PL arm.INF a good 50 knights

And soon after that, we had 50 knights armed.

- (52) I aviat **férem** armar una bona cinquantena
 and soon made.1PL arm.INF a good 50
 de cavallers.
 of knights

And soon after that, we had 50 knights armed.

MCat

These examples show that the verb is located above VP adverbs and below TP adverbs, in TP. It could be argued for preverbal adverbs to be located in the left periphery. Nevertheless, adverbs in these examples receive a neutral reading and are not discourse binding (i.e. deictically pointing at the wider discourse, located in the Frame field, Benincà [2004]), and therefore it can be safely assumed that they are located in their base generated projection corresponding to Cinque's hierarchy.

Thus, OCat contrasts with Old Sardinian (henceforth OSard), a language in which the verb allegedly raises to the CP layer (Wolfe, 2014). (53) contains an OSard sentence where the adverb *osca* 'then', equivalent to the OCat *puyis* in (45) and (47) appears postverbally.

- (53) **Bennit** osca Ithocchor Manutha
 came.3SG then Ithocchor Manutha
 'Ithocchor Manutha came then'.

Il Condaghe di San Nicola di Trullas, 323, *apud* Wolfe (2014, p. 5)

OSard

In this section, it has been established that adverb position provides unambiguous evidence about the position of the verb in OCat: the verb remained in a position lower than the CP layer, within TP. Further evidence from MCat has been used to show that the behaviour of the Old language patterns with that of the Modern one. Together with clitic placement, which is explored in section 3.1.5, this provides us with a further piece of evidence for a non-V2 analysis of OCat.

3.1.5 Clitic Placement

Romance object and adverbial clitic pronouns are syntactically independent but phonologically dependent elements; that is, they need a phonological host with which they

form a phonological word. In this section, we present the different configurations in which clitics appear in main clauses from the LFRJ database; we summarise the main proposals made to describe the distribution of OCat object and adverbial clitic pronouns, to which we refer in this section as ‘clitics’; and we contrast them against the data from the LFRJ database. Finally, clitic placement is shown to be determined independently to verb movement.

The distribution of preverbal and postverbal clitics in Old Romance was first described by Adolf Tobler and Adolfo Mussafia at the end of the 19th century (Mussafia, 1888; Tobler, 1875) and is now referred to as the Tobler-Mussafia Law (henceforth TML). The TML states that in the Old Romance languages, enclisis is found in verb initial clauses, while proclisis only occurs when there is at least one constituent preceding the verb. This observation has been widely refined by extensive research in Old Romance clitic placement that has ensued during the 20th and 21st century. The two main studies of clitic placement in OCat known to me are Fischer (2002) and Batllori et al. (2005). Fischer (2002) proposes that variation in clitic placement between proclisis and enclisis can be explained through verb movement, while Batllori et al. (2005) relate variation in clitic placement to the interaction of syntax and information structure.

In OCat, like in the rest of the Old Romance languages, clitics followed certain distributional patterns:

(54)

- a. In main clauses, clitics had a strong tendency to be enclitic when no constituent preceded the verb, and to be proclitic when a constituent preceded the verb.
- b. In the presence of the negative marker *no* or another negative polarity item (NPI), including the negative coordinating conjunction *ni* ‘neither... nor...’, proclisis was borne.
- c. In embedded clauses, they appear in proclisis.
- d. The clitic pronoun is generally adjacent to the verb.

However, OCat data proves unruly, and does not always abide by these guidelines: proclisis is found in clauses where no constituent precedes the verb, and the reverse is true of enclisis: it is found in clauses where there are one or more constituents preceding the verb. In regard to negation, nevertheless, OCat abides by the rule to the letter.

Fischer (2002) observes that the distribution of OCat clitic pronouns cannot be TML, since proclisis is found even in V1 clauses, clause-initially, and since even when one or more constituents precede the verb, enclisis can obtain. This is indeed also the case for the data from the LFRJ database:

- (55) E Don Ató **atench-lo** a l' eixida d' unes tàpies
 and Sir Ató reached.3SG=him.CL at the;exit of;some fences
 And Sir Ató reached him at the gate of some fences.

Fol. 15r, l 3

- (56) E li **havia** tolt lo castel e la vila d'Alvero
 and to=him.CL had.3SG taken the castle and the village of Alvero
 e bé ·x· míllia kafizes de pa qui eren seus
 and good 10 thousand kafizes of bread that were his

And he had taken the castle and the village of Alvero, and a good 10,000 kafizes²³ of bread that were his.

Fol. 7r, l 23

In (55), in spite of the presence of the preverbal subject *Don Ató* 'Sir Ató', the object clitic *-lo* is enclitic to the verb, and conversely, in (56), while there is no constituent preceding the verb (and we already established in section 3.1.2 that *e* in OCat is extra-clausal or, in any case, located above Force^o), the clitic appears in proclisis. These cases are by no means isolated. Tables 7 and 8 shows the distribution of clitic pronouns in main clauses with respect to linear position of the verb in the clause as well as the presence or absence of a negative marker:

Enclisis	+ overt negative marker	- overt negative marker	Total
V1	-	177	177 (55.3%)
V2	-	121	121 (37.8%)
V3	-	22	22 (6.9%)
Total	-	320	320 (100%)

Table 7 – Distribution of enclisis in the LFRJ database database's main clauses

²³ *Kafizes* are a weight measure that was in use in the Iberian Peninsula during the Islamic Caliphate.

Proclisis	+ overt negative marker	- overt negative marker	Total
V1	12	6	18 (26.1%)
V2	18	25	43 (62.3%)
V3	7	1	8 (11.6%)
Total	37	32	69 (100%)

Table 8 – Distribution of Proclisis in the LFRJ database’s main clauses

As shown in Tables 7 and 8, enclisis is much more frequent in main clauses than proclisis, and it never co-occurs with an overt negative marker. Proclisis, on the other hand, can co-occur with negative markers. In terms of linearisation, enclisis can obtain when one or more constituent precedes the verb (a non-negligible 44.7% of cases of proclisis occur in clauses where at least a constituent precedes the verb). In contrast, proclisis can obtain in the absence of a negative marker, in V1 clauses.

In the light of this evidence, Fischer proposes that the position of OCat clitics was not just subject to the availability of a phonological host, but that it was the result of the interaction between verb position and polarity features. According to her, clitics are DPs whose head moves from its base-generated position in vP to one of the functional heads of I to check a [Specificity] feature, where they remain. In turn, the verb, also base generated in vP , moves to I to check [Person] and [Number] features. Following Laka (1990) and Martins (1994), Fischer puts forward the existence of ΣP , a projection linked to truth value (affirmation, emphatic affirmation and negation), whose features can be checked in different ways depending on the language. For OCat, she proposes that the emphatic affirmative feature in ΣP is checked via verb movement from the inflectional domain to ΣP . In affirmative main clauses the verb did not move. Finally, in clauses where the negative marker *no* appears, it is taken to phonologically realise Σ^0 . This is summarised in (57):

- (57)
- (i) Σ^0 [+V]: emphatic affirmation
 - (ii) Σ^0 [-V]: neutral affirmation
 - (iii) Σ^0 [no]: negation

If OCat Σ^0 had been endowed with the features described by Fischer (2002), the oscillation between proclisis and enclisis could be explained via polarity motivated verb

movement. However, this account encounters two main difficulties. Firstly, as pointed out by Batllori et al. (2005), Fischer's analysis entails that all clauses with enclisis are emphatic affirmative clauses. In the case of the LFRJ database, that would mean that out of the 389 clauses that present clitic pronouns, 82.3% are emphatic affirmative clauses. If we take frequency and unmarkedness to go hand to hand, that would translate into emphatic affirmative clauses being the unmarked option, with the subsequent issue of whether an emphatic configuration can be unmarked. Furthermore, Fischer (2002) does not make clear what contexts would trigger this emphatic word order pattern or which contextual discourse requirements would be needed.

Batllori et al. (2005) offer an alternative analysis of clitic distribution in OCat, which stems from Martins (1994, 2013)'s work on the distribution of clitics in MPt. MCat and MSp present a rather neat split between proclisis and enclisis, related to finiteness: finite forms co-occur with proclisis, and non-finite forms with enclisis. However, in MPt and Modern Galician (MGal), enclisis is still the default option, with the exception of several syntactic contexts:

(58)

- (i) When a *wh*-word precedes the verb.
- (ii) When a focalised constituent precedes the verb.
- (iii) When a quantifier precedes the verb.
- (iv) When the negative marker or a NPI precedes the verb.
- (v) When a focalised adverb precedes the verb.
- (vi) In emphatic declarative clauses.
- (vii) In optative clauses.

Adapted from Batllori et al. (2005)

Proclisis cases from the LFRJ share exactly the same distribution pattern as that sketched by Martins for MPt and MGal. Therefore, against the conclusion reached by Fischer (2002), it is not enclisis but proclisis that represents a marked option linked to emphasis, since its occurrence seems to be linked with the presence of a focalised constituent to the left of the verb. The structure of OCat's left periphery and the encoding of emphasis is extensively discussed in Chapter 5. The relation between polarity and verb movement is considered in length in section 3.1.6 in this Chapter.

We now consider whether clitics found in the LFRJ database were sensitive or not to the presence of any of the proposed triggers by Batllori et al. (2005). Examples (59-62) display clauses with a preverbal constituent that contain a clitic pronoun. In all cases, the context has been provided so that the reader may be able to evaluate the discourse value of the preverbal constituent, signalled by square brackets.

(59) Wh- word

E aquels nobles qui eren venguts ab la
 and those noblemen who were.3PL come with the
 fiyla de l'emperador demanaren [què]·s
 daughter of the;emperor asked.3PL what=REFL.3PL
farien d'aquel engan (...)
 do.3PL.COND of;that trick

And those noblemen who had come with the bishop, asked themselves what they would make out of that trick (...).

Fol. 2r, 1 18

(60) Focalised constituent

a. Context: E levà·s el comte d'Ampúries e dix:

And the count of Empúries stood up and said:

'AÇÒ us **diré** yo (...).
 this to;you.CL say.1SG.FUT I

I will say this to you (...).

Fol. 29r, 1 8

b. Context: E **él** levà·s en peus e dix: 'Seyor, vera cosa és que...

And he stood up and said: 'Sir, it is true that...

On, NÓS vos **deïm**, sobre· ls ·iii· conseyls que
 where we to;you.CL= say.1PL about=the.CL 3 advice that
 vós nos havets demanats(...)
 thatyou to;us.CL = have.2PL asked

Here, we tell you, about the three pieces of advice that you have asked us for (...).

(61) Negative marker

E sobre açò, En Guillem de Muntcada e Don Pero Fferràndez vengren
 and on this Sir Guillem of Muntcada and Sir Pero Ferràndez came.3PL
 ab tot lur poder e [no] y **pogren** entrar
 with all their power and not there.CL = could.3PL enter.INF
 sinó ab aquels que nós havíem manat.
 but with those that we had.1PL ordered

And then, Sir Guillem of Muntcada and Sir Pero Ferràndez came with all their forces and they could not get in but with those that we had decided.

Fol. 10v, 1 2

(62) Focalised adverb

“Açò dit és, e AIXÍ ho **atorgam** tots (...)”.
 this said is and thus it.CL= acknowledged all

‘This has been said, and thus we all acknowledge it’.

Fol. 10r, 1 25

In example (59), the third person plural clitic pronoun *es* ‘themselves’ is preceded by the wh- interrogative pronoun *què* ‘what’. As predicted, it is found in proclisis. In (60), we are presented with two cases of focalised constituents preceding a proclitic clitic pronoun. In (60.a), the constituent preceding the clitic is the neuter demonstrative pronoun *açò* ‘this’. In (60.b), the constituent preceding the clitic is a strong personal pronoun, *nós* ‘we’. The information status of pronouns is discussed at length in Chapter 4. In (61), the clitic follows the negative marker *no* ‘not’ and like in the rest of the Old Romance languages, it is proclitic. Finally, in (62), we encounter a focalised adverb *així* ‘thus’, which is also followed by a clitic in proclisis. The other contexts identified by Batllori et al. (2005) as triggers of proclisis are not found in the LFRJ database.

Examples (59-62) conform to the description of Batllori et al. (2005), and also to descriptions of clitic placement made for OSp (Bouzouita, 2008; Fernández-Ordóñez, 2008). However, they could also be said to fit within Benincà (2004, 2006) and Donaldson (2016)’s approach to clitic position: proclisis is borne when there is a constituent saturating SpecFocP,

the verb having moved to Foc° , and *wh*- words, focalised constituents (DPs and adverbs), quantifiers and emphatic elements can be said to occur in this projection.

Proponents of the V2 hypothesis for the Old Romance languages have used clitic placement as evidence for their analysis of verb position in the Old Romance languages. Following Rizzi's (1997) cartographic analysis of the left periphery, several descriptions of the Old Romance left periphery ensued (Poletto 2002; Benincà 2004; Benincà and Poletto 2004; Benincà 2006). Here, I reproduce Benincà's (2006, p. 32) description of the left periphery of Old Romance:

- (63) [Force C° [Rel wh C°]/{Frame [ScSett][HT] C° } {Topic [LD] [LI] C° } {Focus[I Focus] [II Focus]/[Interr wh] C° } [Fin C°]

According to Benincà (2006), in the Old Romance languages, the verb raised to the Focus field within the CP layer,²⁴ and a constituent needed to be merged in SpecFocP. As exposed in section 3.1.3, Benincà relates the oscillation between preverbal and postverbal clitics to the saturation of SpecFocP: if SpecFocP is saturated or filled, proclisis is expected. If not, enclisis follows. Similarly, Vance, Donaldson & Steiner (2010); Donaldson (2016), base their analysis of the interaction between elements in the left periphery and clitic placement in OFr and OOc on the assumption of the verb raising to Foc° and whether (or not) SpecFocP was filled. According to Donaldson (2016), the sequence XP CL V is found in clauses where SpecFocP is saturated, while the sequence XP V CL stands for clauses in which SpecFocP is not saturated and the preverbal constituent is located higher up in the left periphery, within the Topic or Frame field. This author points out that OOc texts vary in regard to their tolerance of empty SpecFocPs. While the *Vidas* (bibliographies of troubadours, 13th century, of north western origin) have a very strong preference for it to be saturated, *La vida de Santa Doucelina* (hagiography, 13th century, Provençal origin) presents S CL V sequences with topicalised subjects liberally, as it is also the case in OCat (see Chapter 4 for a description of subject distribution in OCat and considerations on Donaldson [2016]).

At this point, it can be established that evidence points towards a relationship between focalisation and proclisis, rather than between clitic placement and the nature of clausal

²⁴ Wolfe (2016), building on Benincà (2006), argues that there were two types of Old Romance languages: those in which the verb rose to FocP , and those in which the verb raises to FinP . This determines the number of constituents that can occur preverbally. When the verb is located in ForceP , only one constituent can occur preverbally, in SpecForceP , to satisfy the edge feature of this projection. When the verb is located in FinP , however, several constituents can appear above it, mainly in the Frame field.

polarity. Focalisation, its syntactic structure and its relationship with emphatic polarity is explored at length in the following section and Chapter 5.

3.1.6 Verb position and polarity

V-to- Σ P movement in OCat, as proposed by Fischer (2002), has been dismissed in section 3.1.5²⁵. Other authors have linked the encoding of assertive force in unmarked main declarative clauses with verb movement to the left periphery in OSp (Rodríguez Molina, 2014; Sitaridou, 2019). In this section, it is shown that OCat did only display V-to-C movement in main clauses in emphatic polarity contexts linked to Focus Fronting, like its Modern counterpart, but not in unmarked main clauses. It is also established that already by the 13th OCat, *sí* had been grammaticalised as an Emphatic Polarity Particle and was no longer a manner adverb, earlier than in OSp, where this grammaticalisation did not take place until the 15th century (Rodríguez Molina 2014, p. 896).

The lexical item, *sí*, etymologically derived from the Latin manner adverb SIC ‘thus’, has received great attention in the literature on historical Romance linguistics, having been provided with several analyses that range from being given expletive status (Ledgeway, 2008; Poletto, 2005; Sitaridou, 2004; Wolfe, 2015a), to being analysed as subject continuity marker (Marchello-Nizia, 1985; van Reenen & Schlösler, 2000, both for OFr), being analysed as an Emphatic Positive Polarity Particle (henceforth EPPA) (Batllori & Hernanz, 2008; Batllori & Hernanz, 2013 for OSp and OCat; Rodríguez Molina, 2014 for OSp).

In this section, we focus on the distribution of *sí* in OCat and its interaction with verb movement. It is shown that in OCat, *sí* was already intimately related to emphatic positive polarity. This section is organised as follows: first, the different analyses proposed for *sí* in OSp and OPt, those linked to verb movement, are presented. This is followed by a presentation of the OCat data from the LFRJ database and an analysis of it.

By the 13th century, *sí* had already split from its manner adverb counterpart *així* (from Latin *ACCU-SIC, SIC plus a demonstrative derived from ECCE.²⁶ The differences between both elements go beyond spelling: their syntactic distribution clearly reflects the fact that they are two distinctive lexical items: an adverb and something that already resembled an emphatic

²⁵ See Martins (2004, 2011, 2013) for analyses that confirm that V-to- Σ P occurred in OPt.

²⁶ This etymology is provided by the Alcover-Moll dictionary.

polarity particle. In the LFRJ database, three different spellings are attested for *així*: *ayxí*, *axí* and *asi*²⁷. Its distribution is akin to MCat *així*: it is base generated in the VP, and can be focalised depending on its discourse value:

- (64) E **façam-** ho axí: (...)
 and did.1PL= it.CL thus
 And we did it in this way: (...)

Fol. 72v, l 15

- (65) E ayxí **féu** -se el matrimoni.
 and thus made.3SG= itself.REFL the marriage
 And thus, the marriage took place.

Fol. 3r, l 9

Batllori & Hernanz (2008) propose that Latin SIC underwent upward grammaticalisation (Roberts & Roussou, 2003) that transformed it from SIC, a manner adverb meaning ‘thus’, into a positive polarity item. There is no doubt of the adverbial status of SIC in Latin: unlike Latin positive polarity particles like ENIM or QUIDEM, which abide by Wackernagel’s Law,²⁸ SIC occurs preverbally (Latin was an SOV language), and it does not require to be verb adjacent.

- (66) Vir autem acerrimo ingenio – sic enim **fuit** –.
 man however great.ABL.SG intelligence.ABL-SG thus indeed was.3SG
 He was a man of great intelligence – he was thus indeed.

Cicero, *Orator*, 18.8

In (66), SIC occurs preverbally, preceding the positive polarity particle ENIM, which needs to appear in sentential second position and in verb adjacent position.

Batllori & Hernanz (2008) identify the context in which SIC was reanalysed as answers to direct and indirect questions involving verb echo (henceforth, V-echo) and verbs of

²⁷ In fragments where there is code-switching into OSp, the spelling *assi* is also attested.

²⁸ Wackernagel’s Law states that ‘enclitics stand in the second position of the sentence’. See Clackson (2007) for an overview of this phenomenon in Indo-European, and Adams (1994) for a Latin-specific analysis.

support, similar to English do-support (Biberauer & Roberts, 2008, p. 34). V-echo and do-support in OSp are illustrated in examples (67) and (68):

- (67) E dixo: ¿**es** este el vuestro hermano el menor
 and said.3SG is this the your brother the younger
 que ·m dixiestes? E dixieron: Sí **es**.
 that =to;me said.2PL and said.3PL si is

And he said: ‘Is this your brother, the little one you told me about?’ And they said: ‘Yes, it is’.

Fazienda, 8v, *apud* Rodríguez Molina (2014, p. 874)

OSp

- (68) Yo le dije todo lo que había pasado.
 I to;him said.3SG all it that had.3SG happened
 Mandome me fuesse a Écija. Dije
 ordered.3SG=to;me.CL I:ACC go.3SG.PST.SBJV to Écija said.1SG
 que sí **haría**
 that yes do.1SG.COND

And I told him everything that had happened. And he told me to go to Écija. And I said that I would.

Vida, 316, *apud* Rodríguez Molina (2014, p. 876)

OSp

In (67), the verb *ser* ‘to be’ that appears firstly in the yes-no question, is repeated in the affirmative answer, preceded by *sí*. In (68), the verb of the indirect question, *irse* ‘leave’, is not repeated. Instead, the answer contains the verb *hacer*, ‘to do’, which refers anaphorically to the predicate of the preceding clause. In both cases, no clitic pronoun occurs between the verb and *sí*.

Batllori & Hernanz (2013) analysis assumes the existence of two projections related to the encoding of polarity in the clause (Holmberg, 2001; 2016; Martins, 2013). Following them, we assume that polarity in main clauses is encoded in two different heads: PolP, in the CP layer, hosts relative polarity features. Relative polarity features are features that encode the agreement or denial of the polarity of a clause in relation to the previous utterance,

referred to as [reverse] and [same] in Farkas & Bruce (2010) analysis. In contrast, ΣP , the highest functional projection in the inflectional domain, hosts absolute polarity features, [+] and [-], corresponding to [+affirmation] and [+negation]. Following Holmberg (2001), Martins (2013) points out that these two heads interact to encode emphatic polarity. With this sentential structure in mind, we can now consider Batllori & Hernanz's (2013) sketch of the upward process of grammaticalisation of *sí* in do-support structures of OSp, by which *sí* starts off as a variant of the manner adverb (*a*)*sí* (from Latin *ACCU-SIC), that firstly undergoes movement to the left periphery to receive Focus Fronting (this happened *in this precise way and no other*, see Chapter 5, section 5.2 for a thorough description of this phenomenon in the LFRJ database's OCat). From there, it gains scope over the whole clause and undergoes reanalysis and starts being directly merged in PolP, undergoing subsequent focalisation:

- (69) a. [CP ... [FocP ... [PolP ... [TP ... [VP fago *así*]]]]]
 b. [CP ... [FocP ... [PolP ... [TP *así/sí*_i [VP fago *t_i*]]]]]
 c. [CP ... [FocP *así/sí*_i [PolP *t_i* [TP ... [VP fago *t_i*]]]]]
 d. [CP ... [FocP *Sí*_i [PolP *t_i* [TP [VP]]]]]]

As it can be seen in (69), *sí* starts off as a manner adverb in its base generated position in VP, and through a process of grammaticalisation and fronting to FocP, it grammaticalises as an emphatic positive polarity marker: it is no longer moved to the left periphery but directly merged in PolP, and focalised in FocP, location of emphatic features, according to Rizzi (1997); Holmberg (2001); Batllori & Hernanz (2008).

Rodríguez Molina (2014) builds on Batllori & Hernanz's (2013) analysis of *sí* in OSp. Departing from Martins' (2013) classification of the Romance languages based on whether verb movement to a polarity related projection is available in a language or not, he establishes that OSp, like MGal and MPt, had verb movement to PolP in unmarked main declarative clauses associated to the marking of assertive force (a similar hypothesis is defended in Sitaridou [2019]). The trigger of movement to PolP are the strong features hosted in this head that attract the verb (note that Martins [2013] defends that verb movement is to ΣP , the highest head of the inflectional domain, rather than PolP in the left periphery).

When a language has verb movement to a polarity related projection, certain syntactic structures become available: it presents V-echo answers with VP ellipsis and answers with the verb *fazer/hacer* 'to do' to yes/no questions and does not have a grammaticalised positive polarity sentential proform or the option to generate emphatic positive polarity clauses in

- (71) “Desto que nos abino, que uos pese, señor.”
 this that to;us = came.3PL that to;you= burdens lord
 Respondio el Rey: “si fago, sin salve Dios!”.
 answered.3SG the King yes do.1SG without save.PRES.SBJV.3SG God
 [ForceP [TopP [FocP S_i [PolP t_j fago_i [FinP [TP t_i]]]]]] sin salve Dios!”.
 si do.1SG without save.3SG.PRS.SBJV God
 ‘What has happened to us should sadden you, lord’/ The King replied: ‘yes it does,
 may God save me!’.

Çid, v. 3040.3042, *apud* Rodríguez Molina (2014, p. 876)

OSp

From the evidence presented in (70) and (71), it can be established that both structures involve the same type of verb movement to the left periphery.

Since OSp was a language that had V-to-Pol, at least according to Rodríguez-Molina (2014), following Martins’ (2013) classification of the Romance languages in relation to this property, it is not expected to have a positive sentential proform. Rodríguez Molina (2014) explains the unavailability of using *sí* as a sentential proform by arguing that it was still categorically an adverb and had not yet been grammaticalised, and because other strategies for the expression of emphatic positive polarity were in place, mainly V-to-PolP-to-FocP. In this sense, OSp differs from OCat or OOc, languages that made use of *òc*, derived from the neuter Latin demonstrative HOC, as a sentential proform.

Òc, unlike MCat *sí*, could only appear as a positive answer to direct or indirect yes/no questions and it never appears within fully fledged main clauses or in combination with the conjunction *que* ‘that’. This is illustrated in examples (72) and (73):

- (72) E dixem -li: “Eres tu ab lo comenador?”
 and said.1PL=to;him are.2SG you with the head knight
 e él dix: “seyor, och”.
 and he said.3SG lord yes

And we said to him: ‘Were you with the head knight?’ and he said ‘Lord, yes’.

Fol. 103r, l 16

(73) E entant atrobam Don Pelegrí d’Atrosillo e
and meanwhile found.1PL Sir Pelegrí d’Atrosillo and
demanam-li si hi havia aygua on poguéssem la
asked.1PL=to;him if there.CL= had.3SG water where could.1PL the
nuyt albergar. E dix que och, (...).
night camp.INF and said.3SG that yes

And in the meanwhile, we bumped into Sir Pelegrí d’Atrosillo and we asked him if there was any water available in a place where we could camp for the night, and he said ‘yes’.

Fol. 39v, l 20

The distribution of OCat *sí*²⁹ differs substantially from OSp and Opt. Data from the LFRJ database suggests that in this language, *sí* behaved in more innovative manner than in other varieties. Instances of *sí* in *El Llibre dels Fets* can be classified in the following 6 categories: (i) occurrences of *sí* in fully fledged main clauses, (ii) with V-echo, (iii) with do-support, (iv) adverbial *sí*, (v) occurrences of *sí* as a sentential proform, and finally, (vi) occurrences of *sí* in the expression ‘may God help ‘x’’. Table 10 shows the frequency in which each of this structures occurs with *sí*:³⁰

²⁹ While in MCat and MSp, the convention is to distinguish emphatic positive polarity *sí* from the conditional conjunction *si* orthographically by means of a diacritic, this is not always the case in OCat and OSp texts. Here, I respect the editor’s decision in adding the diacritic to the emphatic polarity particle, and I rely on the glosses to show the difference in meaning.

³⁰ I have not been able to find any cases of V-echo or do-support without *sí* in *El Llibre dels Fets*. This does not exclude this possibility elsewhere.

	Number of occurrences	%
Main clauses	19	49
<i>Sí</i> V-echo	7	18
<i>Sí</i> do-support	2	5
Adverbial <i>sí</i> (+ scribal error)	2	5
Proform	1	2.5
<i>Sí</i> Déus	8	20.5
Total	39	100

Table 10 – Distribution of *sí* in *El Llibre dels Fets*

As shown in Table 10, *sí* occurs more frequently in fully fledged main clauses than in any other structure. This already establishes a difference between OCat and OSp, where *sí* could not occur in fully fledged main clauses until the 16th century (Rodríguez Molina, 2014 p. 878). From quite a distance, the next context in which *sí* is found is the idiomatic expression '*sí Déus ajut 'x'*', which we are not considering in this section.³¹ It is followed by *sí* V-echo, *sí* do-support, adverbial uses of *sí* and *sí* as a positive sentential proform. Each of these contexts is commented on in order of frequency.

There are 19 occurrences of *sí* in fully fledged clauses with no VP ellipsis in the LFRJ. Syntactically, these clauses present the following features: (i) *sí* has to be verb adjacent, (ii) object and adverbial clitics can appear between *sí* and the verb, and (iii) topical material can precede *sí*. This is illustrated in examples (74-77):

- (74) E metem mà a ·iii· guaytes fer.
and put.1PL hand to 3 watches make.INF
- [La una guayta] sí era als genys, e a les cledes
the one watch sí was to;the war machines and to the fences

³¹ '*Sí Déus ajut 'x'*' is an expression attested throughout the Old Galo-Romance languages. There is no consensus in the literature regarding the category of *sí* (it could be a case of the conditional conjunction 'if'), or the origin of this expression. Hence, we leave it for future research. See Marchello-Nizia (1985, p. 55) for an overview of this expression for OFr.

la altra, sí era contra la porta de Barbelec.
the other sí was against the door of Barbalec

And we decided to set up three watching posts. One, was next to the catapults, the other at the wooden fortress, which was against the gate of Barbalec.

Fol. 48r, l 19

- (75) Si homes del món àn mala fama, [nós] sí
if men of;the world have.3PL bad reputation we sí
la havem bona.
it have.1PL good

If some men of the world have bad reputation, ours is indeed good.

Fol. 29r, l 10

- (76) E sí hi ach altres paraules que començaven ja
and sí there has other words that started.3PL already
de dir mas nós ho vedam.
of say.INF but we it.CL= forbade.INF

And Sir Guillem de Muntcada and Sir Pero Fferàndez had to leave. And there were already other speeches starting **indeed**, but we forbade it.

Fol. 10v, l 29

- (77) E anch per aquesta paraula nengú no·s moch,
and never for this speech no one not=himself.CL moved.3SG
e sí la hoïren tots.
and sí it heard.3PL all

And no one at all budged because of this speech and they all heard it indeed.

Fol. 49r, l 13

In (74), *sí* occurs between the topical subject, *la una guayta*, ‘one of the watches’, and the verb *era* ‘it was’. In (75), *sí* is preceded again by the clause’s subject, *nós* ‘we’. In contrast, in (76), *sí* occurs in clause initial position, only preceded by the coordinating conjunction *e* ‘and’. The adverbial clitic pronoun *hi* ‘there’ appears between *sí* and the verb. The predicate of this clause is existential, therefore, there is no preverbal subject. In this case,

it could be argued that *sí* appears in the clause for the purpose of satisfying the EPP requirement of a subject-less verb (the pivot and the existential verb do not agree in number). (77) could provide extra evidence for this, since *sí* occurs sentence-initially again. In this case, an object clitic intervenes between it and the verb, and the clause's subject, *tots* 'all' appears postverbally, potentially triggering the presence of *sí*. As appealing as this possibility may seem, it can be argued that EPP is actually satisfied by the presence of the locative clitic in (76) (see more on the satisfaction of EPP in OCat in Chapter 4), and by *pro* in (77).

All instances of *sí* in the LFRJ database have one thing in common: they occur within the scope of a non-veridical licensing environment (Giannakidou 1998; 1999; 2001; 2008 and subsequent). Giannakidou's (non)veridicality model accounts for the distribution and licensing of polarity items, defined as follows:

(78) Polarity item

- a. A polarity item α is an expression whose distribution is limited by sensitivity to some semantic property β of the context of its appearance.
- b. B is (non)veridicality.

Giannakidou defines (non)veridical operators as in (79):

- (79) 'An operator Op is nonveridical iff the truth of Op p in c does not require that p be true in some such model in c'.

Giannakidou (1998, p.112)

In other words, a (non)veridical operator does not require a proposition to be true in some of the models (set of worlds) of a specific context. Most importantly, a (non)veridical operator does not entail the falsity of the proposition.

Polarity Items (PIs) fall in two different categories: negative polarity items (NPIs), and positive polarity items (PPIs). However, here we are concerned with PPIs, since *sí* comes to be a PPI in the Modern Romance languages. PPIs, as shown by Ernst (2009), do not occur merely in assertive contexts: they can also appear in nonveridical ones. Giannakidou (1998) identified the following non-veridical contexts as potential licensors of PIs:

(80) Non-veridical contexts in which PIs are licensed

- (i) Monotone quantifiers
- (ii) Modal verbs

- (iii) Intentional verbs
- (iv) Non-declarative clauses: interrogative, imperative and exclamative clauses.
- (v) Protasis of conditionals, conditional clauses, *before*-clauses
- (vi) Habitual tenses
- (vii) Future tenses
- (viii) Nonveridical implicatures

Invariably, all the OCat instances of *sí* in fully fledged clauses occur within the scope of nonveridical contexts. This is shown (81-87), where several cases of *sí* that occur in the LFRJ are reproduced with their wider context, where needed. With the context in hand, it is possible to identify the nonveridical licenser of *sí*, which may be explicitly expressed in the context, or implied (see Giannakidou [1998] for an overview of indirect licensing of PIs):

- (81) E durà la batayla, que no pujà negú. E sí
 and lasted the battle that not climbed.3SG anyone and sí
 y **havia** bon pujador.
 there was good climber

And for the whole battle, no one climbed (the wall). And there was indeed a good climber.

Fol. 7v, l 25

Licenser: implicature derived from the preceding clause, where it is stated that no one climbed the wall for the duration of the battle, implying that there was no one able to do it.

- (82) Context: Guillem de Muntcada stops being loyal to the King and to his trustworthy knight Don Nuno. The King, after marrying the princess of Castile, summons a court meeting in the city of Montsó. He requires that no knight attending the court enters the city with more than two men. Guillem de Muntcada and his ally Pero Ferràndez, who had planned to carry out a coup, and had approached the city with all their manpower, find that they could not get in. The court began, and...³²

³² The context of this clause is rather extensive (Fol. 9v, 10r). For the sake of space, a summary of it is provided.

E sí hi ach altres paraules que començaven ja
 and sí there has other words that started.3PL already
 de dir mas nós ho vedam.
 of say.INF but we it forbade.INF

And there were already other speeches starting **indeed**, but we forbade it.

Fol. 10v, l 29

Licensor: implicature derived from the context. Since Guillem de Muntcada and Pero Ferràndez were leaving, the assembly could have been expected to have stopped. But that was not the case.

- (83) Context: Certa cosa és que·l nostre naximent se féu per vertut de Déu, car no·s volien bé nostre pare ni nostra mare,

It is indeed true that our birth took place by virtue of God, since our father and mother did not love each other,

e sí **fo** volentat de Déu que nasquem en
 and sí was will of God that be born.1PL.PRS.SBJV in
 aquest món.
 this world

And indeed, it was God's will that we were born in this world.

Fol. 28r, l 15

Licensor: Implicature derived from the preceding clause. Since his father and his mother did not love each other, his birth was unlikely to happen.

- (84) Si homes del món àn mala fama, nós sí
 if men of;the world have.3PL bad reputation we sí
 la **havem** bona.
 it have.1PL good

If some men of the world have bad reputation, ours is indeed good.

Fol. 29r, l 10

Licensor: preceding protasis.

- (85) E metem mà a ·iii· guaytes fer.
 and put.1PL hand to 3 watches make.INF
 La una guayta sí era als genys, e a les cledes
 the one watch sí was to;the war machines and to the fences
 la altra, sí **era** contra la porta de Barbelec.
 the other sí was against the door of Barbalec

And we decided to set up three watching posts. One, was next to the catapults, the other at the wooden fortress, which was against the gate of Barbalec.

Fol. 48r, l 19

Licensor: monotone quantifiers ‘three’ and ‘una’.

- (86) E anch per aquesta paraula nengú no·s moch,
 and never for this speech no one not=himself.CL moved.3SG
 e sí la hoïren tots.
 and sí it heard.3PL all

And no one at all budged because of this speech, and they all heard it indeed.

Fol. 49r, l 13

Licensor: implicature derived from preceding clause: since no one moved after the speech, it is possible to infer that they had not heard it.

- (87) E ja·s fos que nós haguéssem a entrar en
 and already be.3SG.PST.SBJV that we had.1PL.PST. SBJV to entry.INF in
 guerra ab lo rey de Castella, ab tot açò, sí·ns
 war with the king of Castile with all this sí=to;us.CL
era bo el pleyt per ·iii· raons: (...)
 was good the dispute for 3 reasons

And even if we had to wage war against the king of Castile, with all this, the conflict would be good for us for three reasons.

Fol. 71r, l 7

Licensor: preceding protasis, introduced by the concessive conjunction *jatsia* ‘even if’.

As it can be seen in examples (81-87), *sí* consistently occurs within the scope of a nonveridical licenser, be it overtly expressed or implied. *Sí* is used to ascertain the truth of a declarative clause, against the nonveridical background, and therefore, it can be established that in the LFRJ, *sí* is already a PPI that occurs after nonveridical contexts, carrying out polarity reversal. Importantly, it does not only occur in answers to yes/no questions, as described by Batllori & Hernanz (2008). The link between emphatic positive polarity and polarity reversal is further explored in Chapter 5, section 5.2.

This far, we have established that *sí* (i) was PPI, and (ii) that it was licensed by nonveridical contexts in fully fledged clauses.

In terms of clausal structure, OCat *sí* requires an analysis that can account for its strict adjacency to the verb, with clitics being the only element that can appear between both. Furthermore, it needs to allow for the presence of topical elements (see (83) and (87)) above *sí* and the lack of cases in which focal constituents precede it.

With all these factors in mind, I propose the following clausal structure to account for occurrences of *sí* in fully fledged clauses (main or embedded clauses with a full left periphery):

(88) *Sí* in OCat fully fledged clauses

[Force [Top [SpecFocP *Sí*_i [FocP V_j [PolP t_i [Pol' +t_j [TP t_j [VP t_j ...]]]]]]]]

The structure shown in (88) assumes that in fully fledged clauses, *sí* is base-generated in PolP, since, as we have shown, it already behaved as a PPI, being licensed by a nonveridical context. From there, it moves to SpecFocP (as mentioned above, verb movement to FocP is linked to the emphatic character of this construction). The verb, like in all cases of Focus Fronting, raises to Foc^o. This explains the compulsory adjacency between *sí* and the verb, except for object clitics: as phonologically dependent on the verb, they raise with it to Foc^o. It also explains why topical elements can occur above it, but not focal ones.³³ By postulating that *sí* is base-generated in PolP and not in VP, as it would be expected if it were still an adverb, we are accounting for the fact that there is not a single case of *sí* co-occurring

³³ There are no instances of *sí* being preceded by focalised constituents in *El Llibre dels Fets* or *sí* co-occurring with a wh- word. We have not been able to find any such examples in the *Corpus Informatitzat del Català Antic* (Torruella, Saldanya, Martines, 2009) either. The lack of evidence of the co-occurrence between *sí* and focal elements leads us to reach the conclusion that *sí* could not co-occur with them because they occupied the same position.

with negation spelled out in ΣP . PolP hosts relative polarity features, that is to say, polarity features that connect the clause's assertion with the wider context, while ΣP hosts absolute polarity features. Since *sí*, a PPI, asserts the 'positiveness' of the clause against the (non)veridical context and is associated with verb movement to Foc° , it is natural to assume that the spelling out of an overt negative marker in the lower polarity head ΣP is not possible, given that it gets valued with a [+] polarity feature when the verb moves from TP to Foc° through it. Furthermore, this hypothesis has an added advantage: it can account for the emphatic nature of *sí*, and for the fact that there is no VP ellipsis, a feature that was present in OSp verb echo and do-support structures.

(89)

- a. Si homes del món àn mala fama, nós sí
 if men of;the world have.3PL bad reputation we sí
 la **havem** bona.
 it.CL= have.1PL good

If some men of the world have bad reputation, ours is indeed good.

Fol. 29r, l 10

- b. [_{Force} [_{Top} *nós* [_{SpecFocP} *sí*_i [_{FocP} *la havem*_j [_{PolP} *t_i t_j* [_{ΣP} *t_j* [_{TP} ~~*la havem*~~_j [_{VP} *bona*]]]]]]]

In (89b), we can see how *sí*, base generated in PolP, moves to SpecFocP, while the verb and the object clitic leave TP and raise to Foc° . The adjective *bona* 'good', which agrees with the object clitic *la*, remains in the VP, proving that there is no ellipsis.

While it has been shown that OCat *sí* was a PPI linked to polarity reversal, the LFRJ also contains a few cases of *sí* V-echo and *sí* do-support, which are examined in what follows.

7 cases of *sí* V-echo have been identified in *El Llibre dels Fets*. Unlike in OSp, all cases co-occur with *sí* and are answers to yes/no questions. *Sí* V-echo answers are also found in OSp (Rodríguez Molina, 2014 and example (70) in this Chapter), but not in OPt (Martins, 2013), where *sí* was not yet involved in answering yes/no questions positively.

In OSp, V-echo answers involve VP ellipsis, since they never present postverbal constituents. Martins (2013, p. 109) connects the availability of these structures to whether a language has verb movement to PolP in unmarked assertive clauses. Rodríguez Molina (2014) legitimises VP ellipsis by arguing that, once the verb climbs to Foc° , FocP can elide its complement. Regarding the lack of clitics intervening between *sí* and the verb, he suggests

that *sí*, still based generated within the VP given its adverbial nature, incorporates with the verb in TP on its way to the left periphery. It is by virtue of this incorporation of *sí* to the verbal head and the ellipsis of FocP's complement that clitics cannot intervene between both. It could also be suggested that the lack of clitics derives from the VP ellipsis and the non-spelling out of their traces in VP. VP ellipsis also accounts for the lack of topical subjects such as those that we found in OCat, displayed in (74) and (75): since the VP is elided, they cannot be moved to the left periphery.

In the LFRJ, *sí* V-echo answers display the same features as OSp. However, there is an important difference: while OSp could generate V-echo answers without the presence of *sí*, as shown in (90), such cases are not found in *El Llibre dels Fets*, as it can be seen in (91-93):

(90) CELESTINA: ¿**Quiereslo** saber?
 Celestina want.2SG=it.CL know.INF

SEMPRONIO: **Quiero.**
 Sempronio want.1SG

CELESTINA: Do you want to know it?

SEMPRONIO: I do.

Celestina, 1.110, *apud* Rodríguez Molina (2014, p. 873)

OSp

(91) E nós resposem que: ‘Ço que nós n’ avem feyt, que
 and we answered.1PL that this which we of;it.CL=have.1PL done that
 u havem feyt ab dret’, e que neguna esmena no
 it have.1PL done with right and that no amend not
 li· n fariem. E éls dixeren que
 to;him.CL= of;it.CL= do.1PL and they said.3PL that
sí **faéssem**, que aquela esmena que nós fariem
 si do.1PL.PRS.SBJV that that amend that we do.1PL
 a él seria molta (...).
 to him be.3SG.COND much

And we answered that: ‘What we have done of it, we have done rightfully, and that we would not make any corrections to it’. And they said that yes, we would make (a correction), that the correction that we would make would be much for him.

Fol. 13r, l 21

- (92) E, hoïdes les paraules, respòs Don Nuno e dix -li:
 and heard the words answered.1PL Sir Nuno and said.3SG= to;him.CL
 ‘En açò que vós deïts que no tenits al rey nostre
 in this that you say.2PL that not have.2PL to;the king ours
 gran tort, sí **tenits**, quan li presés una tarida
 great offence sí have.2PL when to;him.CL. take.2SG a boat
 de son regne.
 of his kingdom

And, after hearing these words, Don Nuno answered back and said to him: ‘What you say that you have not committed a great offence against our king, yes you have, when you have taken a boat from his kingdom’.

Fol. 45, l 21

- (93) E nós volguem saber dels altres si eren
 and we wanted.1PL know.INF of;the others if were.3PL
 en aquel consel, e atorgaren tots que sí **eren**.
 in that meeting and decided.3PL all that sí were.3PL

And we wanted to know whether the others would be in that meeting and they all said that they would.

Fol. 104v, l 26

In (90), we can appreciate a case of V-echo without the presence of *sí* in OSp. This contrasts with (91) and (92), two OCat texts with V-echo, where *sí* has to be present for the verb to appear. In (91), V-echo appears in a reaction to a negative statement and reverses its polarity. In (92), the *sí* V-echo answer also reverses the polarity of the preceding statement, overtly negative. Finally, in (93), the licensing context is an indirect question introduced by *si*. In all cases, V-echo occurs invariably with *sí*.

From these examples, it can be established that *sí* V-echo in the LFRJ does not only occur in answers to direct questions: it can also be licensed by nonveridical contexts, as it was the case with *sí* in fully fledged clauses. Secondly, as we can see in (92), *sí* V-echo can also appear within embedded clauses: it appears within a complement clause introduced by the subordinator *que* ‘that’.³⁴

The appearance of *sí* was not just triggered by nonveridicality: there is one further variable that intervenes in the triggering of *sí*’s presence: emphasis. This becomes apparent below, where we examine a clausal minimal pair in which one clause contains *sí* and the other does not.

This is shown in (94) and (95), a minimal pair:

- (94) E nós resposem que: ‘Ço que nós n’ avem feyt, que
 and we answered.1PL that this which we of.it.CL= have.1PL done that
 u havem feyt ab dret’, e que neguna esmena no
 it.CL have.1PL done with right and that no amend not
 li ·n faríem. E éls dixeren que sí **faéssem**,
 to;it.CL= of;it.CL=do.1PL and they said.3PL that si do.1PL.PRS.SBJV
 que aquela esmena que nós faríem a él seria molta (...).
 that that amend that we do.1PL to him be.3SG.COND much

And we answered that: ‘What we have done of it, we have done rightfully, and that we would not make any corrections to it’. And they said that yes, we would make (a correction), that the correction that we would make would be much for him.

Fol. 13r, l 21

- (95) E pregam-lo que (...) él que romangués en nostre
 and begged.1PL=to;him.CL that he that remained in our
 loch en Maylorques, e que manariem als cavallers
 place in Mallorca and that ordered.1PL.COND to;the knights
 e a tots los altres hòmens que faessen per él
 and to all the other men that do.3PL.PST- SBJV for him

³⁴ Complement clauses have been shown to crosslinguistically display root phenomena (Hooper & Thompson 1973; Haegeman 2003; Haegeman 2009; Haegeman 2010 and subsequent).

axí con farien per nós. E él dix -nos que u
as do.3PL.COND for us and he said=to;us.CL that it.CL=

faria.

do.3SG.COND

And we begged him to remain in Mallorca on our behalf, and that we would order the knights and all the other men that they would act towards him as they act towards us. And he said that he would do it.

Fol. 57v, 15

In (94) (a repetition of (91) for the reader's convenience), *sí* V-echo appears in an embedded clause, complement of *dixeren* 'they said'. In the immediate context, the statement of the embedded complement clause had been overtly denied. In contrast, in (95), the same lexical verb, appearing in the complement clause of the same *verbum dicendi* that reportedly answers to a command, *dix-nos* 'they told us', and containing the same verb as in the preceding clause, the absence of *sí* allows for the presence of an object clitic pronoun. *Sí* is used to convey emphatic assertion, which is licensed within a nonveridical context, that also encompass antiveridical contexts³⁵. In (94), where *sí* reverses the polarity of the preceding clause emphatically, and it triggers the raising of the verb to Foc^0 and the elision of VP, preventing the appearance of the expected clitic *u/ho*, coreferent with the complement clause of *fariem*, the verb that is repeated in the *sí* V-echo construction. In (95), on the other hand, even though there is a nonveridical context that could license *sí*, namely the request introduced through the verb *pregam-lo* 'we begged him', it doesn't appear in the reported answer to the command. Instead, the answer is an unmarked assertive clause in which the verb remains in TP and allows for the appearance of clitics. Therefore, *sí* is linked to emphatic positive polarity, and specifically, polarity reversal.

Examples (96) and (97) reproduce the only two cases of *sí* do-support found in the LFRJ. Structurally, they are very similar to *sí* V-echo clauses: there is no material intervening between *sí* and the verb, not even clitics, and, like V-echo, it does not occur without the presence of *sí*. Both cases of *sí* do-support found in *El Llibre dels Fets* are the answer to an indirect question, a (non)veridical environment.

(96) (...) sol que la guardassen de les segetes. E nós

³⁵ Antiveridical contexts are those that convey the falsity of the positive counterpratt of the clause.

only that her kept.3PL from the arrows and we
 dixem-li que sí **fariem**.
 said=to;her.CL that sí do.1PL-COND

(...) as long as they would guard her from the arrows. And we said to her that indeed we would do it.

Fol. 24v, l 14

- (97) E, quan vench en l'altre dia, tornam -lo veer
 and when came.3SG on the;other day do;again.1PL =him.CL see.INF
 e enviam -lo a pregar que ns prestàs ·c· mília sous,
 and sent.1PL=him.CL to beg.INF that=to;us.CL lend.2SG.PRS.SBJV 100,000 sous
 e él dix que sí **faria** ab que nós li ho
 and he said.3SG that sí do.3PL.COND with that we to;him=it.CL=
 asseguràssem.
 ensured.1PL.PST.SBJV

And, on the following day, we went back to see him, and sent someone to beg him that he would lend us 100,000 sous and he said that he would do it, as long as we would provide a guarantee.

Fol. 74r, l 15

Both (96) and (97) have nonveridical licensing environments. In (96), it is the indirect question conveyed in the protasis that precedes the *sí* do-support answer. In (97) the licensing context is the intentional verb *pregar*. Therefore, we can establish that the presence of *sí* in do-support answers is also licensed by nonveridical contexts, as it is the case with *sí* V-echo structures and *sí* in fully fledged clauses. Given these similarities, we presume that the underlying structure of *sí* V-echo and *sí* do-support is the same.

With this in mind, we need to propose an analysis that can account for the differences between those cases in which *sí* appears in fully fledged clauses, without VP ellipsis and with clitics intervening between *sí* and the verb in contrast with *sí* V-echo and *sí* do-support structures.

On the view of the data presented, 13th century OCat presents a complex picture in regard to the syntax of *sí*. It can appear in two different types of structures that display significant differences, as it is summarised in Table 11:

	<i>Sí</i> in fully fledged clauses	<i>Sí</i> V-echo <i>Sí</i> do-support
(Non)veridical licensing context	+	+
<i>Sí</i> – cl – V	+	-
VP ellipsis	-	+

Table 11 – Features of the constructions with *sí*

For *sí* in fully fledged clauses, we have proposed an analysis in (89b), which we reproduce here for the reader’s convenience:

(98) *Sí* in OCat main clauses

[ForceP [TopP [SpecFocP *Sí*_i [FocP V_j [PolP t_i [FinP [ΣP t_j [TP t_j [VP ...]]]]]]]]]]

This analysis accounts for *sí*’s EPPA status, since it assumes that it is base generated in PolP, and that it moves to SpecFocP to receive an emphatic reading. At that point, similarly to wh-operators, with which it is mutually exclusive, it attracts the verb to Foc^o. This explains why subjects cannot intervene between *sí* and the verb, while they can occur when left dislocated in TopP, as shown in (84) and (85), and they can also present VP based postverbal subjects, as in (86). However, this analysis cannot account for the VP ellipsis involved in the *sí* v-echo and *sí* do-support structures.

Sí V-echo and *sí* do-support structures require a mechanism that would allow for VP ellipsis. We consider two proposals by which this could be explained.

Rodríguez Molina (2014) proposes that OSp allowed for VP ellipsis in V-echo and do-support structures by arguing that in OSp the verb raised to PolP in unmarked assertive declarative clauses, as shown in (99). In negative clauses, the verb would remain in TP, and *no* would be spelled out in ΣP, blocking the movement to PolP.

(99) [ForceP ... [PolP V_i [FinP ... [TP [T' t_i [VP t_i]]]]]]]]

Rodríguez Molina (2014, p. 890)

He considers that in OSp, *sí* had not yet been grammaticalised as an EPPA. According to his analysis, it was base generated in VP, and then, it underwent move to the left periphery to receive focus. By virtue of the verb then moving to FocP, the complement of this projection could be elided, yielding *sí* V-echo and *sí* do-support constructions. These constructions would be derived as in (100):

(100) [_{ForceP} ... [_{SpecFocP} *sí*_j [_{Foc'} V_i [_{PolP} t_i [_{FinP} ... [_{TP} [_{T'} t_i [_{VP} t_i t_j]]]]]]]]]

Rodríguez Molina (2014, p. 890)

We propose that OCat had two competing strategies for the expression of emphatic positive polarity: an archaising one on its way out, akin to that available in OSp, which yielded V-echo and do-support structures and was already associated to *sí*, and an emerging structure, akin to that found in MSp, found in fully fledged main clauses. The key difference between both structures is that, in fully fledged clauses, the lexical item *sí* has to have already been reanalysed as an EPPA that triggers the movement of the verb to Foc^o (see Chapter 5 on Focus Fronting), while to produce *sí* v-echo and *sí* do-support structures, *sí* and the verb must form a complex head to explain the unavailability of clitic pronouns occurring between *sí* and the verb, along the lines of Rodríguez Molina (2014). This sets OCat apart from the other Ibero-Romance varieties. It is also worth recalling that another fact that differentiates OCat from the other Ibero-Romance languages is that it had a positive sentential proform: *òc*, which does not have continuity in the modern language, as it was replaced by *sí*.

Therefore, we conclude that 13th c Catalan from the LFRJ has already grammaticalised *sí* as an EPPA, and that V-to- ΣP/PolP was restricted to a reduced number of archaising structures that involved polarity reversal.

Interestingly, this description of the clausal structure of declarative main clauses with *sí* parallels that of MSp clauses in which (i) *sí* occurs (Rodríguez Molina, 2014) verb adjacently, (ii) clitics can intervene between it and the main verb, (iii) it can co-occur with preverbal topical elements but not focal, (iv) and it triggers subject-verb inversion (Batllori & Hernanz, 2013). MCat offers a different picture, whereby EPPA *sí* necessarily co-occurs with the subordinating conjunction *que*.

In MCat, the EPPA *sí* occurs in emphatic affirmative main declarative clauses that reverse the polarity of previous overt statements or covert implicatures, focalising the polarity of the whole clause, and always with the conjunction *que*, as illustrated by (101) and (102):

- (101) A: En Jaumet no va comprar pomes-
the Jaumet not goes.3SG buy.INF apples
- B: Sí que en va comprar, en Jaumet.
yes that of;them.CL= goes.3SG buy.INF the Jaumet
- A: Jamie did not buy any apples
B: Yes, he did! / He DID buy apples!
- (102) A: Què passa?
what happens
- B.i: Avui arriba en Jaumet.
today arrives the Jaumet
- B.ii ?? Avui sí que **arriba** en Jaumet.
today yes that arrives the Jaumet
- A: *What's up?*
B.i: Jamie is arriving today.
B.ii: Jamie IS arriving today.

In (101), the presence of *sí* in B is legitimised by the polarity reversal of the question in A. In (102), however, reply B.ii compromises the adequacy of the clause since it is not reversing the polarity of an antiveridical context, and therefore, the presence of *sí que* is not legitimised, while B.i, an unmarked affirmative clause, is perfectly grammatical as an answer to a ‘What is happening?’ question.

In MSp, the distribution of clauses with *sí* is parallel to that of MCat, with one main difference: no *que* is required. It differs from OSp in that clitics can appear between *sí* and the verb:

- (103) A: Jaime no compró las manzanas.
Jaime not bought.3SG [the apples]_i
- B: ¡Sí las compró!
yes the_i bought.3SG
- A: Jaime did not buy apples.
B: Yes, he did buy them!

In MCat and MSp declarative clauses conveying emphatic polarity, the expression of emphatic polarity, like in the Old languages, also involves verb movement to the CP layer, specifically, to FocP. Evidence for this movement comes from (i) the fact that no other element can occur between *sí que/sí* in MSp and the verb, (ii) because it triggers subject-verb inversion, and (iii) because it is mutually exclusive with *wh*-words in interrogative clauses.

The Table 12 presents the different features of main clauses with *sí* in MCat, MSp and OCat:

	OSp	MSp	OCat	MCat
Verb adjacency	+	+	+	+
Subject-verb inversion	- (VP ellipsis)	+	+	+
<i>Sí</i> CL V	-	+	+	+
Mutual exclusion <i>sí/wh</i>-words	+	+	+	+
<i>Sí que</i>	-	+/- (Optional)	-	+

Table 12 – *sí* in OSp, OCat, MSp & MCat main clauses

In Table 12, the syntactic features of clauses with the EPPA *sí* in OSp, MSp, OCat & MCat are shown. There seems to be a continuum in terms of grammaticalisation of *sí* and its relation to verb movement in the left periphery, illustrated by the availability of *sí* CL V and the grammaticalisation of *que*. Both in OSp and MSp, clitics cannot occur between *sí* and the verb, for which, following Rodríguez Molina (2014) we have assumed VP ellipsis. However, note that MSp does not allow for *sí* CL V, but allows for *sí* V DO (as shown in (103)) and combinations of *sí* V and overt subjects, suggesting that in the Modern language : V-echo is no longer the only strategy available. OCat, like MCat, allowed *sí* CL V orders, which lets us establish that there was no VP ellipsis in fully fledged clauses (together with the fact that subjects can co-occur with *sí* V sequences, either in their base generated position in VP or left dislocated). Finally, in MCat, the structure of clauses with *sí* requires the presence of *que*, absent in both Old languages but optional in MSp. Therefore, OCat is closer to MSp than to

MCat, in that the use of *sí* as EPPA does not require VP ellipsis like in OSp or the presence of *que*, like in MCat.

To conclude this section, we consider what might be the first instance of *sí* used as a sentential proform attested in OCat. It is reproduced in (104):

(104) E nós dixem: ‘Batayla?’ e dix ell:
 and we said.1PL battle and said.3SG he
 Sí, seyor, que certament hi devem ésser est matí.
 yes sir that indeed there must.1PL be.INF this morning

And we said: ‘Battle?’, and he said: ‘Yes, sir, because we will certainly have to be there this same morning’.

Fol. 100v, l 16

In (104), *sí* appears in the answer to a direct yes/no question. It is followed by a vocative *seyor* ‘sir’, which precedes a clause introduced by the polysemous complementiser *que*. The fact that the vocative occurs between the *sí* and the *que* dispels the possibility of this being an instance of *sí que*, a compounded subordinating conjunction that introduced adverbial clauses in OCat. On the basis of this, we conclude that (104) is the first case of *sí* proform attested in Catalan.³⁶

So far, we have established that clauses conveying emphatic polarity in OCat, have the following characteristics: (i) emphatic positive value that is necessarily triggered by (non)veridical contexts, (ii) *sí*, having already been reanalysed as an EPPA, is base-generated in PolP, and (iii) there is V-to-C movement associated to emphatic positive polarity, the verb targeting FocP. Therefore, we have identified and described a context in which V-to-C movement occurs, which contrasts with unmarked declarative main clauses. As shown in Chapter 5, the points made here in regard to *sí* are also applicable to Focus Fronting, a construction intimately linked to the expression of emphatic polarity. Furthermore, contrary to other analyses of *sí*, we have shown that *sí* in OCat had already been grammaticalised as an EPPA by the 13th century, rendering Catalan innovative in this respect, and to our

³⁶ A search was carried in the *Corpus Informatitzat del Català Antic*. The 53 cases of *sí* attested in texts older than *El Llibre dels Fets* are either cases of adverbial *sí* or reflexive pronoun *sí*. It is on this basis that this claim is made.

knowledge, the first language to kickstart the grammaticalisation process that would end up with *sí* becoming a positive sentential proform.

3.2 VERB POSITION IN EMBEDDED CLAUSES

Vikner (1995, p. 65ff) proposes a widely accepted typology of V2 languages regarding the presence or lack of symmetry between main and embedded clauses. According to his classification, three types of V2 languages are proposed: (i) asymmetric V2 languages like German and Dutch, where the availability of V2 is linked to the absence of a complementiser, (ii) languages that only allow embedded V2 after bridge verbs, like Danish, Swedish and Norwegian, and finally, (iii) languages that allow V2 in both main and embedded contexts with no restrictions, like Yiddish, Icelandic and Faroese (Rögnvaldsson & Thráinsson, 1990; Santorini, 1989; Sigurdsson, 1990). The purpose of this section is to establish whether OCat presented asymmetries between main and embedded clauses that attributed to a V2 grammar. For this, we explore verb position frequencies in embedded clauses as well as the types of constituents occurring preverbally in embedded clauses. The subject position in embedded clauses is discussed at length in Chapter 4.

Examples (105) and (106) illustrate the difference in verb position in Modern German, an asymmetric V2 language:

(105) Den Mann **hat** der Hund gebissen.
 the man.ACC has the hound.NOM bitten

‘The dog has bitten the man’

(106) Jan **sagte**, dass der Hund den Mann gebissen hatte.
 Jan said.3SG that the.NOM dog the.ACC man bitten had.3SG

‘Jan said that the dog had bitten the man’.

Salvesen & Walkden (2016, p. 3)

(105) is a standard V2 main clause, where the pre-verbal constituent is the DO *den Mann* ‘the man’, in the accusative case. The verb appears in second position. In (106), however, the clause introduced by the subordinating conjunction *dass* ‘that’ has the finite verb in clause final position. This asymmetry is due to the fact that in main declarative clauses, the main verb raises to a projection in the left periphery that has an EPP feature that attracts a constituent to its specifier. In Modern German, this projection is ForceP, given the unavailability of constituents to occur to the left of the one fronted preverbal constituent.

Since in embedded clauses this projection is headed by the subordinating conjunction (in (106) *dass* ‘that’), the verb cannot raise to it, and it remains in TP, where it receives TAM features.

V2 analyses of the Old Romance languages show that the Old Romance languages differed in this aspect. OFr, for instance, has been described as an asymmetric V2 language, while OOc dialects present a divide: asymmetric in the north, and symmetric in the south. As for OSp, Fontana (1993, p. 71) proposes that OSp was a symmetric V2 language, while Wolfe (2015a, p. 130) argues that instead, it patterned with Danish and Norwegian in allowing V2 phenomena in embedded clauses depending on the predicate under which the clause was embedded, some predicates embedding clauses at the main point of utterance, where illocutionary force is conveyed and allowing V2 phenomena, and others below, truncating the left periphery of the embedded clause and blocking V2 phenomena from occurring. Asymmetries between main and embedded non-V2 clauses include: (i) a decrease in V2 orders, (ii) an increase in V1 clauses with null subjects, and (iii) frequent use of SVO order.

Old Romance varieties such as OFr (Vance, 1997) and OIt (Poletto, 2014) have been argued to present asymmetries between main and embedded clauses in regard to verb position. In OCat this is not the case. While there are some differences between main and embedded clauses, they do not seem to affect the frequency of verb position.

The frequency of different linear configurations in main and embedded clauses is shown in Table 13:

	V1	V2	V3<
Main Clauses	47.0% (470/1000)	44.5% (445/1000)	8.5% (85/1000)
Embedded Clauses	46.7% (467/1000)	50.3% (503/1000)	3.0% (30/1000)

Table 13 – Verb position in the LFRJ database main and embedded clauses

As in main clauses, V1 and V2 are the most common orders. However, in subordinate ones V2, accounting for 50.3% of clauses, is more frequent than V1, which accounts for 46.7%. V3 is the least attested order, less frequent in embedded clauses, accounting for 3.0%

And we replied to them that it pleased us very much.

Fol. 195v, 1 13

- (108) “Car **só** ferit”
because am wounded
‘Because I am wounded’.

Fol. 38v, 1 12

Embedded V2 clauses

- (109) E per açò no y calia juhiy, car [ja]₁
and for this not there.CL needed.3SG trial because already
era jutjat per nostres antecessors e ells lurs.
was judged by our predecessors and the theirs

And because of this there was no trial needed, since it had already been judged by our predecessors and theirs.

Fol. 195r, 1 22

- (110) E nós dixem-los que [la cosa]₁ **era** certa:
and we said.1PL=to;them.CL that the thing was true

And we told them that the matter was true: (...)

Fol. 195r, 1 25

Embedded V3 clauses

- (111) E açò se feÿa, car [denant nós]₁ [a
and this REFL.3SG= did.3SG since in front of us in
la riba]₂ **estaven** bé ·v· mília sarraïns;
the shore were.3PL well 5 thousand saracens

And this was being done since in front of us, in the shore, there were a strong 5,000 Saracens; (...)

Fol. 35r, 1 28

- (112) E Don Nuno respòs que [pus él no volia
and Sir Nuno answered.3SG that since he not wanted.3SG

sa amor]₁ que [él]₂ no **volia** la sua
 his love that he not wanted.3SG the his

And Sir Nuno answered that, since he did not want his love, that he didn't want his.

Fol. 9v, l 22

The parallelism between verb position frequencies in both types of embedded clauses, contrasts with comparable data from other Old Romance languages. Wolfe (2015a) and Sitaridou (2012, p. 574) find a strong asymmetry in the distribution of V1 clauses in OSp main and embedded clauses. The same is true of OFr and OIt. Under an asymmetric V2 analysis of the Old Romance languages, this asymmetry is accounted for by postulating that in embedded clauses, the verb does not raise to the left periphery. Instead, it remains within TP, in T°. Wolfe (2015a, p. 130) proposes that in these cases, it is T° that is endowed with an edge feature, triggering the movement of a nominal constituent to SpecTP. Therefore, in embedded clauses, (S)VO is the unmarked word order, and the preverbal field is specialised for subjects, which is not the case in main clauses. Data from the LFRJ database cannot be accounted for in this way. As it is explored at length in Chapter 4, the OCat preverbal field in both main and embedded clauses is specialised for hosting subjects, as it is also the case in MCat.

Another important consideration to make regarding the structure of embedded clauses in OCat is the availability of V3< word orders. Complement clauses are generally selected by bridge verbs such as ‘to know, to say, to beg’, predicates that select complements where the highest lexicalised projection is Force°, where the head of the clause, *que* is located. Complement clauses, cross-linguistically display main clause phenomena (Hooper & Thompson, 1973; Haegeman, 2003; Haegeman, 2009; Haegeman, 2010 and subsequent.), therefore, it is not surprising that they present a fully fledged left periphery where various constituents can occur preverbally. The same is true of some types of adverbial clauses,³⁸ as extensively discussed in (Haegeman, 2003, 2009, 2010, 2014). Since both types of embedded clauses have a fully fledged left periphery, with Topic and Focus projections, V3* orders are

³⁸ Haegeman distinguishes two types of adverbial clauses: peripheral adverbial clauses, merged in a higher projection of the main clause in which they occur and with a fully fledged left periphery; and central adverbial clauses, with a reduced left periphery, merged in a lower position within the main clause's structure. It seems that a great majority adverbial clauses from the LFRJ database belong to the former type. A full study of their syntax is left for future research.

expected (the different projections available in the left periphery of complement and adverbial embedded clauses is explored at length in Chapter 5).

So far, we have dispelled the possibility of OCat being an asymmetric V2 language, since embedded clauses from the LFRJ database behave in the same way as main clauses. Further evidence in favour of the verb being located in TP and not above it comes from adverb placement.

- (113) Car [_T anterior/past ja] [_{ASP} repetitive *altra vegada*] l' **avets** pres
 since already again it.CL= have.2PL taken
 per vostra dretura
 by yours righteousness

Since you have caught him yet again with your righteousness (...).

Fol. 29v, 1 23

Ja ‘already’ and *altra vegada* ‘again’ are a tense related adverb, and an aspectual adverbial expression respectively. What they tell us is that the verb is located below the CP layer, since it appears below TP adverbs.

Like in main clauses, the verb appears below TP adverbs, as shown in (113), but above VP adverbs, as in (114) and (115).

- (114) E respòs-vos (...) que **devíets** [bé] saber
 and answered.3SG=to.you.CL that should.2PL well know.INF
 lo nom de son seyor
 the name of his lord

And he answered that you should know the name of his lord.

Fol. 45v, 1 3

- (115) E féu-nos per tal que us **servíssem** [bé
 and made.3SG=to;us.CL so that to.you.CL serve.1PL.SBJV well
 e leyalment].
 and loyally

And (God) made us so that we would serve you well and loyally.

Fol. 29v, 1 3

Example (114) is especially helpful: the adverb *bé* ‘well’ occurs between the control verb *devevs* ‘ought to’ and the lexical verb *saber* ‘to know’. *Bé* is a relatively low VP adverb. What this shows is that the control verb has raised to a position outside VP, while the lexical verb remains in situ. The adverb shows that there has been verb movement from VP to a TP internal position, below CP.

A further piece of evidence that contributes to the identification of the position of the verb in embedded clauses is bipartite negation. The one instance of bipartite negation found in the LFRJ database is reproduced in (116). Bipartite negation, present across Gallo-Romance varieties, consists of a syntactic structure in which two polarity related heads are phonologically realised: the higher one, ΣP , the highest head of the TP, and a lower one, above the vP-VP cluster. In Modern French (MFr), the first head hosts the preverbal negative marker *ne*, while the postverbal negative marker *pas* is hosted in the latter. The fact that the verb is found above the low negative marker, tells us that the verb has left the vP-VP complex (Roberts, 1993, p. 49; Zanuttini, 1997, p. 88ff; Vance, 1997, p. 73; Cinque, 1999 p. 121; Salvesen & Bech, 2014).

- (116) (...) és obra de Déu, que no és pas nostre.
 is deed of God since not is step.NEG ours
 ‘(...) this is God’s doing, since it is not at all our (doing)’.

Fol. 37r, l 23

No cases of postverbal negation are found in main clauses.

In some Old Romance languages (OFr and OVen), asymmetries in the distribution of pro-drop subjects is strikingly asymmetrical between main clauses, where they are licensed, and embedded clauses, where they are not. This asymmetry has been explained by relating it to the oscillation in verb position between CP and TP in main and embedded clauses. This matter is addressed at length in Chapter 4.

3.3 SUMMARY

In this Chapter, the position of the verb in OCat main and embedded clauses has been explored, allowing us to reach the conclusion that OCat was not a V2 language, the verb remained in TP in unmarked main and embedded clauses. This has been done by establishing the absolute verb-initial condition of V1 clauses, with the exception of the case of unaccusative verbs, which in clauses presenting VS orders, involve a null *pro*_{LOC} element in SpecTP. This element, like subjects occurring in SpecTP, would ground the clause to the

discourse. Oscillations between enclisis and proclisis have been shown to interact with the saturation of FocP, rather than with the raising of the verb in the left periphery. The position of the verb has been identified as being in the inflectional domain has been confirmed by using Cinque (1999) adverb hierarchy. A context which involves verb movement to the left periphery: the expression of emphatic positive polarity. Finally, it has been shown that verb position is parallel in main and embedded clauses. On the grounds of these findings, we establish that OCat did not present a V2 grammar, and that instead of aligning with neighbouring Old Romance varieties such as OOc and OFr, it aligns with MCat, an SVO language, in regard to verb position. A final important conclusion that can be drawn from the data presented in this chapter is that OCat seems to be innovative when compared to the other Old Romance varieties, in that by the 13th century, *sí* had already been reanalysed as an EPPA, a century earlier than in OSp.

4 Subject Position in Old Catalan

Subject position is intimately linked to verb position. Depending on the position of the verb in the clause, different word order patterns are expected. For instance, in V2 languages, by virtue of the verb raising to a projection in the left periphery with an EPP feature that attracts a maximal projection to its specifier, XP V strings where XP need not be the subject are expected, unlike in SVO languages, where the subject is the canonical preverbal constituent. This is often referred to in the V2 literature as ‘non-specialisation of the preverbal field’. A consequence of the non-specialisation of the preverbal field in V2 language is Germanic-style inversion, whereby, when a constituent other than the subject raises to the left periphery, the subject remains in SpecTP, verb-adjacent. In Chapter 3, it has been shown that the OCat’s verb in unmarked declarative main and embedded clauses was located within the inflectional domain. Therefore, the distribution of subjects is expected to pattern with that of SVO languages like MCat, and not with V2 languages. In section 4.1, it is shown that OCat was a null subject language like MCat. In section 4.2, considerations on MCat subject position are provided, so as to establish a standard of comparison for the OCat description that follows. In section 4.3, a description of the distribution of OCat subjects in main clauses is brought forth, as well as a discussion on the nature of OCat’s SpecTP. In section 4.4, the distribution of subjects in embedded clauses is described and shown to be parallel to that of main clauses. Finally, in section 4.5, the findings of this Chapter are summarised.

4.1 Old Catalan: a null subject language

OCat was, like its modern counterpart, a Null Subject Language (NSL). NSLs can have null subjects, *pro*, in finite clauses. Rizzi (1982, p. 142) formulated the parameter that regulates this possibility as follows:

- (1)
 - a. INFL can be specified [+ pronoun].
 - b. INFL which is [+pronoun] can be referential.

Rizzi (1982, p. 142)

A lot of work has followed this first formulation of the Null Subject Parameter (NSP), creating more nuanced typological classifications of languages depending on the relationship overt verbal morphology with the licensing of null subjects (Biberauer, Holmberg, Sheehan, & Roberts, 2010; Biberauer & Roberts, 2008; Koenenman & Neeleman, 2001, to mention a

few). Work on the NSP has also been tightly linked to the debate around the existence and the nature of *pro* and the existence of a subject-dedicated A-position in SpecTP linked to an EPP feature (see Rosselló, [2001]; Solà [1992], for accounts that dispense with the existence of an EPP requirement in MCat, which take preverbal DPs as non-argumental and left dislocated, Alexiadou & Anagnostopoulou [1998]; Ordóñez & Treviño, [1999], for accounts against the existence of *pro* in NSLs, and Costa [2001, 2008] for an analysis that favours it). Following Cardinaletti (2004); D'Alessandro (2014); Holmberg (2005); Rizzi (1982, 2006, 2018); Roberts (2010); Sheehan (2005), among others, it is assumed that *pro* exists and that its availability is one of the defining features of NSLs. The assumed position of *pro* for NSLs depends on the features linked to the Extended Projection Principle which requires elements with matching featural makeup to appear in SpecTP.

Biberauer et al. (2010) and Biberauer & Roberts (2008) present a four-fold typology of languages in regard to the NSP, built around the richness and strength of two types of features found in TP: tense and agreement, the former responsible for V-to-T movement, and the latter for the licensing of null subjects. The four types of languages are the following:

(2)

- a. Rich agreement and rich tense inflection: hence V-to-T and null subjects, e. g. Italian, Greek, Spanish.
- b. Poor agreement but rich tense: hence V-to-T but no null subjects, e. g. French, Middle English.
- c. Poor tense and poor agreement: hence no V-to-T and no null subjects, e. g. Modern English, Mainland Scandinavian.
- d. Rich agreement and poor tense: null subjects, but no V-to-T; no clear example.

Adapted from Biberauer & Roberts (2008, p. 27)

OCat patterns with type (a) languages: it has rich agreement and tense inflection. Furthermore, Sheehan (2015, p. 1)³⁹ lists other surface features linked to null subject languages (group (a) in Biberauer et al. [2010]; Biberauer & Roberts [2008]), such as the obligatoriness of (non-referential) null expletive subjects and free inversion. These features

³⁹ Other phenomena connected to the null subject parameter, such as a violation of the that-trace filter, cannot be readily described or analysed in OCat, given the unavailability of native speakers.

are also found in OCat, contrary to what has been found in other Romance languages (see Poletto [2005] for OIt, Mathieu [2006] for OFr). OCat had null expletive subjects, non-referential and referential null subjects, as shown in (3-6):

- (3) Context: Les gèns d'aqueles terres que dessus havem dites vengren a nostre pare e dixeren-li que él podia ésser seyor d'aqueles terres (...).

And the people from those regions that we have mentioned above, came to my father and told him that he could be lord of those lands.

E d'una part, li ó **daven** de paraula (...)
and from;one side to;him.CL=it.CL= gave.3PL of word

On the one hand, they gave it to him by word (...).

Fol. 4v, l 14

- (4) Context: E descavalcà Don Exemèn Lópeç de Rigols e dix a Don Pero Ahonés que cavalcàs en lo seu caval, que·l seu era cansat, e que pensàs d'estorçre.

And Sir Exemèn Lópeç of Rigols dismounted his horse and told Sir Pero Ahonés that he should ride his horse, because he should think of saving himself.

E mentre que açò faïen, **gitaven** pedres grosses e poques (...)
and while this did.3PL threw.3PL stones big and few

And while they did this, they threw big and small stones (...).

Fol. 15r, l 26

- (5) 'e és mester que vós fassats tals obres ab
and is need that you do.2PL.PRS.SBJV such work with
nostra ajuda (...).
our help

And it is necessary that you do such work with our help.

Fol. 29r, l 11

- (6) e **féu** tanta de mar, que (...)
and did.3SG so much of sea that

And the sea was so rough, that (...).

Fol. 33r, l 21

In (3), the 3rd person plural subject is coreferent with a 3rd person plural DP previously introduced in the discourse: *les géns d'aquelles terres* 'people from those lands'. In (4), again, the referent of the 3rd person plural subject of the verb can also be recovered from the context. These examples contrast with (5) and (6), two cases of non-referential null expletive subjects. In (5), the expression *és mester que* 'it is necessary that', occurs with a non-referential 3rd singular null subject. In (6), the subject of the verb *féu* 's/he did' is null and non-referential, as expected in weather verbs in a null subject language, such as the verb 'to rain', in MCat, as in (7):

- (7) Plou.
 rains
 It rains.

MCat

OCat has no overt expletives.⁴⁰ This, together with the fact that OCat has rich tense and agreement inflection, proves that OCat was a null subject language of type (a) in the sense of Biberauer et al. (2010) and Biberauer & Roberts (2008), like its Modern counterpart.

Apart from licensing non-referential null expletives, Sheehan (2015, p. 2) also lists free inversion as one of the features associated to NSLs of type (a). OCat allowed free inversion. This phenomenon can have the same surface expression as Germanic verb-subject inversion, which is a surface effect of verb movement to the left periphery. In Chapter 3, it has been concluded that in OCat, the verb was located in the inflectional domain in unmarked declarative clauses, and predicates prone to occur in V1 configurations have been discussed at length, dismissing the possibility of null elements triggering Germanic inversion in these cases. In Chapter 5, further arguments against the existence of this phenomenon in OCat are presented, definitely discarding the occurrence of this phenomenon in OCat unmarked declarative clauses.

V-to-T movement takes place not only in main, but also in embedded clauses, as shown in Chapter 3, section 3.2. As a consequence, the licensing of null subjects is parallel in both clause types, as it is shown in the subsequent sections. Table 1 displays the frequency of overt and null subjects in main and embedded clauses in the LFRJ database:

⁴⁰ See Hinzelin (2006, 2010) for an analysis that shows that OCat neuter pronouns were used as discourse markers, rather than expletives and Pujol i Campeny (2018) for a dismissal of *si* as a potential expletive.

	Main clauses	Embedded clauses
Null subjects	532 (53.2%)	521 (51.1%)
Overt subjects	468 (48.6%)	479 (47.9%)
	1000 (100%)	1000 (100%)

Table 1 – Null and overt subjects in main and embedded clauses

Table 1 shows that there is no asymmetry in the frequency of overt subjects in main and embedded clauses, overt subjects being 0.7% more frequent in main than embedded clauses, an unexpected result since overt subjects are, as it is discussed below, more frequent in embedded clauses in neighbouring Old Romance varieties. Thus, OCat was a symmetric pro-drop language and differed from asymmetric pro-drop Old Romance languages (Adams [1987]; Roberts [1993]; Vance [1997 p. 182]; Labelle [2007]; Vance et al. [2009]; Sitaridou, [2004]; Steiner [2014]; and Salvesen & Walkden [2015] for OFr, Poletto [2013] for OIt, Wolfe [2015] on several Old Romance languages, including OSp).

A closer look at the distribution of overt subjects in main and embedded clauses provides further evidence for a symmetric analysis of OCat main and embedded clauses in terms of the NSP:

	Nominal	Pronominal	Total
Preverbal	110/259 (43%)	149/259 (57%)	259 (55%)
Postverbal	159/209 (76%)	50/209 (24%)	209 (45%)
Total	269 (57%)	199 (43%)	468 (100%)

Table 2 – Frequency of postverbal and preverbal overt subjects in main clauses (out of all clauses with overt subjects)

	Nominal	Pronominal	Total
Preverbal	155/377 (41%)	222/377 (59%)	377 (79%)
Postverbal	76/102 (75%)	26/102 (25%)	102 (21%)
Total	231 (48%)	248 (52%)	479 (100%)

Table 3 – Frequency of postverbal and preverbal overt subjects in embedded clauses (out of all clauses with overt subjects)

Tables 2 and 3 show the frequency of preverbal and postverbal nominal and pronominal subjects in main and embedded clauses respectively. While the frequency of nominal and pronominal subjects in both clause types is parallel, the frequency of preverbal and postverbal subjects is not. Preverbal subjects are significantly more prominent in embedded clauses, where they make up 79% of overt subjects, while in main clauses they make up 55%. In both cases, the most frequent type of preverbal subjects are pronouns. Nevertheless, the proportion of postverbal nominal and pronominal subjects in main and embedded clauses is almost the same: in main clauses, 75% of postverbal subjects are nominal, while 25% are pronominal, while in embedded clauses, 76% of postverbal subjects are nominal and 24% pronominal. This suggests that there are no differences in the licensing of nominal and pronominal subjects in main and embedded clauses and provides evidence for establishing that OCat was a symmetric pro-drop language. The position of postverbal subjects in main and embedded clauses is described and studied at length in section 4.3 in this Chapter.

Another piece of evidence in favour of a symmetric analysis of pro-drop in OCat is that overt subjects in complement clauses introduced by *que* ‘that’ trigger disjoint reference effects, as expected in an NSL such as MCat or MIIt (Franco, 2009, p. 192). This is illustrated in (8) and (9):

- (8) Context: E aquest apostoli, papa Innocent, fo el meylor apostoli, que de la saó que faem aquest libre en ·c· anys passats no hac tan bo apostoli en la Església de Roma. Cor él era bon clergue en los sabers que taynen a apostoli de saber, e havia sen natural, e dels sabers del món havia gran partida.

And this Pope_i, Pope Innocent, was the best pope, because during the period in which we are making this book, in the last 100 years there has not been such a good Pope in the Church of Rome. Since he was a good clergyman in the knowledge that had to be

known to a pope, and he had natural common sense, and he had a great part of worldly knowledge.

E emvià tan forts cartes e tan forts misatge
 and *pro*_i sent.3SG such strong letters and such strong messages
 al [comte Simon]_j, que él_i **hac** a atorgar
 to;the count Simon that he had to grant.INF
 que·ns retrie a nostres hòmens.
 that;to.us.CL= return.3SG.SBJV to our men

And he (*the Pope*) sent such great letters and great messengers to count Simon, that [he] (*count Simon*) had to return us to our men.

Fol. 5v, l 14

- (9) Context: E quan los frares viren que [·l comte de Proença]_i se n'era anat, sí que no·ls ho féu saber, enteseren que la nostra estada no·ls era bona. E al comte Don Sanxo, quan ho hoí, pesà-li molt la anada del comte de Proença.

And when the friars saw that the count of Provence had left, even though he had not let them know, they understood that our stay was not good for them. And when count [Sir Sanxo]_i heard about it, he was very sad about the leaving of the count of Provence.

E, quan él_i **entès** que él_i se n'
 and when he understood.3SG that he REFL.CL=from.there.CL=
era **anat** ab aquels qui eren del seu
 was.3SG gone with those who were.3PL of;the his
 bando d'Aragó (...).
 side of;Aragon

And when he (*the count Sir Sanxo*) understood that he (*the count of Provence*) had left with those Aragonese men that sided with him (...).

Fol. 6v, l 23

In this section, it has been established that OCat was an NSL with very similar frequencies of null and overt subjects in main and embedded clauses, as well as the same frequency of preverbal and postverbal pronominal and nominal subjects.

4.2 Subject distribution in Modern Catalan main clauses

It is widely accepted that in M_{Cat}, IS is mostly encoded syntactically⁴¹, while its prosody remains rather static: the primary sentence accent falls invariably on the last element that can bear stress in the core of the clause, marking its boundary. It is also generally accepted that IFoc is to be contained in the core of the clause and old information is to be dislocated to the right or the left (Vallduví, 1992; Forcadell, 2013). However, there are discrepancies regarding the canonical word order of M_{Cat}, specifically around subject position. While Vallduví (1994, 1995, 2002), Solà (1992) and Bonet (1990) have defended that it is a VOS language, Forcadell (2013) and Gallego (2013) advocate that it is, in fact, SVO. In this section we offer (i) an overview of Vallduví's VOS underlying word order hypothesis, and (ii) an overview of Forcadell's SVO hypothesis.

4.2.1 Modern Catalan as a VOS Language

According to Vallduví (1994, 1995, 2002) and Vallduví & Engdahl (1996), the most frequent word order of a language is generally considered as its canonical order. However, they defend a more thorough linguistic analysis (beyond frequency rates) to establish what the canonical order of a language is. In the case of M_{Cat}, this translates into the following: although the most frequently attested word order is SVO, the fact that preverbal subjects can be separated from the verb by interjections and *wh*-words leads him to conclude that M_{Cat} is in fact VOS.

His argumentation departs from the fact that, in M_{Cat}, thematic material is dislocated to the right and left edges of the clause, leaving a clitic pronoun to retake the dislocated element in the core of the clause. M_{Cat} has two sets of personal pronouns: strong and weak ones, as we can see in Table 1, weak ones being used to retake dislocated elements and having allomorphs depending on the form of the verb they attach to, as well as the phonological environment:

⁴¹ However, in Forcadell & Llopis (2014), the emergence of the possibility of marking IS by means of prosody is confirmed: during the last 30 years, there has been an increase in the use of prosody to signal IS, partly due to the contact with Spanish and English, languages in which prosody is the main means to convey IS (Leonetti & Escandell-Vidal, 2009b, sec. 2.1).

		Singular		Plural	
		Strong	Weak	Strong	Weak
Nominative	1 st person	Jo	-	Nosaltres	-
	2 nd person	Tu	-	Vosaltres	-
	3 rd person	Ell/ella	-	Ells/elles	-
Accusative	1 st person	A mi	Em/me/m'/'m	A nosaltres	Ens/nos/'ns
	2 nd person	A tu	Et/te/t'/'t	A vosaltres	Us/vos
	3 rd person	A ell/ella	lo/la/l'/'l	A ells/elles	Los/les/'ls
Dative	1 st person	A mi	Em/me/m'/'m	A nosaltres	Ens/nos/'ns
	2 nd person	A tu	Et/te/t'/'t	A vosaltres	Us/vos
	3 rd person	A ell/ella	Li	A ells/elles	Los/'ls

Table 4 – M-Cat Personal Pronoun Clitic Paradigm

However, the paradigm is defective: nominative pronouns lack clitic counterparts. This is illustrated in examples (10-12). In (10.a) we find the strong pronoun *ell* and in (10.b) we do not find an overt realisation of the subject. This is not the case with direct objects, as shown in (11.a) and (11.b) which, when dislocated, leave a clitic copy in the core of the sentence. Vallduví explains this asymmetry between nominative/accusative-dative in the Catalan clitic paradigm, by proposing that null subjects are the nominative equivalent to clitics. There is, nevertheless, an exception: indefinite subjects which are replaced by the clitic *en*, as we can see in (12.a) and (12.b). In (12.a), the indefinite noun *trens* “trains” appears in the expected postverbal position for unaccusative verbs, however, in (12.b), it is substituted by the clitic pronoun *en* (literally, it can be translated as ‘of it’).

- (10) a. Ell **vindrà** demà.
he come.3SG.FUT tomorrow
He will come tomorrow.
- b. **Vindrà** demà.
come.3SG.FUT tomorrow
(He) will come tomorrow.

- (11) a. **Vull** la poma.
 want.1SG the apple
 I want the apple.
- b. La_i **vull**, [la poma]_i.
 it.CL.ACC.FM.SG= want.1SG the apple.
 I do want the apple.
- (12) a. **Venen** trens.
 come.3PL trains
 There are trains coming.
- b. En **venen**.
 of them.PTCP.CL come.3PL
 There come (trains).

As we have said above, whenever a constituent is dislocated (be it a DO taking the accusative case or an IO taking the dative), the presence of a retaking clitic in the core of the clause is mandatory, as we can see in (13) and (14), left and right dislocations respectively. Furthermore, (15) shows the object appearing *in situ* and allows us to conclude that there are three grammatical positions for the object to appear: left dislocated, right dislocated and *in situ*.

- (13) [La poma]_i, [la]_i **tinc** a la bossa.
 the apple it.CL.ACC.FM.SG= have.1SG in the bag
 The apple, I have it in my bag.
- (14) [La]_i **tinc** a la bossa, [la poma]_i.
 it.CL.ACC.FM.SG= have.1SG in the bag the apple.
 I have it in my bag, the apple.
- (15) **Tinc** la poma a la bossa.
 have.1SG the apple in the bag
 I have the apple in the bag.

As it is discussed in Chapter 5, in the case of subjects, due to the lack of clitic nominative pronouns, it is not possible to identify dislocations as readily as in the case of DOs and IOs. This is illustrated in (16-19). In example (16), the subject occurs in SpecTP. (17) contrasts with

(16), in that the subject is left-dislocated (the comma between the subject and the verb indicates a prosodic pause). In (18), the subject occurs in its base-generated postverbal position, contrasting with (19), where it is right dislocated. In this case, the dislocation is also prosodically marked: it occurs after the core sentential stress, which naturally falls on the last constituent of the clause, in this case, *ahir* ‘yesterday’.

(16) La Quimeta va venir ahir.
 the Quimeta came.3SG yesterday

Quimeta came yesterday.

(17) La Quimeta, va venir ahir.
 the Quimeta came.3SG yesterday

It is Quimeta that came yesterday.

(18) Va venir la Quimeta ahir.
 came.3SG the Quimeta yesterday.

Yesterday that she came, Quimeta.

(19) Va venir ahir, la Quimeta.
 came.3SG yesterday the Quimeta.

Yesterday she came, Quimeta.

While, as stated above, SVO is the most frequent word order pattern for M_{Cat} declarative clauses, VOS and VSO are also attested. The former is rarer, but not unusual or ungrammatical, according to Vallduví. Example (20) reproduces one of Vallduví’s (2002, p. 1247) examples of VOS:

(20) Toca el piano ma germana.
 plays the piano my sister

My sister plays the piano.

Therefore, there are four positions available for subjects: preverbally in SpecTP, postverbally, in its base-generated position, and right and left dislocated.

However, the fact that there are only 3 positions available for objects, while there are 4 positions available for subjects, creates an asymmetry that is the cornerstone of Vallduví’s proposal that M_{Cat} is a VOS language. He tries to solve the asymmetry in the distribution of objects and subjects by proposing that all preverbal subjects in M_{Cat}, are, in fact, left

dislocations, which do not look as such due to the fact that subjects do not leave a clitic in the core of the clause to be retaken, given the presence of verbal morphology (Alexiadou & Anagnostopoulou, 1998). This makes the distribution of objects and subjects symmetric, and suggests that M_{cat} is a VOS language, since only postverbal non-right dislocated subjects are *in situ*. He proves this hypothesis by providing tests to show that the subject is dislocated, in spite of not having a coreferent clitic in the core of the clause. Firstly, he shows that the interjections *tio* and *xec*, both meaning ‘mate, dude’ (which can be translated as ‘man’ or ‘dude’ in English) can appear between dislocated constituents and the verb, while they cannot be inserted in the core of the clause. Furthermore, dislocated constituents can also precede interrogative words, while this should be ungrammatical if the subject belonged to the core of the clause:

- (21) La Sandra, tio, no (*tio) **menja** pomes.
 the Sandra man not man eats apples.
 Sandra doesn’t eat apples, man.

- (22) La Sandra, com que no **menja** pomes?
 the Sandra how come not eats apples
 How come Sandra doesn’t eat apples?

Thus, he concludes that M_{cat} is a VOS language and that the greater frequency of SVO is not due to it being the canonical word order, but to the fact that there is a greater preference for X-V-Y clauses than V-X-Y, where X does not necessarily have to be the subject (Vallduví 2002, p. 1247).

It is worth noting that these tests simply show that a subject *can* be left dislocated, not that it always is. Consider (23):

- (23) A: Què passa?
 what happens
 B: La Sandra (*tio) no **menja** pomes, tio.
 the Sandra man not eats apples man
 A: What’s happening?
 B: Sandra doesn’t eat apples, man.

In (23), the out of the blue answer to a ‘What’s happening?’ question, does not admit the interjection between the subject and the verb. Nothing, apart from negation, can intervene between them. This casts doubt over the validity of Vallduví’s test for left dislocation and shows that (21) and (22) are merely a minimal pair between a clause with a left dislocated subject and a one with the subject in SpecTP, the former not being a suitable out-of-the-blue answer, and the latter being perfectly adequate as such. Therefore, this adds evidence against Vallduví’s dismissal of SpecTP as a subject position. This also adds further evidence in favour of Forcadell (2013) proposal that M_{cat} is indeed an SVO language.

4.2.2. Modern Catalan as an SVO Language

Forcadell (2013) argues against Vallduví’s proposal and defends that M_{cat} is an SVO language, by offering information structure arguments, as well as alternative analyses of Vallduví’s data.

In terms of information structure, Forcadell states that the preverbal subject position is vital for the encoding of all-focus, a structure long associated with canonical word order:

(24) [You know what?]

La	Marta	adopta	un	nen.
the	Marta	adopts	a	child.

Marta is adopting a child.

Adapted from Forcadell (2013, p. 43)

Furthermore, Forcadell argues that the postverbal subject position is inherently linked to a narrow focus reading. If that were the case, in a VOS order where the focus would be placed on the subject, and according to Vallduví’s own description of the distribution of rhematic and thematic information, the object, being thematic, should be dislocated, since the core of the clause is only to host rhematic information. Thus, (25) is deemed bizarre, if not ungrammatical by Forcadell, and (26) is offered as the grammatical alternative, with narrow focus on the subject:

(25) ??	Adopta	un nen	la	Marta.
	adopts	a child	the	Marta

Marta adopts a child.

Adapted from Forcadell (2013, p. 43)

- (26) [L]_i'**adopta** la Marta, [un nen]_i.
 he.CL=adopots the Marta a child

It is Marta, who is adopting a child.

Adapted from Forcadell (2013, p. 43)

Forcadell then proceeds to describe in which cases subjects are grammatically allowed to appear postverbally, limiting the possibilities to specific linguistic contexts: (i) with weak direct objects, (ii) in announcement contexts, (iii) with corrective/contrastive focus, (iv) with pure intransitive verbs, (v) in all-focus contexts.

The reasoning lying behind the possibility of having VOS with certain light verbs is that they are normally combined with weak objects, which can be incorporated into VP to a certain degree or which add little meaning to the clause. Therefore, phrases like *parar la taula* “to lay the table” or *fer petons* ‘to kiss’, should not be taken as VO but as V (especially in the case of *parar la taula*, which exists in the form *parar taula*, where the noun does not present the determiner *la*, indicating further grammaticalisation). In (27), *la lliga* further specifies what Barça will win. However, in most conversations, the fact that it is the league what Barça will win would be taken for granted, and therefore, the noun phrase *la lliga* does not add much information. This is illustrated in examples (27-29):

- (27) **Guanyarà** la lliga el Barça.
 win.3SG.FUT the league the Barça

Barça will win the league.

Adapted from Forcadell (2013, p. 44)

- (28) **Guanyarà** el Barça.
 open.3SG.FUT the Barça

Barça will win.

- (29) **Para** taula en Pere.
 lays table the Pere

Pere lays the table.

From Forcadell (2013, ft. 28)

In contexts where there is contrast expressed, usually combined with the ‘announcement’ feature, VOS is also allowed. In (30), it is made clear in the announcement of

who will open the conference that it will be Dr Brown, cancelling all other options or correcting wrong hypotheses.

- (30) **Inaugurarà** la conferència el Dr Brown.
 open.3SG.FUT the conference the Dr Brown

Dr Brown will inaugurate the conference.

In the case of purely intransitive verbs, like unaccusatives, postverbal subjects are preferred, as shown in (31) and (32):

- (31) Bufa el vent.
blows the wind

The wind is blowing.

- (32) Ara **arriben** mons pares.
 now arrive my parents

Our parents are arriving now.

Finally, Forcadell states that all-focus sentences, if also exclamative, can present VOS provided that the subject receives extra emphasis in the already very emphatic context. The example she gives is reproduced in (33)⁴².

- (33) [What happened?]
 (Que) [al nen]i, li₁ **ha** fet un petó
 that to;the child to;him.CL= has made a kiss
 aquell BORRATXO!
 that drunkard

That drunkard has kissed the child!

Therefore, Forcadell (2013) brings two types of arguments in favour of an SVO analysis of M_{Cat}: on the one hand she dismisses VOS orders as grammatical in unmarked M_{Cat} main clauses, and on the other, she describes the specific contexts in which V(O)S orders are

⁴² The author, as well as 5 other Easter Catalan native speakers consulted, find this example completely ungrammatical under any reading, which makes us doubt of (i) the narrow focus condition of postverbal subjects in M_{Cat}, as well as (ii) their acceptability in declarative clauses in contexts other than psychological, unaccusative verbs.

acceptable, including all-focus clauses, exclamations, constructions with unaccusative predicates and constructions with light predicates.

4.2.3 Considerations about Subject Position in Modern Catalan

Unfortunately, neither of the approaches above account for speakers' judgements. In this section we aim to disprove Vallduvi's VOS hypothesis and argue that M_{Cat}, like its old counterpart, is an SVO language with an EPP feature in SpecTP, that might be filled either by a nominal expression or by a locative, following the proposal made by Sheehan (2010); Sifaki (2003); Tortora (2001), bringing it together with Rigau (1993) analysis of existential constructions in M_{Cat} (see Chapter 4, section 4.2.2 for a discussion of existential constructions in O_{Cat}).

Vallduvi's (1994, 1995, 2002) hypothesis that the position of subjects in M_{Cat} is sentence-final is based on two main points: (i) the fact that the distribution of subjects should be parallel to that of objects, and (ii) the fact that subjects cannot be substituted by a clitic counterpart. Firstly, there is no robust cross-linguistic evidence supporting the fact that the distribution of objects and subjects should be symmetric, while it seems to be common to have specialised positions for both the internal and the external argument. Secondly, the fact that subjects do not have clitic counterparts, again, does not seem to be a stable feature throughout languages that have object clitics.

Regarding the overwhelming presence of preverbal subjects, we can posit the existence of a strong EPP feature in M_{Cat} that can be either checked by *pro*, by an overt nominal or pronominal subject, and potentially, by a locative in existential clauses (along the lines of Sheehan (2010)'s analysis of free subject inversion in Spanish).

And thirdly, his examples have proved to be ungrammatical to a pool of 5 native speakers, along the lines of Forcadell's analysis. Even in contexts where Forcadell justifies the presence of postverbal subjects, such as in exclamative clauses as in (33), informants have homogeneously judged those as ungrammatical judgements.

Based on the reasonings exposed above, we assume that M_{Cat} is linearly an SVO language, that allows (and sometimes favours) VS orders with certain predicate types (including unaccusatives as well as *verba dicendi*), in all-focus contexts, as illustrated in (26-28), and in Focus Fronting constructions.

4.3 Subject distribution in Old Catalan main declarative clauses

In this section, we describe the positions available to OCat subjects in declarative main clauses, and the distributional properties of different subject types. Firstly, we focus on OCat preverbal and postverbal subjects, including nominal subjects, pronominal subjects and null subjects, identifying which positions are available to them. It is shown that (i) OCat data supports views that defend that NSLs have a position available to subjects in SpecTP; that (ii) elements other than lexical and pronominal subjects can occur in SpecTP, mainly locatives (overt or null); and that (iii) the informational value and heaviness of the subject interact with the predicate type determining its position.

Together with Cardinaletti (2004); Sheehan (2010, 2015) and Gallego (2007), Rizzi (2015, 2018), among others, we defend that the Modern Romance languages have a preverbal projection whose featural makeup involves an EPP feature that triggers the movement of an element to it that can satisfy the featural makeup of the projection, a position that is often (but not always, as we shall see) targeted by subjects. Other authors defend that preverbal subjects can only be focal, or clitic left dislocated (Solà, 1992; Vallduví 1993, 2002; Alexiadou & Anagnostopoulou, 1998; Ordóñez, 1998; Rosselló, 2000). Gallego (2007) argues in favour of the presence of a preverbal projection specialised for hosting subjects, showing that all preverbal subjects need not be left dislocated. Gallego dismisses the hypothesis of preverbal subjects in NSLs being Left Dislocated by providing a minimal pair in MCat, where it can be seen that there is a difference between Left Dislocated subjects, located in a projection within the CP layer, and those merged in SpecTP⁴³:

(34) La Maria **ha** plorat.
the Maria has cried
Maria has cried.

(35) La Maria_i, t_i **ha** plorat.
the Maria has cried
As for Mary, she has cried.

MCat

Adapted from Gallego (2007, p. 51)

⁴³ 5 native speakers have certified that the difference between both sentences is clear and prosodically marked.

Gallego (2007) states that (34) and (35) differ substantially, but he does not specify which one it is. However, the difference between (34) and (35) is prosodically clear for native speakers and it can be further characterised by the questions that could elicit (34) and (35) as answers: in (34), *la Maria*, a definite referential subject, is located in SpecTP in a clause that could be all-focus, answering to the question ‘what’s happened?’, whereas (35) could be an answer to ‘what’s up with Mary?’, which requires Mary to have already been introduced to the discourse.

Furthermore, Gallego associates SpecTP with interpretation effects at clause level: while preverbal subjects in SpecTP receive a Categorical Interpretation, by which something is predicated of them, DP subjects elsewhere receive athetic interpretation, by which the clause lacks a new/old information divide, and licenses only all-focus reading. This fact is relevant for the analysis of preverbal subjects in OCat, a matter that is considered in the following paragraphs.

Sheehan (2015) uses the following tests to determine whether SpecTP is an A position in Modern Romance or not: (i) adverb placement, (ii) wide/narrow scope of preverbal subjects, (iii) binding of postverbal subjects, (iv) non-referential subjects, (v) floating quantifiers, (vi) subjects vs. topics, (vii) hortative contexts, (viii) basic word order, (ix) disambiguation, and (x) parasitic gaps. Contrasting those views that defend that agreement morphology is pronominal and that all preverbal subjects are left dislocated (Solà, 1992; Alexiadou & Anagnostopoulou, 1998; Rosselló, 2000 among others), and those that assume that agreement is the spell out of uninterpretable features (Holmberg, 2003, 2004; Roberts, 2004), Sheehan (2015, p. 9) concludes that at least some preverbal subjects in MIIt, MSp, Modern Romanian and MPt are likely to be located in an A-position, therefore favouring analyses that advocate for the existence of a preverbal A position linked to the satisfaction of EPP that would attract preverbal subjects to SpecTP. This stance is also supported in Rizzi & Shlonsky (2006) and Rizzi (2015, 2018), where the preverbal position is described as a position in high TP which ‘expresses the argument ‘about which’ the event expressed by the predicate is presented’ (Rizzi, 2018, p. 510). According to this author, this position is a halting position for A-movement: once a constituent moves here, it cannot move further up within the clausal structure. Halting is associated to criterial positions (i.e. positions linked to IS values). This is also the case of SpecTP, which is

associated to an ‘aboutness effect’ (Rizzi, 2018, p. 514).⁴⁴ What differentiates this position from Topic positions, which have often been connected to the notion of aboutness, is ‘givenness’. While topics need to be connected to a contextually given set established in the discourse, elements occurring in SpecTP do not. They are to be interpreted as the argument which the predicate is about. This does not impede subjects to occur in the left periphery in a topic position: they can bind *pro* in SpecTP when, in addition to being the argument which the predicate is about, they are also given (Rizzi, 2018, p. 527).

Rizzi’s (2018) approach is in line with Cardinaletti (2004) proposal of a cartography of SpecTP. Cardinaletti proposes that SpecTP actually contains three heads linked to the checking of different features associated with SpecTP: SubjP, EPPP and AgrP. SubjP is associated with the checking of a *subject of the predication* feature; EPPP, which requires a constituent to move to it to check the Extended Projection Principle; and AgrP, which is associated to grammatical features (subject to being split further into different heads depending on the range of features that it hosts, Cardinaletti [2004, p. 135]). The checking of the *subject of the predicate* feature is linked with the clause receiving a categorial judgement: something is being predicated of the subject, and the clause has bipartite information structure, similar to Rizzi’s ‘argument ‘about which’ the predicate is presented’. When the *subject of the predicate* feature is not checked, the clause isthetic judgement reporting an event (Cardinaletti, 2004, p. 151) and are linked to all-focus.

Thetic judgements are crosslinguistically linked to VS orders (see Petrova & Hinterhölzl [2009] for views on VS in Germanic languages, Lahousse [2011] for Modern French, Leonetti [2017] for an overview of word order in the Romance languages), where the subject remains *in situ*, and therefore, does not check the *subject of the predicate* feature in Cardinaletti’s terms, or where the subject does not raise to the position associated with ‘aboutness’, in Rizzi’s terms. If VS orders are indeed linked to all-focus thetic readings, it is not surprising for pronouns to be less frequent in postverbal position in both main and embedded clauses (see Tables 2 and 3), since they are inherently topical and referential to elements that are already active in the discourse or in the discourse’s background.

In cases where the subject remains in postverbal position, we have suggested in Chapter 3, section 3.1.2 that a stage topic (overt or covert) can occupy its position. Along

⁴⁴ As pointed out by Rizzi (2018), a test that confirms this association is passivisation, which shifts aboutness from the subject to the object by turning it into the subject.

these lines, (Tortora, 2001) and Sifaki (2003) propose the existence of a null locative that would satisfy EPP in SpecTP without being able to check case, person and number features (a hypothesis that has already been discussed at length in relation to predicate types in Chapter 3, section 3.1.2). This brings Cardinaletti (2004) to propose the presence of one further projection in the Subject field, consecrated to EPP. EPP has to be checked: it is this projection that attracts an argument to the left periphery, given that AgrP does not require movement value its features (Cardinaletti, 2004, p. 152), and that the *subject of the predicate* feature only needs to be checked if a clause is to receive categorical reading.

In (36.a) we find a thetic clause with an unaccusative predicate, where the *subject of the predicate* feature does not have to be checked, and therefore, the subject remains postverbal. Nevertheless, EPP needs to be satisfied and it attracts a null locative, *pro*_{Loc} in Tortora (2001) terms in the case of unaccusative verbs. This null locative refers to the ‘here and now’ of the communication act, akin to stage topics. In (36.b), we have an example of a thetic clause with a transitive verb, with a preverbal subject. According to Cardinaletti, while the subject is preverbal, it does not raise to SubjP, and therefore, it does not check the *subject of the predication* feature:

- (36) a. È venuto Gianni.
 is come Gianni
 Gianni has come
 [SubjP [EPPP \emptyset _{Loc} [AgrP è ...[VP arrivato Gianni
- b. Gianni a chiamato Piero.
 Gianni has called Piero
 [SubjP [EPPP Gianni_i [AgrP t_i a [VP t_i chiamato Piero

Adapted from Cardinaletti (2004, p. 153, examples 142 & 147)

Note that in Cardinaletti’s approach, auxiliaries raise from T° to AgrP in the Subject field, where they check tense features, but as she states, there are other ways in which the valuing of these features could take place (Cardinaletti, 2004, p. 152). Furthermore, the presence of *pro* in cases like (36.a) is ruled out given that *pro* would be able not only to check AgrP and EPP, but also SubjP: it would render the clause categorical since *pro* is only ever available with discourse-active referents. Interestingly, *pro*_{Loc} is not specified for Agreement features, and therefore, AgrP probes for a goal endowed with the relevant features in the

highest argument in ν P (if this stance is taken, there is no need for the auxiliary or the verb to raise to the Subject field), as in (36.b). Nevertheless, the overt locative used in existential clauses does carry valued agreement features, as it is shown and discussed in 4.3.2. This is shown to be important since it correlates with the ability of checking the *subject of the predicate* in Cardinaletti's terms.

One of the main drawbacks of Cardinaletti's (2004) dissection of the left periphery is that it may propose unnecessary projections. Instead, variation between the different elements that can satisfy SpecTP can be accounted for if the featural make-up of SpecTP is described to account for the different elements that can appear in this position. It could be hypothesised that T° 's featural make-up consists of a feature stack containing [case, person, number] features, as well as a [predication] feature and an Edge Feature that triggers the movement of a phrasal category to SpecTP to value all or a subset of these features, the EF being the only one that has to be absolutely checked by means attracting an XP, overt or covert to SpecTP. A stacking analysis of the featural make-up of SpecTP has the desirable descriptive consequence of reducing the amount of hypothesized projections, one of the problems of the Subject field hypothesis. For the purpose of this Chapter, her framework is useful to identify different properties associated to SpecTP: person, number, and predication. Note that when a clause receives categorial reading, it is assumed that it has bipartite Information Structure. Therefore, the content of SpecTP either connects the clause to the wider discourse (something is predicated of something that has already been introduced in it) or, when the *subject of the predication* feature is not valued, it receives athetic reading: it reports an event, decontextualised. In the case of left dislocated subjects, Rizzi proposes that they bind *pro* in SpecTP, which would check this feature.

A final piece of evidence for the existence of a position devoted to subjects within TP is the fact that subjects can occur below TP adverbs, as located in Cinque's adverbial hierarchy. This is illustrated in examples (37-38), where adverbs with identifiable positions within Cinque's hierarchy are used to locate subjects within the clause:

- (37) E lavorens En Guillem de Muntcada **féu** ses covinençes d'amor
 and then Sir Guillem de Muntcada did.3SG his pacts of;love
 ab Don Pero Ferrandes e ab la partida sua.
 with Sir Pero Ferrandes and with the party his

And then Sir Guillem de Muntcada did his pledge of love to Sir Pero Ferrandes and his party.

Fol. 9v, l 24

- (38) E ladonchs ell **respondria** a aqueles demandes
and then he reply.3SG.COND to those requests

And then he would answer to his requests.

Fol. 197r, l 28

In example (37), the nominal subject is located between the adverb *lavorens* ‘then’, a Future adverb located within TP in Cinque (1999) hierarchy. Similarly, in (38), the adverb *ladonchs* ‘then’, also a TP Future adverb, precedes the subject NP, which, in turn, precedes the verb, also located in TP. This indicates that the subject is placed within a projection in TP. The scarcity of TP adverbs that can be located within Cinque’s hierarchy makes it impossible, with data from the LFRJ database alone, to establish whether nominal and pronominal subjects occupy different positions within the left periphery depending on their category.

Subjects can also be moved to the left periphery for the purposes of focalisation or topicalisation. (39-41) show clear cases in which pronominal and nominal subjects have been topicalised: they are either followed by constituents that are located in a lower projection within the left periphery (Benincà, 2004, 2006), or they do not trigger proclisis (Batllori et al., 2005), contrasting with cases where the subject is focalised in which proclisis is borne.

- (39) E nostra mare, [sempre que nós **fom** nats],
and our mother as soon as we were.1PL born
envià·ns a Sancta Maria.
sent.3SG=us.CL to Saint Mary

And as soon as we were born, my mother sent us to Saint Mary’s church.

Fol. 3v, l 13

- (40) Aquest mal nós no **podem** adobar sinó per dues maneres:
this evil we not can.1PL mend.INF but for two manners

And this evil, we cannot mend but in two ways: (...)

(41) On yo us **prec (...)** que em retats
 hence I to;you.CL= beg.1SG that to;me.CL = return.2PL.SBJV

Balaguer (...)

Balaguer

Hence, I beg you that you return Balaguer to me.

In (39), the nominal subject is followed by a temporal adverbial clause. Temporal adverbial clauses are located within the Frame field (Haegeman, 2009, 2010). Therefore, in this case, the subject can only be a Hanging Topic, and not a CILDed topic, since it is undoubtedly located in the Frame Field, rather than the Topic field. In (40), the pronominal subject *nós*, ‘we’, follows a direct object that has been fronted to receive Focus Fronting (see Chapter 5, section 5.2.4 for a thorough analysis of this structure in OCat). Since we do not assume Topic projections to be recursive, in the lines of Benincà (2004) and Benincà & Poletto (2004), the only position in which the subject can appear is SpecTP. Finally, in (41) the subject pronoun *yo* ‘I’ is focalised. This is clear from the fact that the clitic pronoun *us* ‘to you’ is found proclitically.⁴⁵

Up to this point, we have reached two important conclusions: firstly, that OCat has a specialised subject position in SpecTP; secondly, that preverbal subjects need not occur only in SpecTP, they can also be moved to the Left Periphery. Note that all the preverbal subjects that have been examined up to this point contain old information. This conforms with the categorical/thetic reading divide described by Cardinaletti (2005), Gallego (2007), Petrova & Hinterhölzl (2009), by which preverbal subjects occur in categorical sentences, where something is being predicated of them.

This conclusion provides a further piece of evidence in favour of a non-V2 analysis of OCat. As stated in Chapter 3, one of the features that is often cited by the proponents of a V2 analysis for the Medieval Romance languages, is the lack of a specialised preverbal field for subjects. This is the case because, by virtue of the verb raising to the left periphery, any

⁴⁵ It is worth noting that for (41) *on* is taken to be an adverb, and not a relative pronoun, given that it has no clear antecedent in its context, and therefore, (41) is a main clause. See the entry of *on* in the *Diccionari Català – Valencià – Balear* for more information on the different meanings and functions of *on*.

constituent is susceptible to be moved to the specifier of the projection to which it has been moved, not only subjects. In contrast, in SVO languages the verb remains in TP, below the projection(s) that attracts elements that provide a connection between the clause and the discourse, thus having a preverbal field that is often targeted by subjects (amongst other elements, like spatio-temporal arguments).

In what follows, we explore the different postverbal positions that can be occupied by subjects in OCat. It is shown that OCat postverbal subjects can occupy three positions: they can either appear *in situ*, in *vP*, where they are base generated, in the low left periphery of the clause, or right dislocated. Once we identify the positions available for subjects, we proceed to describe which types of subjects are more prone to occur in each of them.

Postverbal subjects occurring *in situ* belong to three categories: (i) subjects of predicates crosslinguistically prone to present postverbal subjects, (ii) subjects of transitive verbs of clauses with thetic reading, and (iii) non-specific personal pronouns (personal pronouns with no readily recoverable reference). Table 4 shows the predicate nature of clauses with postverbal subjects in the LFRJ database:

Predicate type			Total
Postverbal nominal subjects – verb adjacent	Copula	9	87,2% (109/125)
	Verba dicendi	25	
	Existential	4	
	Passive	9	
	Reflexive	3	
	Unaccusative	59	
	Transitive	16	12,8% (16/125)
Total	125	100%	
Postverbal pronominal subjects – verb adjacent	Copula	3	73,3%
	Verba dicendi	21	
	Existential	4	
	Passive	-	
	Reflexive	2	
	Unaccusative	3	
	Transitive	12	26,7%
Total	45	100%	

Table 5 – Predicates in clauses with postverbal subjects in main clauses

Table 5 confirms that subject position is tightly connected with the nature of the predicate, as it was already established in Chapter 3, section 3.1.2.1. Nevertheless, there is a number of clauses with transitive predicates that present postverbal subjects. Transitive clauses with postverbal subjects need not have a preverbal constituent: out of the 16 clauses with a transitive predicate and postverbal nominal subjects, in 11 (69%), the verb occurs in clause initial position, 3 (19%) in verb second, and 2 (12%) in verb third clauses. Therefore, location of the subject a postverbal location cannot be linked to the presence of a constituent in the left periphery (see section 3.1.2.3 for an analysis of absolute verb initial clauses).

In cases in which the verb is preceded by other constituents, the interaction between the preverbal constituent and the postverbal subject is not obvious. This is shown in (42-43), examples with postverbal lexical subjects:

- (42) E aquela nuyt que abdós foren a Miravals, volch
 and that night that both were.3PL in Miravals wanted.3SG
 [Nostre Seyor] que nós **fóssem** engenrats.
 our Lord that we be.1PL.PRES.SBJV conceived

And that night in which both were at Miravals, Our Lord wanted for us to be conceived.

Fol. 3v, 18

- (43) E puyts, passat açò, **demanaren-nos** [nostres naturals].
 and then passed this asked.3PL=to;us.CL our relatives

And then, after this happened, our relatives requested us.

Fol. 5v, 12

In (42), the preverbal adjunct, containing exclusively old information, is located within the Frame field, as will be extensively shown in Chapter 5. Therefore, it cannot be the trigger of Germanic-style subject-verb inversion. Similarly, in (43), the adverb *puyts* ‘afterwards’, and the participial sentence *passat açò* ‘after this happened’ are also located in the Frame field, and unable to trigger subject verb inversion, and nonetheless, the subject is located postverbally. We propose that the reading of these postverbal subjects is secondary focus withinthetic or all-focus clauses. (42-43) are clauses in which the entire predicate is focalised (i.e. they have athetic reading), in contrast with clauses that only host information focus. Examples (42-43), the CP layer hosts various types of constituents, in the Frame field.

Absolute verb initial clauses have been associated with all-focus andthetic readings in the literature (see section 3.1.2 in this Chapter). This is the case of example (44), the opening line of *El Llibre dels Fets*:

- (44) **Retrau** mon seyor sent Jacme que fe sens obres morta és
 states my lord saint James that faith without deeds dead is

‘My lord Saint James states that faith without deeds dies’.

Fol. 1r, 11

In (44), the clause's subject occurs postverbally. This word order is expected if theticity is linked to the subject remaining *in situ* instead of raising to the *subject of the predication* projection (Cardinaletti, 2004) or checking an 'aboutness' feature (Rizzi, 2018). Instead, the clause has no informational divide and receives thetic reading. Conversely, when subjects raise to SpecTP, it is what remains in the *v*-VP complex that receives focal reading (narrow informational focus, see Chapter 5). In these cases, there can be sharp divide between new and old information, and the subject can be taken to act as *subject of the predication*.

Pronominal postverbal subjects are much less frequent than nominal ones, as shown in Table 2: nominal subjects make up 76% of postverbal subjects while pronominal subjects only account for 24% of cases. Most cases of postverbal pronominal subjects, 73.3%, are linked to predicates that are crosslinguistically prone to present postverbal subjects. The remaining 26.7% clauses contain transitive predicates. In XP V S_{Pron} clauses with transitive predicates, the verb is consistently in the second linear position within the clause. Preverbal constituents include focalised adverbs, focalised prepositional phrases (argumental and non-argumental) and focalised direct objects in 66.7% (8/12) cases (the relation between postverbal verb-adjacent subjects and focalised constituents is discussed at length in Chapter 5). There are four cases, 33.3% of XP V S_{Pron} with non-focalised preverbal constituents. Examples (45-48) contain two cases of XP V S_{Pron} with preverbal focalised constituents, and the two cases in which the constituent preceding the verb is a Scene Setter. The context of each example has been provided so that the informational value of each constituent may be available to the reader.

- (45) Context: E en tant levà·s en peus lo bisbe de Barçalona, per nom En Berenguer de Palou, e dix: (...). E sobre açò dis lo bisbe de Gerona:

And in the meanwhile, the bishop of Barcelona, named Sir Berenguer of Palou, stood on his feet and said: (...). And the bishop of Girona said the following about this:

‘A	Nostre Seyor	graesch	<u>jo</u>	la	bona	voluntat	que	Déus	
to	our	Lord	thank.1SG	I	the	good	will	that	God
ha	donada	a	vós	e	a	vostra	cort’.		
has	given	to	you	and	to	your	court		

I thank our Lord for giving good will to you and your court.

Fol. 31v, 17

- (46) Context: E nós parlam ab En Guillem de Muntcada e dixem-li que aquest pleyt nos volien fer, ço és assaber: que retriem lo castell en mà d'En Ramon Berenguer d'Àger, e ell que·l tendria en faeltat, e que aquel que goanyaria lo pleyt entre En Guerau e la comtessa que li retés él lo castel. E respòs En Guillem de Muntcada e dix:

And we spoke with Sir Guillem of Muntcada and told him about the dispute that they had with us, which is: that they would render us the castle by the hand of Sir Ramon Berenguer of Àger, and that he would render it in fealty, and that he who would win the dispute between Sir Guerau and the countess, that would render him the castle. And Sir Guillem of Muntcada answered and said:

‘AÇÒ no **tench** yo com cosa que vós dejats fer.’
 this not have.1SG I as thing that you should.2PL do.INF

I do not hold this as something you should do.

Fol. 25v, l 3

- (47) Context: E tots vengren al dia de la cort, levat Don Fferrando e·l comte Don Sanxo, car havien esperança que cascú fos rey.

And they all came the day of the assembly, except of Sir Ferrando and the Count Sir Sanxo, since they had the hope that either of them could be king.

E aquí, **juraren-nos** tots que·ns gardarien nostre
 and here swore.3PL=to;us.CL all that=to;us.CL guard.3PL.COND our
 cors e nostres membres e nostra terra (...).
 body and our limbs and our land

And here, all of them swore to us that they would guard for us our body, our limbs and our land (...).

Fol. 6r, l 2

- (48) Context: E féu fer ·xii· candeles, totes de ·i· pes e d'una granea, e féu-les encendre totes ensemps, e a cada una mès sengles noms dels apòstols, e promès a Nostre Seyor que aquela que pus duraria, que aquel nom auríem nós. E durà més la de Sent Jacme be ·iii· dits de través que les altres.

And she had 12 candles made, all of the same weight and size, and she had all of them lit at the same time, and each of them was labelled with the name of an apostle,

and promised to Our Lord that that candle which would last the most, that that name we would bear. And that of Saint James last a good three fingers more than the others.

E	per	açò	e	per	la	gràcia	de	Déu	havem	<u>nós</u>
and	for	this	and	for	the	grace	of	God	have.1PL	we
nom	En	Jacme.								
name	Sir	Jaume								

And because of this and by the grace of God, we are called Sir Jaume.

Fol. 3v, l 26

Examples (45-48) contain a focalised constituent in the left periphery. In (45), the focalised constituent is the argumental prepositional phrase *a Nostre Seyor* ‘to Our Lord’. In (46), the focalised constituent is the direct object *açò*, ‘this’. In both cases, the focalised constituent contains old information (*Nostre Seyor* is an active referent in the reported conversation from which (45) is extracted, and *açò* makes reference to what has been said immediately before in the reported conversation). (47-48) present Scene Setters in the left periphery. In (47), the adverb *aquí* ‘here’, is clearly not focalised since the clitic pronoun that accompanies the verb is enclitic (if it were focalised, proclisis would be expected, as discussed in Chapter 3, section 3.1.5). In (48), the coordinated prepositional phrases *per açò e per la gràcia de Déu* ‘because of this and by the grace of God’, both active referents in the preceding discourse, act as Scene Setters of the clause. Note that in (45), (47) and (48), the three examples where a non-direct object (including indirect objects and other prepositional complements) is fronted, the subject is followed by the direct object of the clause, remaining *in situ*. In (48), the direct object is not heavy. Therefore, it could have been amenable to be fronted to the left periphery and comply with the V2 requirement, were it to be active in OCat. Instead, SpecFocP is left empty, while a Scene Setter occurs in the Frame field.

A further consideration about postverbal verb-adjacent subjects needs to be made regarding their position in relation to periphrastic verbal forms. Both postverbal pronominal and nominal subjects occur between the auxiliary and the lexical verb, be it *haver* ‘to have’, *ésser* ‘to be’, or the modals *poder* ‘to be able to’ and *voler* ‘to want’, as shown in (49-52). Subjects are not the only elements that can appear between the auxiliary and the lexical verb: direct objects and adverbs can also appear in this position, as in some NSLs like Modern

Greek. This contrasts with MCat, where subjects cannot appear between any auxiliary (be it modal or tense) and the lexical verb.

- (49) E aytal semblança **pot** hom fer de vós.
 and such likeness can.3SG one do.INF of you

And one can find such likeness in you.

Fol. 31r, 1 26

- (50) E quant fo açò fet, **fou** la mar abonança.
 and when was.3SG this done was.3SG the sea appeased

And when this was done, the sea had become calm.

Fol. 33v, 1 2

- (51) a. Hom **pot** veure tal semblança en vós.
 one can.3SG see.INF such likeness in you

One can see such likeness in you.

- b. ***Pot** hom veure tal semblança en vós.
 can.3SG one see.INF such likeness in you

One can see such likeness in you.

MCat

- (52) a. I la mar **fou** abonança.
 and the sea was.3SG appeased

And the sea calmed down.

- b. *I **fou** la mar abonança.
 and was.3SG the sea appeased

And the sea calmed down.

MCat

(49-50) show how both pronominal and nominal subjects can appear between auxiliaries and lexical verbs in OCat. In (49), the subject appears between the modal verb *poder* ‘to be able to’ and the lexical verb that follows, with the subject *hom* ‘one’ sandwiched between the auxiliary and the non-finite verb, but there DOs, certain adverbs (*bé* ‘well’, *ja* ‘already’ and *més* ‘more’) and prepositional phrases can also be found between the auxiliary

and the non-finite form. There are no cases of lexical subjects occurring between a tense auxiliary and a non-finite form. In (50), a nominal subject, *la mar* ‘the sea’, appears between the auxiliary of a passive construction and the lexical verb, a past participle that agrees with the subject and that is located in ν -VP. Passive voice clauses like (50) are the only cases of AUX S_{LEX} V found in the LFRJ database, even though they could also be analysed as copular clauses with predicative adjectival participles acting as subject arguments, rendering (50) a case of V S XP_{SubjectComplement} rather than passive clauses. The fact that passive clauses are a mechanism by which the aboutness of the clause shifts from object to subject (Rizzi, 2018) favours this hypothesis: if one of the motivations for passivisation is to shift aboutness to the subject, it would not be expected for the subject of the passive voice to occur in a postverbal position. No constituents other than subject pronouns are found between passive auxiliaries and past participles, and none of these constructions are available in M_{Cat}, as shown in (51) and (52).

According to Poletto (2014) and Wolfe (2015), the presence of subjects between an auxiliary and a non-finite form could constitute evidence for a V-to-C analysis, since complements of auxiliary verbs signal the edge of the ν -VP complex.⁴⁶ However, following Belletti (2004), it is also possible that these subjects, instead of occurring in SpecTP, given their position to the left of the infinitival complement, might simply be fronted to the low left periphery in ν P, where informational structure-related projections parallel to those present in the left periphery would be available. This possibility is backed up by the fact that all subjects (and other elements) found between finite modal verbs are mostly elements that tend to receive focal readings. In the case of pronouns, the two that are found in AUX_{MODAL} S_{PRON} V_{INF} configurations are *hom* ‘one’ and *negú* ‘no one’, both indefinite pronouns. The make-up of the O_{Cat} low left periphery and the other elements that can be found in AUX_{MODAL} XP V_{INF} configurations are briefly considered in Chapter 5, section 5.6. (51) and (52) show how it is impossible for M_{Cat} subjects to appear in such position. This might be due to the fact that M_{Cat} low left periphery has undergone a drastic simplification through the centuries, ceasing to have a Focus position available. Alternatively, it could also be explained by relating it to the grammaticalisation of modal periphrases into monoclausal structures, by which the non-finite form raises with the finite form to TP, preventing any element from appearing between

⁴⁶ Poletto (2014, p. 2) states that the V2 systems should yield AUX S V sequences in tense periphrases, without discussing whether similar sequences in passive clauses or clauses with modal auxiliaries.

the auxiliary and the verb. Nevertheless, the availability of examples like (53) and (55) in MCat favours the former hypothesis:

- (53) Ja ho pots ben dir!
indeed it.CL= can.2SG well say.INF

It goes without saying!

- (54) Pots recitar bé el poema.
can.2SG recite.INF well the poem

You recite the poem well.

- (55) L'església de Sant Pau ha estat ben restaurada.
the;church of Saint Pau has been well restored

The church of Saint Pau has been well restored.

- (56) L'església de Sant Pau ha estat restaurada bé.
the;church of Saint Paul has been restored well

The church of Saint Pau has been restored well.

MCat

Example (53) displays one of the few MCat cases in which an element can occur between a finite auxiliary and a non-finite verb.⁴⁷ This possibility has been retained for the low adverb (Cinque, 1999) *bé* 'well'. When moved to the low left periphery, it receives emphatic reading. In (54), in contrast, *bé* occurs in its base generated position, after the infinitive, and does not receive emphatic reading. In the case of passive clauses, the presence of adverbs between the finite verb and the non-finite is also allowed in MCat, as shown in (54). The difference between (54) and (55) is that in (54), the adverb has been focalised to the low left periphery of vP (Belletti, 2001), and therefore, it receives an emphatic reading, whereas in (56), the reading is neutral.

⁴⁷ *Ben*, the allomorph of *bé* 'well' when it precedes adjectives, adverbs and verbal forms, can also be found between tense auxiliaries and their associated non-finite forms in MCat. This configuration is not instantiated in the LFRJ database.

- i. Ara m' has ben cardat.
now me.CL= have.2SG well screwed up.PPT

You got me there.

MCat

So far, it has been established that postverbal subjects may occur in their base generated position or they can be focalised in the low left periphery. Postverbal subjects can occupy a further position within the clause: they can be right dislocated, when they have been previously introduced in the discourse. Topical elements dislocated to the right periphery strictly require the presence of a coreferent clitic pronoun in the core of the clause (Villalba, 2009).⁴⁸ However, OCat did not have subject clitic pronouns, making the identification of such structures trickier. By being right dislocated, subjects escape the possibility of being focalised. The information structure value of right dislocated elements are explored in Chapter 5. Example (57) and (58) contain two cases of right dislocated subjects.

(57) E, feyt aquest matrimoni, **entram** en Aragó e
 and done this marriage entered.1PL in Aragon and
 en Cathalunya, nós e nostra muyler la reyna.
 in Catalonia we and our wife the queen

And, once this marriage was done, the queen and I went to Aragon and Catalonia.

Fol. 9v, l 12

(58) E ab ·i· trebeyl que feya ab uns moços,
 and with 1 game that did.3SG with some boys
donaren-li d'una teula en lo cap ·i· d'aquels qui
 gave.3PL=him.CL of;one tile in the head 1 of;those who
jugaven a aquel joch contra él.
 played.3PL at that game against him

And when he was playing with some boys, they hit him with a tile on the head, one of those that were playing at that game against him.

Fol. 9r, l 14

⁴⁸ There are no identifiable cases of right dislocation of DOs or prepositional phrases in the LFRJ database. Therefore, we use a MCat example to illustrate this phenomenon. As it can be appreciated, the DO *la poma* ‘the apple’ is right dislocated and coreferent with a preverbal object clitic that triggers agreement of the past participle.

i. Me li' **he** menjada, [la poma].i.
 myself.REFL.1SG.CL it.CL have.1SG eaten the apple
 I have eaten it, the apple.

In (57), the right dislocation is both orthographically and syntactically marked: it is separated from the core of the clause by a comma and it occurs after an argumental prepositional phrase, base-generated and located within *v*-VP and marking its boundary. In (58), the subject also follows an argumental prepositional phrase. In this case, it is ambiguous whether the subject is right dislocated or it is an afterthought, since it does not agree with the main clause's verb: *donaren* 'they gave' is in the 3rd person plural, while the subject is a 3rd person singular *i d'aquels qui (...)* 'one of those who (...)'. The lack of agreement could be explained if the subject is added as an afterthought, rather than moved from its base generated position to the right periphery (see Villalba [2000, p. 133] for considerations on right dislocation and afterthoughts in MCat).

Up to this point, seven different positions have been identified for subjects in OCat. Preverbally, they can appear in four positions: in SpecTP, in the Focus field, in the Topic field, and in the Frame field as HTs. Postverbally, we have identified three positions in which they can appear: *in situ*, in their base generated projection, in left periphery of *v*P, and right dislocated.

4.3.1 The Old Catalan preverbal field: specialised for hosting subjects?

As mentioned above, OCat is an NSL with V-to-T movement. The basic word order of Romance languages that present these two features is assumed to be SVO (Sheehan [2010], Forcadell [2013], *pace* Alexiadou & Anagnostopoulou [1998], Vallduví [1999]). In what follows, we endeavour to identify whether the OCat preverbal field is specialised for hosting subjects, and to describe the distribution of different subject types.

Table 5 shows the different word order patterns found in V2 and V3< clauses with overt subjects:

	Non Pro-Drop		Total
	Structure	Occurrences	
Preverbal Subjects	S _{LEX} V	78	72.1% (259/359)
	S _{PRO} V	100	
	S _{LEX} V O	28	
	S _{LEX} V XP O	4	
	S _{PRO} V O	45	
	S _{PRO} V XP O	4	
Postverbal subjects with verb – subject adjacency	ADV V S _{LEX}	36	23.2% (83/359)
	ADV V S _{PRON}	8	
	DO V S _{LEX}	2	
	DO V S _{PRON}	10	
	PP V S _{LEX}	16	
	PP V S _{PRON}	7	
	SC V S _{LEX}	4	
Non-verb adjacent postverbal subjects	DO V PP S _{LEX}	1	4.7% (17/359)
	PP V PP S _{LEX}	1	
	PP V PP S _{PRO}	1	
	ADV V PP S _{LEX}	8	
	ADV V PP S _{PRON}	1	
	ADV V DO S _{LEX}	2	
	PP V DO S _{LEX}	1	
	ADV V SC S _{LEX}	3	
	SC V SC SN	1	
Total		359 (100%)	

Table 5 – Preverbal constituents in non-prodrop main clauses (V2, V3<) in the LFRJ database

In Table 5, we can see the nature of preverbal constituents in the LFRJ database. Noun phrases have been classified according to their syntactic function, rather than their phrasal type, so that the distinction between subjects and direct objects is readily available to the reader. As we can see, 72.1% of V2 clauses present their subject as the clause initial element. 23.2% of them have another type of constituent in clause initial position, and could be candidates for Germanic inversion, since the subject appears postverbally, adjacent to the

verb. Finally, 4.7% of clauses display a non-subject constituent in clause initial position, and their overt subject does not appear adjacent to the verb in a postverbal position.

The frequency of non-subject constituents in preverbal position is high enough to be considered a cue for the acquisition of V2 orders, as established by several authors. Lightfoot (1999, p. 153), examining German and Dutch data, proposes that the frequency of XP V S clauses within V2 clauses should be of 30% of Primary Linguistic Data (PLD); Yang (2000), examining Dutch data, suggests that it should be of 23%; and Westergaard (2009), examining Norwegian data, proposes that it should be of 13.6%. Nevertheless, a close analysis of each of the instances in which a non-subject constituent occurs pre-verbally, lowers the percentage of instances that could represent a cue for the acquisition of V2, especially in the case of most adverbial and prepositional phrases, which act as frame setters, and are base-generated in the higher projections of the left periphery (see Chapter 5 for more on the syntax and distribution of these elements in the OCat left periphery). There are a few cases of preverbal adverbs and prepositional phrases whose location in the left periphery is motivated by other factors, mainly focalisation and clitic left dislocation (CLLD). Focalised constituents are moved to a position in the left periphery to receive an emphatic reading, while the CLLD constituents are base-generated in it (Cinque, 1990, p. 60; Sheehan, 2005, p. 3; Gallego, 2007). Finally, a small number of clauses present clause initial objects and subject complements, which are clearly focal. Table 6 shows the position of adverb phrases, prepositional phrases, and direct objects within the left periphery, as described by Poletto & Benincà (2004) for the Old Romance languages:

	Frame Setters	CILDed	Topicalised	Focalised
Adverbs	75% (42/56)	-	-	25% (14/56)
Prepositional Phrases	73% (19/26)	4% (1/26)	-	23% (6/26)
Direct Objects	-	-	-	100% (13/13)
Subject Complements	-	-	40% (2/5)	60% (3/5)
Total (out of XP V S)	61% (61/100)	1% (1/100)	2% (2/100)	36% (36/100)

Table 6 – Position of preverbal constituents in non-prodrop sentences with postverbal adjacent subjects

Examples (59-61) illustrate the various types of constituents other than nominal expressions that can appear immediately preverbally, with different informational readings:

Adverbs

- (59) E puys **levà·s** lo probost de Terragona.
and then stood.3SG=himself.CL the chief of Tarragona
And then the chief of Tarragona got to his feet.

Fol. 31v, l 16

- (60) E ja **vench** la ora del vespre.
and already came.3SG the time of;the evening
And it was already evening.

Fol. 33r, l 9

- (61) E Don Exemen Corneyl **era** ja de dies (...)
and Sir Exemen Corneyl was.3SG already of days
And Sir Exemen Corneyl was already old (...).

(62) is an example of a prepositional phrase with Scene Setter meaning, located within the frame field. In (63), however, the experiencer PP, acting as an adjunct, is CILDED to the Topic field, located between the Frame and Focus fields. This is clear since there is a resumptive clitic pronoun coreferent with the PP in the core of the clause. Finally, in (64), the recipient PP, an argument of the predicate *agrair* ‘to thank’, is fronted to the Focus field to receive Focus Fronting (once more, we refer the reader to Chapter 5 for an extensive discussion on this construction). Note that the information contained in the PP is old: *nostre Seyor* ‘our Lord’, is an active referent throughout the text, and especially within the context of (64), since the sentence is uttered by the Bishop of Girona, on the one hand, and since it has just been mentioned that the conquest of Mallorca (the subject of the conversation), is God’s will. Therefore, it cannot receive information focus reading. The last type of preverbal constituent that are found are Direct Objects (DOs):

Direct Objects

(65) AQUESTA PARAULA **volch** Nostre Seyor **complir** en los
 this word wants Our Lord accomplish.INF in the
 nostres feytz.
 our deeds

Our Lord wanted to accomplish this statement in our deeds.

Fol. 1r, 12

(66) ‘AÇÒ no **tench** yo com cosa que vós dejats fer.’
 this not have.1SG I as thing that you should.2PL do.INF

I do not hold this as something you should do.

Fol. 25v, 13

(67) ‘Och’ **dixem** nós.
 yes said.1PL we

We said ‘yes’.

Fol. 25 v, 126

In (65), the direct object *aquesta paraula* ‘this statement’, appears preverbally. It is one of the 13 objects that occur preverbally in clauses where the subject is postverbal and adjacent to the verb. They constitute a 3.2% out of the total XP V S instances in the LFRJ

database. In example (66), the fronted DO is a demonstrative pronoun, *açò* ‘this’. In (67), the fronted DO contains a demonstrative adjective. Focus Fronting refers to a construction that involves verb movement to the left periphery and subject-verb inversion and constitutes one of the few contexts in which the verb moved to the left periphery in OCat. It is explored at great length in Chapter 5.

Example (67) constitutes an especial case, as *Och* ‘yes’ is a reported speech direct object. Therefore, its preverbal location obeys rules linked to the predicate type, *verba dicendi*, as it has been shown in section 3.1.3.2, rather than to sentential information structure.

Examples (65-67) provide crucial information about the nature of the OCat prefield in main clauses. On the one hand, most constituents (62%, 62/100 of XP V S) are located within the frame field, and base-generated in their position. Only argumental prepositional phrases and noun phrases are topicalised (2%, 2/100 of XP V S respectively). Finally, there is a small number of constituents containing old information that are fronted to the Focus field to receive an emphatic reading, including Subject Complements, DOs containing old information, argumental and adjunct prepositional phrases and adverbs generated in the lower layers of the clause (36%, 36/100 of the total of XP V S clauses). They only represent 3.6% out of the total of main clauses analysed in the LFRJ database, and 8.1% of V2 clauses from the database. Therefore, they can hardly constitute a cue for the acquisition of a V2 grammar, if the percentages of frequency of XP V S orders within V2 clauses provided by Lightfoot (1999), Yang (2000) and Westergaard (2009) are on the right track.

The fact that XP V S orders do not constitute cue for the acquisition of a V2 grammar, together with the fact that S V (XP) orders represent 72.1% of V2 clauses, lead us to establish that the preverbal field, in OCat main clauses, together with the structural arguments presented in Section 4.2, was specialised for hosting subjects, as Sitaridou (2011, 2012, 2019) points out for OSp.

Another piece of evidence in this direction can be found in the type of verbal predicate found in XP V S clauses:

Predicate type		Number of occurrences	Total
Predicates crosslinguistically prone to have postverbal subjects	Copula	10	77.1% (64/83)
	Verba dicendi	8	
	Existential	4	
	Passive	4	
	Reflexive	1	
	Unaccusative	37	
	Transitive	19	
	Total	83	100% (83/83)

Table 7 – Predicate types in XP V S V2 clauses

As shown in Table 7, 77.1% of predicates appearing in XP V S clauses are crosslinguistically prone to present postverbal subjects (see Corr [2016] for a null-locative analysis of XP V S in Ibero-Romance), as extensively explored in Chapter 3, section 3.1.2. It is worth mentioning that the remainder 22.9% are transitive verbs that present focalised direct objects, focalised arguments or focalised adverbs, as well as Scene Setters, in preverbal position. In these cases, the XP V S order can be linked to the informational weight of the subject (Belletti, 2004; section 4.2 in this Chapter for MCat) or to Focus Fronting, a construction explored in depth in Chapter 5, section 5.2.

In this section, quantitative data suggesting that the OCat preverbal field was specialised for hosting subjects has been presented. Three pieces of evidence have been provided: (i) the high rate of S V (XP) orders, (ii) the lack of a clear cue for the acquisition of Germanic verb-subject inversion, and (iii) the fact that most V S orders can be accounted for by means of their predicate type. In what follows, we consider the distribution of subjects within OCat main clauses, cases of XP V S that have not been explained here and showing that the distribution of subjects in OCat conforms with that of SVO languages.

4.3.2 Locatives and the information structure value SpecTP

In this section, we show that the satisfaction of EPP (understood as a requirement on the specifier of a projection to be filled by an element, Svenonius [2001, p. 16]) in the subject field in OCat could be carried out by elements other than nominal and pronominal constituents. Pinto (1998), Tortora (2001), Cardinaletti (2004), Borer (2009), Sitaridou (2010), Sifaki (2003), Sifaki & Tsoulas (2016), Sheehan (2010, 2015) among others, propose that locative elements are able to satisfy EPP. As discussed in 4.2.1, different types of constituents can occupy different positions within the subject field depending on their semantic nature, locatives filling the specifier of the highest projection of the field, namely EPPP in Cardinaletti (2004) description of the subject field. The semantic differences and featural make-up of potential satisfiers of EPP within the subject field determine which position within the subject field they target, how they interact with agreement, and whether they occur preverbally.

In the preceding section, we have also seen that clauses with athetic reading, where the whole verbal predicate receives focal reading, present postverbal subjects. These facts suggest that subject position is intimately linked not only to the informational status of the subject, but also to the informational reading of the clause as a whole, strengthening the hypothesis that SpecTP, is connected to topical or categorical readings, while postverbal subject positions are linked to presentational focus or thetic readings.

In what follows, we concentrate in OCat existential clauses, which provide valuable information to widen our understanding of OCat subject positions and confirm those hypotheses that connect the satisfaction of EPP with locatives (*pace* Bentley & Ciconte, 2015) and discourse grounding.

Bentley & Ciconte (2015, p. 1), building on McNally (2011, p. 1830), define existential constructions as ‘constructions with non-canonical morphosyntax which express a proposition about the existence or presence of someone in a context’. OCat had two existential predicates that occur in constructions that correspond to this definition: *ésser* ‘to be’ and *haver* ‘to have’. The elements occurring in existential constructions include a locative expression, the verb, a pivot, and optionally, a coda. The locative expression may be overt (be it an adverb, prepositional phrase, a referential clitic pronoun, or a non-referential proform (Bentley 2015, p. 1). The pivot is the element of which the existence in a certain context is being asserted. Finally, the coda can be optionally present. Codas can belong to different

categories and have different functions within the existential clause. (68) and (69) show the canonical structure of existential clauses with both predicates (the order of constituents may change in relation to information structure).

(68) [locative_{constituent/proform}] [to have_{3SG}] [pivot_{indefinite}] [coda_{optional}]

(69) [locative_{constituent/proform}] [to be_{Agr i}] [pivot_{definite i}] [coda_{optional}]

There are two noteworthy differences between (68) and (69) in the Romance languages that present both options: the fact that *haver* tends to take indefinite pivots, while *ésser* tends to take definite ones. Furthermore, *ésser* agrees in person and number with its pivot, while *haver* invariably appears in the 3rd person singular. We first explore existential sentences with *haver*, considering the different locatives that can appear in them, the lack of agreement between the verb and the pivot, the relation between postverbal subjects and indefiniteness, and the restrictions for certain types of subjects to appear in *haver* existential clauses. Secondly, we consider existential clauses with *ésser*.

Haver existential clauses invariably present a locative expression preceding the verb. It can either be an adverb, as in (70), a prepositional phrase, as in (71), a deictic location pronoun, as in (72), or the clitic pronoun *hi* (from Latin *IBI/HIC*), as in (73):⁴⁹

(70) Adverb:

<u>Aquí</u>	no	ha	altre	acort.
here	not	had.3SG	other	agreement

Here there is no agreement.

Fol. 64v, l 14

(71) Prepositional phrase:

E	[<u>en</u>	<u>aque</u>	<u>yla</u>]	ha	rey.
and	in	that	island	had.3SG	king

And in that island, there is a king.

Fol. 27v, l 13

⁴⁹ Three cases of existential clauses with *haver* where there is no overt locative have been found in the LFRJ database. These three cases correspond to answers to negative implicatures with the positive polarity adverb *sí*. The distribution of *sí* and its interaction with clitics is extensively discussed in Chapter 3, section 3.1.6.

(72) Relative adverb:

E nós, qui conexíem que gran mal era de la
and we who knew.1PL that great damage was of the
vila, on tan gran gent **havia** (...)
village where such great people had.3SG

And we, who knew that the village, where there was a great multitude, was
undergoing great difficulties (...)

Fol. 132, l 22

(73) Clitic pronoun *hi*:

E **havia** -hi ·cl· almogàvers.
and had.3SG=there.CL 150 mercenaries

And there were 150 infantry soldiers.

Fol. 112r, l 27

Adverbs, prepositional phrases and adverbial relative pronouns can be anaphoric: they can refer back to a location that has already been introduced in the discourse. The adverbial clitic *hi* can either make reference to a location present in the same clause, which we refer to as specific reference *hi*, or it can make reference to a location overtly mentioned or implied in the discourse but not overtly present in the near vicinity of the clause, which we refer to as unspecific reference: it anchors the event to the same spacio-temporal frame of the discourse (Alexiadou & Carvalho 2017, p. 58), and according to Bentley & Ciconte (2015, p. 89), its deictic interpretation refers to the speaker's proximal physical space. In the case of our text, it refers to the proximal space of the intra-homodiegetic narrator.⁵⁰ Both options are illustrated in (74) and (75).

(74) Specific referential *hi*

E car és [en terra de montanya]_i e **ha-y**_i
and since is in land of mountain and has=there.CL

⁵⁰ Especially in narrative texts, as well as in reported speech, the nonspecific reference *hi* need not refer back to the proximal physical space of the speaker, as shown in (i). This suggests that *hi* might be coreferential with a *binding* discourse operator located higher within the left periphery, rather than with a participant of the communicative exchange, as suggested by Alexiadou & Carvalho (2017, p. 60).

tan gran compaya (...)
such great company

And since it is a mountain land and there was such a big company...

Fol. 146v, l 14

(75) Unspecific referential *hi*

Context: E puyas enviam Bertran d'Aunés que pujàs al pug Escardeyno ab ·iii·
cavallers e ab ·iiii· escuders; e él pujà-hi e dix-nos que no y poriem parar tendes (...).

*And then, we sent Bertran d'Aunés so that he would climb the Escardenyo hill with 3
knights and four squires, and he climbed there and told us that we could not set
camp there.*

E nós pujam en altre pug que y **ha** (...)
and we climbed.1PL in other hill that there.CL has

And we climbed another hill that there is.

Fol. 131r, l 24

In (75), *hi* is coreferent with the prepositional phrase *en terra de montanya* 'in mountainous land'. In contrast, in (76), *y* is not coreferent with any of the overt referents introduced in the immediately preceding discourse or within the same clause. Instead, it refers to the location in which the action narrated in the discourse is taking place. *Y* provides a spatio-temporal frame in which the *altre pug* 'the other hill' exists.

Up to this point, we have described the type of locatives that can occur in OCat existential clauses with the predicate *haver*. It has been established that two main types of locative elements occur with *haver*: elements with semantic content, which include adverbs, prepositional phrases, and pronouns coreferent with overt locative expressions (relative or clitic), and *hi* with non-specific referential properties. Furthermore, all clauses have in common that the verb occurs in the 3rd person singular. The lack of agreement between the pivot and the verb in existential clauses with the verb *to have* is attested throughout the Romance languages, especially those in which Latin HABERE is not only used as an existential

predicate, but also as a possession one (Bentley & Ciconte, 2015, Ch. 1), as it was the case in OCat.⁵¹

The lack of agreement between the verb and the pivot was apparent in example (73), where the plural pivot ‘*cl·almogàvers*’, which would agree with another verb in the 3rd person plural, occurs with the verb in the 3rd person singular). We follow Rigau (1993, 1997), Pinto (1997), Tortora (2001), Cardinaletti (2004), Sheehan (2006), Sifaki (2003), Espinal & McNally (2010) in attributing the lack of agreement to the presence of the locative. Different analyses have been proposed as to why that is the case. Rigau (1997) analyses MCat *haver-hi* as the spelt-out result of the incorporation of an abstract preposition of central coincidence into the verb *to be*, where *hi* is the overt realisation of the head of a prepositional phrase, and blocks agreement between the verb and the pivot. On the other hand, Espinal & McNally (2010, p. 123) suggest that in MCat, *hi* checks an unspecified [*uLoc*] feature of verbs that are specified with it. Although the situational argument realised by *hi* does not correspond to a subject, it raises to SpecTP as it is the only candidate apt to fill the position (Espinal & McNally 2010, p. 125), and it is responsible for the verb occurring in the 3rd person singular. Tortora (2001), for MIIt, Sheehan (2005) for Italo and Ibero-Romance languages, and Borer (2010, p. 320) for MCat and Hebrew suggest that the presence of a locative element in SpecTP is responsible for VS orders, not only with existential predicates, but also with unaccusative ones (see Chapter 3, section 3.1.2 for more on unaccusative predicates and postverbal subjects).

A consequence of the presence of *hi* in the SpecTP and the lack of agreement on the finite verb is the inability for *haver-hi* predicates to assign nominative case (Rigau, 1993). Instead, Rigau (1993, 1997) and Espinal & McNally (2010) argue that the preposition of central coincidence incorporated into the verb assigns an oblique case to the pivot. Examples (76-81) illustrate the facts that have just been described for MCat, mainly: (i) the compulsory

⁵¹ In MCat, *haver* has lost its lexical use and has grammaticalised into (i) a tense auxiliary, and (ii) one of the strategies to create existential clauses when occurring in the 3rd person singular and with the locative clitic *hi* (*haver-hi*). However, some lexical usages remain in fixed expressions such as (i), always with the meaning *to get/to obtain*:

- (i) No ho ha pogut **haver.**
 Not it.CL= has could have.INF
 He hasn't been able to get it.

occurrence of *hi* in *haver-hi* existential clauses, (ii) the lack of agreement between the verb and the pivot, and (iii) the fact that the pivot is not assigned nominative case, but oblique.

(76) [A casa]_i hi ha pomes.
 at house there.CL has apples

There are apples at home.

(77) Hi ha pomes.
 there.CL has apples

There are apples.

(78) ** A casa ha pomes.
 at house has apples

There are apples at home.

(79) ** A casa hi ha ell.
 at house there.CL has he

He is at home.

(80) a. A casa hi ha pomes.
 at home there.CL has apples

At home there are apples.

b. N_i' hi ha.
 of;it.CL= there.CL has

There are some (of apples).

(81) a. Hi ha un cavall.
 there.CL has a horse

There is a horse.

b. N' hi ha un.
 of;it.CL= there.CL has one

There is one (of it).

c. ??Hi ha el cavall.
 there.CL has the horse

There is the horse.

d. * L' hi ha.
 it.CL.ACC.SG= there.CL has

There it is.

MCat

In (76), we have an existential clause in which apart from the clitic pronoun *hi* required by the predicate *haver-hi*, there is a prepositional phrase that is coreferential with it. It contrasts with (77), where *hi* occurs without a coreferential locative expression, yielding a completely grammatical sentence. In (78), we have the proof that *hi* is necessary for existential *haver-hi*, and that the presence of a locative expression alone does not suffice to render the clause grammatical. In (79), we see how a nominative personal pronoun, *ell* ‘he’, is banned from occurring as a pivot, while in (80-81) we see that an indefinite pivot can be pronominalized by the partitive clitic *en* ‘of it’, but that definite pivots, as in (84.c-d), cannot be pronominalized by accusative clitic pronouns.⁵²

OCat differs from MCat in that the presence of *hi* is not compulsory (as we saw in examples (76-78), where locative elements other than *hi* occur with existential *haver*). However, it patterns with it in not allowing nominative pronouns in pivot position (there are no cases attested), and in not allowing the co-occurrence of accusative clitic pronouns and *hi* (there are no instances of this combination in the LFRJ database), but in allowing the co-occurrence of *hi* with the partitive clitic *en*, as shown in (82) where *en* is coreferent with the CILDed pronoun *molts* ‘many’, and occurs cliticised to the verb together with *hi*, in preverbal position.

(82) *quar molts_i ni'í havia a qui no plaÿa que*
 since many of;them.CL=there.CL has to whom not please.3SG that
València fos presa
 Valencia was.3SG taken

(...) since there were many who were not pleased by Valencia being taken.

Fol. 116v, 120

⁵² As pointed out by Rigau (1997), this might be due to the fact that *hi* and *el/la* have the same person specification: 3rd person singular.

The fact that pivots can be substituted by the partitive clitic *en* but that there are no clear cases of nominative pivots in *El Llibre dels Fets* in existential clauses with the verb *haver* suggests that this predicate is unable to assign nominative case. This is intimately connected that *haver* existential clauses invariably show 3rd person singular morphology. We follow Rigau (1997, p. 104) in attributing the lack of agreement between the pivot and the verb to the presence of the clitic locative. According to Rigau (1997), *hi*, the subject of an abstract preposition of central coincidence, is responsible for the lack of agreement between the pivot and the verb. Taking Cardinaletti (2004) map of the subject field, *hi* raises to SpecEPPP, after raising to AgrP, where its failure to check [*uperson*] and [*unumber*] results in the verb taking the unmarked 3rd person singular (number and person heads being found within AgrP, in the subject field [Cardinaletti, 2004]). Rigau notes that the assignment of nominative case is linked to the presence of a valued person feature. Like *pro*_{Loc} found with unaccusative predicates, *hi* ‘is always deictic: it can either refer back to an overtly expressed location in the preceding discourse or make reference to the speaker’s proximal physical space. In terms of Talmy (1972), *hi* is the ground in which the pivot, the figure, is located. Therefore, *hi* conveys old information against which the pivot is contrasted. With this in mind, it could be argued that *hi* not only checks agreement and EPP in the Subject field, but that it also check the *subject of the predicate* feature with a [+ground] feature, instead of a [+figure] feature. This could be the key difference betweenthetic and categorical readings: in clauses receivingthetic reading, something is predicated of the ground, while in clauses receiving categorical readings, a [+figure] subject occurs in SubjP, of which something is predicated. This hypothesis is also proposed for MSp VSX and clauses with copula verbs by Leonetti (2012, 2015).

OCat had another strategy to build presentational clauses: the verb *ésser* ‘to be’ with a locative element, which, like in *haver* clauses, need not have an explicit referent in the discourse. *Ésser* and *haver* presentational clauses display structural differences, parallel to those of languages that use these two predicates to build presentational clauses (Bentley & Ciconte, 2015): Firstly, in *haver* presentational clauses there is no agreement between the pivot and the verb, while *ésser* presentational clauses do show agreement between the pivot and the verb. Secondly, *haver* presentational clauses systematically present indefinite pivots, which are often referents newly introduced in the discourse, while *ésser* presentational clauses invariably have pivots which have already been introduced in the discourse. The distribution of pivots seems to be sensitive to their finiteness: in *haver* presentational clauses, indefinite

pivots tend to appear postverbally, while in presentational clauses with *ésser*, definite pivots occur preverbally. This distribution suggests that in OCat, there was an active definiteness effect, as described by Diesing (1992) for Yiddish, Wallenberg (2009, p. 28) for Old Icelandic, and Light (2015, p. 27) for Middle English. In OCat the asymmetric distribution of definite and indefinite subjects is limited to *haver* and *ésser* presentational clauses, as illustrated by examples (83-88):

- (83) e conegren fort bé que·l nostre escut hi **era**
 and knew.3PL very well that;the our crest there.CL= was.3SG
 e que nós hi **érem**;
 and that we there.CL= were.1PL

And they knew very well that our crest was there and that we were there.

Fol. 83r, l 25

- (84) (...) e·l rey **era** -hi (...)
 and;the king was.3SG= there.CL

And we went to the Hill, where there was our company (...)

Fol. 178r, l 5

- (85) E Don Garcia no **era** en la tenda (...)
 and Sir Garcia not was in the tent

And Sir Garcia was not in his tent.

Fol. 132 v, l 16

- (86) E en aquela yla **ha** rey.
 and in that island had.3SG king

And in that island, there is a king.

Fol. 27v, l 13

- (87) (...) que greu seria que barayla no hi
 That serious be.3SG.COND that quarrel not there.CL=
hagués gran entre aquels de la ost
 have.3SG.PST.SBJV big between those from the company

e nós.

and us

(...) that it would be serious for a quarrel not to take place between those from the company and us.

Fol. 119v, 127

- (88) e En Rocafort anà ab nós e trobà una egua
and Sir Rocafort went.3SG with us and found.3SG a mare
e cavalcà-hi en es dos, car no hi
and rode.3SG=on;it.CL on the back since not there.CL=
havia son caval.
was.3SG his horse

And Sir Rocafort came with us and he found a mare and he rode on it, since his horse was not there.

Fol. 37v, 122

Examples (83-88) are presentational sentences with the verb *ésser*. In (83), the two coordinated presentational clauses have preverbal pivots that refer back to entities that had already been introduced in the discourse: *el nostre escut* ‘our shield’, which is marked as a definite and known entity by the definite article *el* ‘the’, and *nós* ‘we’, a subject pronoun. In the first clause, the verb is in the 3rd person singular, agreeing with *nostre escut*, ‘our shield’, while in the second one it agrees with the subject pronoun *nós* ‘we’, displaying 1st person plural morphology. In (84), the pivot *el rey* ‘the king’, again, had been previously introduced in the discourse, and it agrees with the 3rd person singular verb. Example (85) contrasts with (83) and (84) in being a main clause. In this case, the pivot, *Don Garcia* ‘Sir Garcia’, also previously introduced in the discourse, occurs preverbally. Examples (86-88) are presentational sentences with *haver*. Like in all the presentational clauses with *haver* that have been used up to this point, they have indefinite pivots, that in both cases, introduce new referents in the discourse. Example (87) is especially interesting, since the pivot *barayla* ‘quarrel’ occurs preverbally, but the adjective *gran* ‘great’, which belongs to the same noun phrase, has been left *in situ*. This example contributes to dispelling the possibility of OCat indefinite postverbal subjects being scrambled rightwards, as suggested by Wallenberg (2009), and backs up our hypothesis of them remaining in their base-generated position. In

(87), the head of the noun phrase, *barayla* ‘quarrel’, undergoes movement to the left periphery to receive emphatic reading. (88) is a presentational clause with *haver* where the pivot can be argued to be an old referent. Sir Rocafort’s horse, while not explicitly and overtly mentioned in the discourse, is elicited by the fact that Rocafort is a knight, and therefore, has a horse. Nevertheless, the reader does not know which particular horse that would be, apart from being Sir Rocafort’s. While the presence of the possessive entails that the noun is finite, it is evoked for the first time in the discourse, and therefore, it has the same information status than the other pivots of *haver* presentational clauses.

Both *ésser* and *haver* presentational clauses are stage-level predicates: they state the existence of *x* in context *y* at a certain point in time. Therefore, they are alsothetic predicates: they report an event, as opposed to categorical predicates, which predicate something of an entity. Presentational clauses with *haver* are associated with all-focus readings, since they introduce new elements to the discourse. *Ésser* presentational clauses, on the other hand, establish the existence of a known element within a location. In these cases, the pivot, containing old information preferentially occurs in a Topic position in the left periphery. This is exemplified by example (84), where the locative clitic *hi* ‘there’, is enclitic to the verb in spite of the presence of a preverbal constituent, the pivot *el rey*. As discussed in Chapter 3, section 3.1.5, proclisis is borne if the preverbal constituent is located in FocP (Batllori et al., 2005). Therefore, enclisis confirms that the pivot *el rey* is located within the Topic field.

In MCat,⁵³ no strict definiteness effect is observed in the distribution of pivots *ésser* and *haver-hi* existential clauses. Definite pivots can appear with either predicate, depending on the information reading that they are to receive:

(89)

- (a) Existential clause with *haver-hi* – postverbal pivot, neutral reading.

Hi **havia** el president.
 there.CL has the president

The president was there.

- (b) Existential clause with *ser-hi*, ungrammatical with postverbal pivot.

Grammatical with right dislocated pivot.

⁵³ When referring to ‘MCat’, we refer to Central Catalan, on which Catalan Standard is modelled. Existential clauses in MCat are subject to dialectal variation. See Rigau (1993, 1997) for an account of this variation.

- i. * Hi **era** el president.
 There.CL was.3SG the president
- ii. Hi **era**, el president.
 there.CL was.3SG the president

(c) Existential clause with *ser-hi*, preverbal pivot, neutral reading.

El president hi **era**.
 the president there.CL was.3SG

(d) Existential clause with *haver-hi* – preverbal pivot. Grammatical if the preverbal pivot receives contrastive focus reading.

- i. * El president hi **ha**.
 The president there.CL has
- ii. EL PRESIDENT hi **ha**.
 The president there.CL has

Adapted from Rigau (1997, p. 397)

While definiteness plays a role in the choice of existential predicate, it does not interact with the location of the subject in clauses containing other predicates. Definite and indefinite lexical and pronominal subjects display preferences in their distribution, but there are no hard and fast rules determining it. Definite subjects from both categories tend to appear preverbally, while indefinite ones have a preference for postverbal positions. We first consider preverbal definite lexical and pronominal subjects, and then preverbal indefinite lexical and pronominal subjects. This is followed by a description of postverbal lexical and pronominal definite subjects, and of postverbal lexical and pronominal indefinite subjects.

Preverbal definite lexical subjects

Preverbal definite lexical subjects can occupy either of the positions that have been listed above in SpecTP or the left periphery. For the reader’s convenience, we reproduce here some of the examples that have been used above to identify the different positions available to subjects with definite lexical subjects:

(90) Lexical subject in frame field, hanging topic (see example (39) for an explanation):

E [nostra mare], sempre que nós **fom** nats, envià·ns
 and our mother as soon as we were.1PL born sent.3SG=us.CL

a Sancta Maria.

to Saint Mary

And as soon as we were born, my mother sent us to Saint Mary's church.

Fol. 3v, l 13

- (91) Lexical subject in topic field, topic shift. This can be seen when considering the context of the clause, where the clause's subject had already been introduced as an IO, *a-N Guerau de Cabrera* 'to Guerau of Cabrera'. In the following clause, the overt subject *En Guerau* appears, shifting the topic back to it.

Context: E en tant enviam ·i· missatger a·N Guerau de Cabrera e els de la vila que atorgàvem la feeltat que·ns tingués En Ramon d'Àger.

And in the meanwhile, we sent a messenger to Sir Guerau of Cabrera and to those from the village, saying that we accept the fealty that Ramon d'Àger should hold towards us.

E·[N Guerau] no **havia** ·l sen de Salamó.
And;Sir Guerau not had.3SG;the judgement of Salamó

And Sir Guerau did not have Salomon's wisdom.

Fol. 25v, l 20

- (92) Lexical subject in Focus field, contrastive reading. In this sentence, the subject *la major part de la partida* 'the greatest part of the squadron', contrasts with the pivot of the preceding existential clause, *una partida de l'estol* 'a part of the squadron', and is thus focalised in the left periphery.

E ach una partida de l'estol a Cambrils, e la
And was a part of the;squadron to Cambrils and the
[major partida], on nós érem, **fo** en lo port de Salou
greater part where we were was.3SG in the port of Salou
e en la plaja (...)
and in the beach

And there was a part of the squadron in Cambrils, and the greater part, where we were, was in Salou's harbour.

Fol. 32r, l 21

Preverbal definite pronominal subjects

Pronominal subjects, like lexical ones, can occur either in SpecTP or in the left periphery. In the LFRJ database, only cases of pronominal subjects in the topic and Focus field are found. There are no cases of pronominal subjects in the frame field.

- (93) Pronominal subject in SpecTP, as indicated by the presence of the high TP adverb (for a more thorough description, see example (38)):

E ladonchs ell **respondria** a aqueles demandes
and then he reply.3SG.COND to those requests

And then he would answer to his requests.

Fol. 197r, l 28

- (94) Pronominal subject in the Focus field, receiving emphatic reading (it is me, the king, that is carrying out the action).

On yo us **prec (...)** que em retats
hence I to;yo.CL beg.1SG that to;me.CL= return.2PL.SBJV

Balaguer (...)

Balaguer

Hence, I beg you that you return Balaguer to me.

Fol. 24v, l 24

- (95) Pronominal subject, topic shift. The null subject of the preceding clause in the context refers to *los missatgers de l'emperador*, 'the emperor's messengers'. In the glossed clause, the pronoun makes reference to another plural collective, previously introduced in the discourse. Thus, it carries out topic shift.

Context: E quan los missatgers de l'emperador enteseren la lur voluntat e que d'altra manera no poria ésser, demanaren-los acort.

And when the messengers from the emperor heard their will, and that it could not be in any other way, they asked them for an agreement,

E éls **donaren-lo·ls** tro en l'altre dia.
And they gave.3PL=it.CL=to;them.CL until in the;other day

And they gave it to them the following day.

Postverbal definite lexical subjects

Definite postverbal subjects are found in three positions: in situ, within the left periphery of *v*P, and right dislocated.

- (96) Verb adjacent postverbal lexical subject discussed in example (65):

Aquesta	paraula	volch	<u>Nostre Seyor</u>	complir	en	los
this	word	wants	Our Lord	accomplish.INF	in	the
nostres	feytz.					
our	deeds					

And Our Lord wanted to accomplish this statement in our deeds.

Fol. 1r, 1 2

- (97) Non-verb adjacent postverbal lexical subject:

E	sempre	envià	ns	missatge	<u>en Berenguer de Finestres</u>	(...).
and	quickly	sent.3SG=to;us.CL	message	Sir Berenguer de Finestres		

And quickly, Sir Berenguer de Finestres sent us a message.

Fol. 25v, 1 22

- (98) Right dislocated postverbal lexical subject, discussed in example (57):

E,	feyt	aquest	matrimoni,	entram	en	Aragó	e	en
and	done	this	marriage	entered.1PL	in	Aragon	and	in
Catalunya,	[<u>nós</u>	<u>e</u>	<u>nostra</u>	<u>muyler</u>	<u>la</u>	<u>reyna</u>].		
Catalonia	we	and	our	wife	the	queen		

And, once this marriage was done, the queen and I went to Aragon and Catalonia.

Fol. 9v, 1 12

Postverbal definite pronominal subjects

Definite pronouns can be found postverbally, verb-adjacent. There are no instances of V XP S_{Pron}, where the pronoun is not verb-adjacent in main clauses from the LFRJ database.

- (99) E per açò e per la gràcia de Déu **havem** nós
and for this and for the grace of God have.1PL we

nom En Jacme.

name Sir Jaume

And this is why, and by the grace of god, we are called Sir Jaume.

Fol. 3v, 127

Preverbal lexical subjects

There is only one case of preverbal lexical subject in the LFRJ database. It is reproduced in (100):

(100) E [i] cavaler **prestà**·ns ·i· gonió leuger (...)
and 1 knight lent.3SG=us.CL 1 chain mail shirt light

And one knight lent us a light chain-mail shirt.

Fol. 7r, 18

Preverbal quantifier subjects

There is only one case of a preverbal quantifier subject, the generic pronoun *tots* ‘all’, in the LFRJ database, reproduced in (101).

(101) E tots **vengren** el dia de la cort (...)
and all came.3PL the day of the court

And all of them came the day of the court.

Fol. 6r, 11

Postverbal indefinite lexical subjects

Postverbal indefinite subjects are more common than preverbal ones. In the LFRJ database, there are 8 cases of postverbal indefinite lexical subjects, contrasting with only two preverbal cases, both verb-adjacent.

(102) E **feriren-lo** [ii] cavalers (...)
and wounded.3PL=him.CL 2 knights

And two knights wounded him.

Fol. 15r, 112

Postverbal pronominal subjects

Like in the case of lexical subjects, postverbal pronominal subjects are more common than preverbal ones: in the LFRJ database, there is only one case of preverbal non-specific

pronominal subject, whereas there are 12 cases of non-specific or indefinite postverbal pronominal subjects, all verb-adjacent. Among them, we find the only negative pronoun found in the LFRJ database's main clauses, *negú* 'no one', as shown in (103), generic referential pronouns, as in (104), and indefinite non-referential pronouns, as in (105). They all appear adjacent to the verb. In (106) we see that it was also possible for indefinite subjects to be right dislocated.

- (103) E çó que Déus vol no **pot** negú desviar ni
 and that which God wants not can anyone divert.INF nor
 tolre.
 leverage.INF

And what God wants, no one can divert or leverage.

Fol. 27v, l 15

- (104) E aquí, **juraren-nos** tots que·ns gardarien nostre
 and here swore.3PL=to;us.CL all that=to;us.CL guard.3PL.COND our
 cors e nostres membres e nostra terra (...).
 body and our limbs and our land

And here, all of them swore to us that they would guard for us our body, our limbs and our land (...).

Fol. 6r, l 2

- (105) E aytal semblança **pot** hom fer de vós.
 and such likeness can.3SG one do.INF of you

And one can find such likeness in you.

Fol. 31r, l 26

- (106) E ab ·i· trebeyl que feya ab uns moços,
 and with 1 game that did.3SG with some boys
 donaren-li d'una teula en lo cap [·i· d'aquels qui
 gave.3PL=him.CL of;one tile in the head 1 of;those who
 jugaven a aquel joch contra él].
 played.3PL at that game against him

And when he was playing with some boys, they hit him with a tile on the head, one of those that were playing at that game against him.

Fol. 9r, l 14

In the light of this evidence, we establish that OCat did not have a definiteness effect, and that the distribution of definite and indefinite subjects depends on their informational value, rather than their [\pm definite] specification.

4.3.3 Summary

In 4.2, we have explored the distribution of subjects in the LFRJ database's main clauses. It has been shown that in OCat, there was a specialised subject position in a preverbal position: SpecTP. Other positions available to subjects have been identified: subjects can appear dislocated in the left periphery (be them CLDed or HTs, a matter that has been shown to be difficult to disentangle), in SpecTP, in their base-generated position, and finally, right dislocated. Furthermore, it has been shown that elements other than nominal or pronominal expressions can occupy SpecTP, suggesting that the featural make-up of this projection is not only linked with a [+D] feature, but also with a discourse grounding one. Finally, the distribution of subjects has been shown to be linked to its informational reading and the information structure of the clause (whether it receives a categorical orthetic reading) rather than its definiteness, even though the two sometimes correlate.

4.4 Subject distribution in Old Catalan embedded clauses

Positions available to subjects in embedded clauses are the same as those available in main clauses. Furthermore, the tendency of occurrence of certain subject types in certain positions does not show significant discrepancies between main and embedded clauses, and differences can be explained by the main/embedded clause asymmetry. This contrasts sharply with data from other Old Romance languages, most notably OFr, where SVO orders are the overwhelmingly frequent in embedded clauses, while they are much more restricted in main clauses given the non-subject specialisation of the immediate preverbal position. Nevertheless, there are some differences in the distribution of definite and indefinite pronominal and nominal lexical subjects.

In OFr (Adams 1987b, p. 5; Vanelli, Renzi & Benincà 1987, p. §4.2; Roberts 1993, p. 142; Vance 1997, p. 133; Salvesen 2013, p. 140; Wolfe 2015, p. 101), as well as in other Old Romance varieties, among which OSp (Wolfe 2015, p. 129), OOc (Jensen 1994, p. 386), OVen, OSic (Wolfe, 2015a), have an overwhelming preference for SVO orders with overt

subjects in embedded clauses. In Table 7, the frequency of SVO orders in OOc and OSp, neighbouring varieties of OCat, as well as OFr, the best studied Old Romance language, are shown. Percentages for varieties other than OCat have been drawn from Wolfe (2015a).

Language	% of SVO in main clauses	% of SVO in embedded clauses
OFr	46.32%	82.07%
OSp	35.32%	59.76%
OOc	23.78%	76.47%
OCat	24.7%	36.4%

Table 8 – Frequency of SVO orders in main and embedded V2 clauses

As shown in Table 8, OFr SVO order is almost twice as frequent in embedded than in main clauses. The same is true of OSp. In OOc, the frequency of SVO triples from main to embedded clauses. However, in OCat the difference between the frequency of SVO in main and embedded V2 clauses is not as striking: while SVO main clauses account for 24.7% of main V2 clauses and main V3 clauses with preverbal verb adjacent subjects, in the case of embedded V2 clauses and V3 clauses with preverbal verb adjacent subjects, they account for 36.4% of embedded V2 and V3 clauses, with a 11.7% difference in the occurrence of SVO between main and embedded clauses, smaller than in the other Old Romance languages. Furthermore, OCat is the only language in which the percentage of SVO in embedded clauses is below 50%. Therefore, there is no sharp asymmetry in terms of subject position between main and embedded clauses, as it has been described for Romance varieties that have been argued to possess a V2 grammar. This provides us with a further piece of evidence against a V2 analysis of OCat.

Embedded clauses included in the LFRJ database are exclusively completive and adverbial. Completive and adverbial clauses, as it has been discussed above and is discussed at length in Chapter 5, crosslinguistically tend to have a fully fledged left periphery. In Tables 1 and 2 in this Chapter, it was shown that the proportion of null subjects in main and embedded clauses are very similar, with a difference of less than 6%.

Table 9 offers a breakdown of the distribution of subjects in non-pro-drop embedded clauses.

	Non Pro-Drop		
	Structure	Occurrences	%
Preverbal subjects with verb adjacency	S _{LEX} V	151	77.7% (372/479)
	S _{PRO} V	217	LEX: 40.6% (151/372)
	DO S _{PRO} V	3	
	ADV S _{PRO} V	1	PRO: 59.4 % (221/372)
	Total	372	
Preverbal subjects - non-verb adjacent	S _{PRO} DO V	3	2.3% (11/479)
	S _{PRO} DO V	2	LEX: 45.5% (5)
	S _{LEX} PP V	5	
	S _{LEX} ADV V	1	PRO: 55.5 (6)
	Total	11	
Postverbal subjects with verb adjacency	ADV V S _{LEX}	5	16.3% (78/479)
	ADV V	4	LEX: 69.2%(54/78)
	S _{PRON}		
	DO V S _{LEX}	2	PRO: 30.8% (24/78)
	DO V S _{PRON}	5	
	PP V S _{LEX}	8	
	PP V S _{PRON}	2	
	SC V S _{LEX}	1	
	V S _{LEX}	38	
	V S _{PRO}	13	
	Total	78	
Postverbal subjects non-verb adjacent	ADV V DO	2	3.7% (18/479)
	S _{LEX}		LEX: 100% (18/18)
	PP V PP S _{LEX}	1	
	PP V ADV	1	100% (479/479)
	S _{LEX}		
	V ADV S _{LEX}	3	
	V PP S _{LEX}	3	
V DO S _{LEX}	1		

	V SC S _{LEX}	7	
	Total	18	
		479	

Table 9 – Subject distribution of non-prodrop embedded clauses in the LFRJ database

Preverbal subjects with verb adjacency account for 77.2% of the total. In main clauses, that percentage is of 69.5%, with a difference of 7.7%. Within preverbal subjects, in embedded clauses, lexical subjects make-up for 40.5%, and pronominal ones for 59.5%. In main clauses, those percentages are of 46% and 54% respectively, therefore, differences between the frequency of preverbal verb adjacent lexical and pronominal subjects is of less than 5%. Verb adjacent postverbal subjects are also distributed in a parallel fashion: in main clauses, they account for 28.6% of the total, while in embedded clauses, they only account for 16.6% (this difference of 12% can be linked to the nature of predicates that occur in embedded clauses, a matter that is covered below). In main clauses, postverbal verb adjacent subjects are, in 74.4% of cases nominal, and pronominal in 25.6% of them. In embedded clauses, percentages are quite similar: lexical subjects represent 69.2% of cases of postverbal verb adjacent subjects, while pronominal ones 30.8%. These percentages differ 5.2% between main and embedded clauses. These percentages confirm the intuition that, at least linearly, the distribution of subjects in main and embedded clauses is parallel.

The low frequency of certain predicate types interferes with the frequency of postverbal subjects in embedded clauses. Table 9 shows the predicate types of embedded clauses with postverbal subjects. When compared to Table 4, which shows the frequency of predicate types in main clauses with postverbal subjects, it becomes apparent that embedded clauses present a much lower rate of *verba dicendi* (1.8% in embedded clauses vs. 27.1% in main clauses), unaccusative verbs (36.4% in main clauses and 25.6% in embedded clauses), and a much higher rate of transitive verbs (16.5% in main clauses, and 44.8% in embedded clauses).

Predicate type			Total
Postverbal nominal subjects – verb adjacent	Copula	5	53.2% (41/77)
	Verba dicendi	1	
	Existential	7	
	Passive	4	
	Reflexive	1	
	Unaccusative	17	
	Unergative	6	
	Transitive	16	20.8% (16/77)
	Total	57	74% (57/77)
Postverbal pronominal subjects – verb adjacent	Copula	-	20% (4/20)
	Verba dicendi	-	
	Existential	-	
	Passive	-	
	Reflexive	-	
	Unaccusative	4	
	Transitive	16	80% (16/20)
	Total	20	26% (20/77)

Table 10 – Predicates in clauses with postverbal subjects in embedded clauses

The difference in frequency of predicate between main and embedded clauses can be attributed to the very nature of embedded clauses: completive embedded clauses tend to occur as direct objects of *verba dicendi*, and therefore, this predicate type is expected to be less frequent in embedded clauses. Unaccusative verbs are also significantly more frequent in main clauses. This might be related to the tendency for unaccusative verbs to denote thetic actions rather than categorical ones, an informational configuration associated with main clauses.

So far, we have established that main and embedded clauses present the same linear distribution of subject types, and that differences are connected to foreseeable variation connected to the different incidence of predicate types. In what follows, it is shown that the same positions are available to subjects in embedded clauses. This results from embedded clauses using material that has already been introduced in the discourse, and therefore, they are definite. It is for this reason that, unlike in section 4.3, we do not make a distinction between definite and indefinite subjects when describing the positions available to them.

Preverbal lexical subjects

Given the presence of the complementiser, in embedded clauses there is no clear test available to identify whether subjects are topicalised or focalised in the left periphery. This is made apparent by (107), where the preverbal subject is followed by a clitic, as expected in embedded clauses. In (108), the subject precedes a focalised DO, and therefore it is topicalised (see Chapter 5 for an overview of the OCat left periphery). No examples of lexical subjects below focalised constituents are found in the LFRJ database, but this configuration is found with pronominal subjects, as it is shown below. However, there are a few cases that could suggest that preverbal lexical subjects can appear in the Frame Field, as HTs. This is the case in (108), where the subject is separated from the finite verb by a subordinate temporal clause, which can be argued to be located in the Frame Field (see Chapter 5 for an extensive discussion on the make-up of the OCat left periphery).

(107) per ço car [ell e N' Exemèn d'Orrea, e N' Artal de Luna
 for this since he and Sir Exemèn d'Orrea and Sir Artal de Luna
 sobre perferta de dret], li **havién** feit mal.
 about offering of law to;him.CL=had.3PL done harm

because he, and Sir Exemèn d'Orrea, and Sir Artal de Luna, by right, had harmed him.

Fol. 196v, l 13

(108) Car diu Salamó que [qui perdona a son fiyl
 since says Salomon that he who forgives to his son
 les vergues de castigament], que MAL li **fa**.
 the rods of punishment that harm to;him.CL= does

Since Salamon says that he who spares his son from birching, that he is harming his son.

Fol. 1r, l 22

Preverbal pronominal subjects

Preverbal pronominal subjects can be topicalised and appear above focalised elements, as in (109), and they can also appear in SpecTP, below focalised elements, as in (110). Example (111) shows a prototypical case of S_{PrON} V O where the location of the subject cannot be determined. Therefore, their distribution is parallel to that of pronominal subjects in main clauses.

(109) Car nós allò ne **seguiríem**.
since we that not follow.1PL

Since we would not follow that.

Fol. 195r, l 20

(110) E dixem-los que esta cosa, nós no **faríem**
and told.1PL=them.CL that this thing we not do.1PL.COND
per re (...)
for nothing

And we told them that this particular thing, we would not do for anything.

Fol. 32v, l 26

(111) Si nós **volíem** acabar lo feyt de Balaguer
if we wanted.1PL finish.INF the issue of Balaguer

Postverbal lexical subjects

Postverbal lexical subjects, like in the case of main clauses, can appear verb-adjacent or non-verb-adjacent, and while being more common with certain predicate types, they also occur with transitive verbs. Like in main clauses, there is no need for the presence of a preverbal constituent for the subject to be postverbal.

In (112), the heavy postverbal subject *lo rei Karles de Nàpols* ‘King Charles of Naples’ appears postverbally with a transitive predicate and a preverbal constituent. In (113), the transitive predicate is not preceded by any constituent, and the subject, heavy again, also

appears postverbally. In these two cases we assume the subject to be in its base generated position. In (114-116), postverbal subjects are separated from the verb by a DO in (118), a PP in (115), and an adverb in (116). Since objects mark the boundary of *vP* (Wolfe, 2015), we can safely assume that in (114) the subject is in its base generated position, within the *vP*-VP complex. In (119), the PP is argumental and also located in the *vP*-VP boundary. In (116), the adverb *bé* ‘well’ is located between the verb and the subject. *Bé* ‘well’, is one of the adverbs that Cinque (1999) locates within the lower section of the clause, within *vP*-VP. Therefore, the fact that it precedes the subject suggests that it is located in its base generated position. No cases of right dislocated subjects are found in embedded clauses from the LFRJ database. It is worth noting that right dislocations are scarce in the medieval Romance record.

(112) car per l’Església lo **tenia** [lo rei Karles de
 since for the;Church him.CL=had.3SG the king Charles of
Nàpols] (...)
 Naples

(...) since because of the Church, king Charles held him (...)

Fol. 194r l 18

(113) E acordaren-se que ns **nodrís**
 and agreed.3PL=themselves.CL that =us.CL reared.3SG
 [lo mestre del Temple en Montsó].
 the master of;the Temple in Montsó

And they agreed that we would be brought up by the Master of the Templar order.

Fol. 5v l 20

(114) car si no hi anats, no **haurà** lo castell
 since if not there.CL= go.2PL not have.3SG.FUT the castle
 [la comtessa] (...)
 The countess

Since, if you don’t go there, the countess won’t have her castle (...).

Fol. 26v l 16

(115) Que la **presés** per muyler [Don Pero Aonés].
 that her took.3SG as wife Sir Pero Aonés.

That Don Pero Aonés should take her as wife.

Fol. 8r l 16

(116) (...) car no·s volien bé [nostre pare
 because not=each other.CL wanted.3PL well our father
ni nostre mare].
 nor our mother

(...) because our father and our mother did not love each other.

Fol. 28r l 15

Postverbal pronominal subjects

Postverbal pronominal subjects can appear verb-adjacent, like in (117) and (118). In (117), the indefinite 3rd person subject *hom* ‘one’ appears between the inflected modal *podia*, ‘was able to’, and its infinitival complement (see section 4.3.1 in this Chapter for a more detailed analysis of V SUB INF sequences in sentences with modal predicates), possibly located in the left periphery of *vP*. Postverbal pronominal subjects can also be separated from the inflected verb by a constituent, as in (118), where the pronoun *nós* ‘we’ is separated from the verb by an argumental prepositional phrase, *en aquel siti* ‘in that siege’. Since the subject follows an argumental phrase, it is taken to be in its base-generated position. Alternation between V S XP and V XP S is considered in section 4.3 in this Chapter and in Chapter 5, where the low periphery of OCat is considered.

(117) (...) bé durà per ·viii· dies que no **podia**
 well last.3SG for 8 days that not could.3SG
hom fer carrera en la host.
 one make.INF way in the army

it last for a good 8 days that one could not make way with the army.

Fol. 40r, l 14

(118) E creem que **estiguem** en aquel siti nós entorn
 and believe.1PL that remained.1PL in that siege we around
 de ·ii· mesos
 of 2 months

And we believe that we stayed in that siege around two months.

In this section, two main conclusions have been reached: (i) that there is no asymmetry in subject distribution between OCat main and embedded clauses from the LFRJ database, a fact that sets OCat apart from its neighbouring varieties, and that (ii) the same subject positions are available in main and embedded clauses.

4.5 SUMMARY

In this Chapter, the position of subjects in the LFRJ database has been explored in depth. Firstly, it has been demonstrated that like M_{Cat}, O_{Cat} was an NSL. This has been followed by a thorough description of the different positions available to subjects in O_{Cat}: they can appear in the Frame field as HT, in the Topic field, left dislocated, in SpecTP, where they are attracted by the featural make-up of the projection, in their base generated position, when something else satisfies the EPP feature on SpecTP, and finally, right dislocated. The possibility of elements other than nominal expressions occurring in SpecTP (locative and temporal adverbial expressions) has been considered, confirming what had already been hinted in Chapter 3: that SpecTP is a position that can host elements grounding the clause in the discourse (in the ‘here and now’). Finally, it has been extensively shown that the distribution of subjects is parallel in main and embedded clauses. Therefore, just like in the case of verb position in Chapter 3, we establish that the O_{Cat} data from the LFRJ aligns with M_{Cat} in terms of (i) positions available to subjects, (ii) the lack of expletives, and (iii) the symmetry in subject distribution in main and embedded clauses.

5 OLD CATALAN INFORMATION STRUCTURE

5.1 Basic notions

This chapter discusses the link between syntactic positions and Information Structure in OCat, understanding Information Structure (IS) as defined in Gundel (2012, p. 585): information structure refers to the morphosyntactic and prosodic organisation of an utterance depending on what it is about, and what the common ground (knowledge shared by the speaker and the addressee) is assumed to be. IS has been described and analysed from several perspectives: focusing on the prosody of utterances (Büring, 1997, 2003); from a semantic-pragmatic perspective (Chafe, 1976; Frascarelli & Hinterhölzl, 2007; Reinhard, 1981); or from a syntactic angle (Batllori & Hernanz, 2013; Batllori et al., 2004; Benincà, 2004, 2006; Benincà & Poletto, 2004b; Donaldson, 2016; Eide & Sitaridou, 2014; Martins, 2011; Rizzi, 1997; Sitaridou, 2015, 2012; Wolfe, 2016, to mention a few). The study of the syntactic encoding of IS in the Romance languages has greatly contributed to our understanding of the relation between syntax and pragmatics, culminating in the Cartographic Project, which we assume in this thesis given its suitability to describe relations between information value of elements and their syntactic distribution, as well as the wealth of empirical data that supports it. Furthermore, given the nature of our data, exploring IS through its syntactic reflexes allows us to describe OCat IS without having to rely on a semantico-pragmatic analysis of the informational value of utterances and their parts, since, as Benincà (2004, p. 249) puts it, ‘judgements are elusive and slippery’ enough for modern languages, and even more so for languages without accessible speakers.

IS is generally considered to have a bipartite structure, whereby old and new information receive a different treatment. The terms ‘topic’, ‘theme’ and ‘ground’ have been used to refer to old information – information shared by the speaker and the interlocutor, while ‘focus’, ‘rhema’ and ‘comment’ have been used to refer to newly introduced information. As stated by Gundel (2012) and Gundel & Fretheim (2004), there is no agreement in the literature regarding IS primitives or their nature (whether it is syntactic and linked to the derivation of the clause or simply prosodic, for instance). For example, there is controversy around the notion of contrast and whether it is an information structure primitive with syntactic reflexes, a pragmatic interpretation that can be combined with other notions linked to IS (Vallduví [1992] for MCat, Vallduví & Vilkuna [1998] for MCat and Modern Finnish, Frascarelli & Hinterhölzl [2007] for MIIt and Modern German) or a feature inherently linked with the notion of Focus (Chafe, 1976; Rooth, 1992). However, there is crosslinguistic

evidence pointing towards the existence of syntactic encoding of contrast (Kiss [1998] for Hungarian, Rizzi [1997], Belletti [2001, 2004] for MIT, Benincà & Poletto [2004] for Medieval Romance).

In line with the cartographic literature on Old Romance word order and information structure, we analyse OCat's IS by identifying syntactic positions in the left periphery that can be linked to specific IS readings (Benincà, 2004), thus allowing us to draw a map of OCat left periphery. In (1), we reproduce Benincà's (2006, p. 13) map of the Old Romance left periphery, against which OCat is contrasted in this Chapter:

- (1) [Force C° [Rel*wh* C°] / {Frame [ScSett][HT] C°} {Topic [LD] [LI] C°} {Focus [I Focus] [II Focus] / [Inter*wh*] C°} [Fin C°]

Benincà (2006, p. 13, example 18)

Benincà (2006) divides the left periphery in three fields: The Frame Setter field, the highest in the left periphery, where she argues that Scene Setters (ScSett) and Hanging Topics (HT) are located in Medieval Romance, a Topic field, located just below, which hosts Left Dislocated Topics (Left Dislocation, LD) and Contrastive Topics (List Interpretation, LI), and finally a Focus field, where Information Foci (Focus I) and Contrastive Foci (Focus II) are hosted respectively. Benincà (2004, p. 246) states that '[i]t is important to preliminarily underscore that I use the labels Topic and Focus to refer to syntactic objects,⁵⁴ putting aside their precise pragmatic values'. While syntactically we follow Benincà's description of left-peripheral topics, we attempt to offer a more precise pragmatic description of the types of foci and topics that can occur in the OCat left periphery. In what follows, the notions topic and focus are explored, and the labels used in this work presented.

The term 'common ground' (Prince, 1988; Stalnaker, 1978), used by Vallduví (1994) and Vallduví & Vilkuna (1998) as 'ground', is used to refer to 'what the speaker assumes the hearer knows or believes to be true and is attending at the time of the utterance' (Vallduví 1996, p. 465). Therefore, something need not have been mentioned in the preceding discourse to be part of the common ground, which can contain shared knowledge between the speaker and the hearer that is not overtly mentioned. Gundel & Freitheim (2004) use the term 'relational' to refer to whether an element is given or new in relation to the common ground

⁵⁴ When referring to the syntactic Topic and Focus projections, the term is capitalised, while it is spelt in lower case when referring to the semantico-pragmatic notions.

(Fraser [1998] ‘discourse topic’), while they use the term ‘referential’ to refer to whether an element or not has been overtly mentioned in the discourse. Relational givenness and newness need not go hand in hand: an element can be active in the mind of the speaker and the hearer without having been mentioned in the discourse. The term ‘topic’ is used to refer to given information, be it relational or referential. Another term often used to refer to the common ground is ‘background’ (Leonetti, 2017).

Different types of topics and foci have been identified in the literature. For instance, Steiner (2015) and Petrova & Solf (2009) follow Frascarelli & Hinterhölzl (2007) in distinguishing three types of topics: aboutness topics, which establish what the sentence is about, contrastive topics, which create opposition among possible topics, and finally, familiar topics, which usually express topic continuity and normally take a pronominal form. Frascarelli & Hinterhölzl (2007) argue that each topic type is hosted in a different projection in the left periphery, hierarchically organised as follows:

(2) Shifting Topic [+aboutness] > Contrastive Topic > Familiar Topic

Nevertheless, Cruschina (2012, p. 20) presents several points arguing that the syntactic differences between these topic types are purely interpretative and subject to the context (among the arguments he presents in this direction there is the possibility for an unlimited number of topics to occur in the left periphery [Cinque, 1990] and the fact that associating [+aboutness] with shifting topics does not account for the fact that [+aboutness] is also a feature of continuing topics). Cruschina (2012, p. 11) distinguishes between aboutness topics, ‘what the sentence is about’, and referential topics, which crucially share ‘anaphoric information with respect to the previous discourse and mental state of the interlocutors’ (they are referentially and relationally given). Thus, ‘aboutness topics’ encompass Frascarelli & Hinterhölzl (2007) ‘shifting topics’ and ‘referential topics’ are equivalent to Frascarelli & Hinterhölzl (2007) ‘familiar topics’.

With this in mind, the labels used to describe the interpretation attributed to Topics in this work is the following: we use the term ‘discourse topic’ to refer to relational non-referential topics that are inferable from the common ground. The term ‘shifting topic’ is used for referential topics that have been overtly mentioned in the discourse ([+anaphoric]), to which the speaker returns. Finally, the term ‘continuing topic’ is used to refer to a topic that simply reiterates the topic of the preceding clause. The distinction between shifting and

continuing topics has proved important for historical Romance linguistics for the analysis of the function of *et*, as shown in Chapter 3, section 3.1.2, and thus, we decide to maintain it.

Regarding contrastive topics, Cruschina (2012) rules out their existence in the Romance languages since they do not have a specific syntactic configuration associated to them, contrary to contrastive foci. As it is shown below, data from the LFRJ database supports Cruschina (2012) stance. However, it is worth noting that Frascarelli & Hinterhölzl argue for contrastive topics to be a specific tonal event, something that cannot be tested for OCat. Nevertheless, we assess whether other elements have a contrastive flavour.

Finally, ‘stage topics’ (Ertschik-Shir, 1997; Lahousse, 2011; Leonetti, 2015, 2017) introduced in Chapter 3, refers to the spatio-temporal context of the utterance, and they can be implicit or explicit, and overt or covert, as we have seen in Chapters 3 and 4. Stage topics have also been referred to as ‘scene setters’ (Bentley et al., 2015, p. 65) or ‘frame setters’ (Chafe, 1976, p. 50; Krifka, 2007). In Chapters 3 and 4, the term ‘scene setter’ has been used to refer to what has otherwise been labelled as *pro*_{LOC}, a null locative element that grounds predicates in the ‘here and now’ of the discourse, which in M_{Cat}, is realised by the locative pronoun *hi* in existential clauses (see Chapter 4, section 4.3.2 for more on this matter). Therefore, we will use the term ‘scene setter’ for overtly realised non-clitic elements that realise this function.

Two main types of foci are distinguished in terms of interpretation. Firstly, ‘information focus’ (IFoc) introduces new referents in the discourse. Secondly, Contrastive Focus (CFoc), signalling out an element against a set of alternatives (Molnár, 2002; Lahousse, 2011). Some authors classify the syntactic operation consisting on the fronting of an element that has already been introduced in the discourse to the left periphery, triggering subject-verb inversion as focal, as reflected used by them to refer to it: ‘*foco débil*’ or ‘weak focus’ (Batllori & Hernanz, 2009), ‘focus fronting’ (Cruschina, 2012). Other authors do not link this operation to focality, as is the case with Leonetti & Escandell-Vidal (2009, p. 182), who name this construction ‘*verum focus*’, or Leonetti (2017, p. 908), who refers to it as non-focal fronting, following Kaiser & Zimmermann (2011), given the fact that fronted elements always contain given information. Since our analysis of this construction in OCat within a cartographic framework locates the landing site of the fronted constituent within the Focus field, and since this construction is inherently connected to emphasis, we refer to it as Focus Fronting.

The scope of foci can vary. Narrow focus only takes scope over one constituent. When more than constituent falls within the scope of focus, broad focus is borne. Finally, when a clause is singled out as a single informational unit and there is no new/old information partition, the label ‘all-focus’ or ‘thetic’ clause is used (Leonetti, 2017).

Cruschina (2012), giving emphatic reading to a constituent that is neither contrasted nor new information, but linked to polarity. Unlike in the case of topics, there is no controversy regarding the link between contrast and focalisation.

In what follows, the left periphery of main and embedded clauses from the LFRJ database is explored and described.

5.2 The Left Periphery in the LFRJ database

In Chapter 4, section 4.2.1, we have explored the different types of constituents that can appear preverbally: noun phrases, functioning as direct objects and subjects moved or based generated in the left periphery, prepositional phrases, functioning as Scene Setters, prepositional complements and adjuncts moved or based generated in the left periphery depending on their function, pronouns, functioning as subjects, adverbs, acting as Scene Setters or adjuncts, moved or generated in the left periphery depending on their function, and embedded adverbial clauses, acting as adjuncts, based-generated in the left periphery (Ernst, 2009; Gallego, 2013; Haumann, 2007; Potsdam, 1999; Störzer & Stolterfoht, 2015).

Benincà (2004, p. 249. ft.5) makes a clear distinction between topics and foci in the left periphery in terms of movement and base generation. She takes all topics to be based generated in the left periphery and all foci to be moved to the left periphery. Among other arguments presented in favour of this hypothesis, there is the fact that a different type of movement (not operator-driven) would have to be hypothesised for topics, which should also account for the formal differences between hanging topics (HT) and clitic left dislocated (CILDed) constituents (see Frascarelli [2004] and De Cat [2007] for base-generated analyses of left dislocation (LD), Cecchetto [1999], Villalba [2000], López [2009] for movement analyses and Ott [2015], Fernández-Sánchez [2017] for biclausal analyses of left and right dislocations respectively). We follow Benincà (2004), De Cat (2007) and Frascarelli (2004), in taking CILDed constituents to be base generated in the left periphery. Together with HTs, we also take scene setting elements to be based generated in the left periphery.

In what follows, we explore the different values that constituents can have when occurring in the left periphery of main and embedded clauses. Following Benincà (2004,

2006) description of the Old Romance left periphery, we start considering scene setting elements, followed by HTs, CILDed topics, and finally, foci.

5.2.1 Scene Setters

Several categories of constituents can act as scene setters, providing a spatio-temporal context for the clause: embedded adverbial clauses (finite and non-finite), prepositional phrases and adverbs. Scene setters can be verb adjacent or can precede other preverbal constituents. This is illustrated in examples (3-7):

- (3) Context: E, ans que fos la batayla, volie·s metre En Simon de Monfort en son poder per fer sa volentat, e volia·s avenir ab él. E nostre pare no u volch pendre.

And, before the battle would start, Simon of Monfort wanted to gain power over him so as to do as he pleased and wanted to reach an agreement with him. And our father did not want to take it.

E	[<u>quan</u>	<u>viren</u>	<u>açò lo comte</u>	<u>Simon e</u>	<u>aquels de</u>	<u>dins</u>],
and	when	saw.3PL	this thecount	Simon and	those of	inside
preseren	penitència (...)					
took.3PL	penitence					

And when count Simon and those inside saw this, they took penitence (...)

Fol. 5r, l 19

In example (3), the adverbial temporal clause headed by the conjunction *quan* ‘when’ occurs preverbally. It anchors the new information of the clause in the discourse by specifying the point in time when the action took place and its participants, ‘count Simon and those inside’. According to Haegeman (2006), peripheral adverbial clauses (i) deictically link their main clause to the discourse, and (ii) are located in a higher projection in the left periphery. This is the case of the embedded clause in (3): the *quan* ‘when’ clause contains deictic elements that link the main clause to the discourse (*açò*, ‘this’). Based on this, it can be hypothesised that the *quan* ‘when’ clause is in the highest layer of the left periphery. No other evidence is found in (3) to locate the *quan* ‘when’ clause in the left periphery.

- (4) Context: E nós romanguem en Carcassona en poder del comte, car él nos nodria e tenia aquel loch.

And we remained in Carcassonne under the control of the count, since he was bringing us up in that place.

E [puys], [passat això], demanaren-nos nostres naturals.
and afterwards happened.PTCP this asked.3PL=us.CL our relatives

And afterwards, after this, our relatives asked for us.

Fol. 5v, l 2

In (4), the scene setters establishing a link between the discourse and the clause are an adverb and a non-finite participial clause consisting of a past participle and a subject, paralleling Latin ablative absolute. Different scene setters can co-occur in the left periphery. Unfortunately, there are not enough instances of co-occurrence of scene setters to establish whether these constituents were ordered hierarchically or whether there is a limit to the number of scene setters that can co-occur.

(5) E, [fenides estes paraules], levà· s En
And finished.PTCP these words stood= himself.CL Sir

Pere Grony.

Pere Grony

Once these words were over, Pere Grony stood up.

Fol. 31v, l 19

In (5), the scene setter element is a non-finite participial clause. Like in the examples above, it connects the new information of the clause to the discourse. Again, there is no evidence in this clause for locating the non-finite clause within the left periphery.

(6) Context: E, quan fom en Ampurdà, sabem que l'infant En Jacme, fiyl nostre, tenia assetjada la Rocha, ·i· castell del comte d'Ampúries. E anam là e faem-lo·n levar.

And, when we reached Empordà, we got to know that prince Sir Jacme, our son, had the Rocha, a castle of the count of Empúries, under siege. And we went there and made him lift it [the siege] from there.

E [en l'endemig] entram en Perpinyà per ver
and in the meantime entered.1PL in Perpignan to see.INF
a la regina de Castella, nostra fiyla (...)
to.ACC the queen of Castile our daughter

And in the meantime, we entered in Perpignan, to see the queen of Castile, our daughter (...)

Fol. 196v, l 28

In (6) we have an instance of a prepositional phrase acting as a scene setter. Once more, the phrase is a signpost for the hearer to know how the new information is (temporally) related to the discourse, and again, the it is found verb adjacently.

- (7) Context: E quan vench aenant, fo parlat matrimoni entre·l rey Don Pedro, nostre pare, e la fyla d'En Guillem de Montpestler, que era dona de Montpestler e de totes ses pertinències, e ela que daria son cors e Montpestler ab totes ses pertinències. E ayxí féu-se el matrimoni, e fo lo seu nom crescut, que hac nom la reyna Dona Maria.

And when he came forward, the marriage between the king Sir Pero, our father, and the daughter of Sir Guillem of Montpellier, who was the lady of Montpellier and all its lands, and that she would give her body and Montpellier with all its lands. And thus, the marriage was done, and her name gained status, since she was called the queen Lady Mary.

E	[<u>puy</u> s]	En	Guillem de Montpestler,	[<u>estant ella</u>	<u>viva</u>],	près
and	then	Sir	Guillem of Montpellier	being she	alive	took.3SG
·i_a·	altra	dona (...)				
one	other	woman				

And then, Sir Guillem de Montpellier, while she was alive, took another woman for wife (...)

Fol. 3r, l 10

In the preceding examples, we have encountered Scene Setters that either immediately precede the verb, or that precede another scene setting element. However, in example (7), a different configuration is found: the first scene setting element, the adverb *puy*s ‘then’, is followed by the DP *En Guillem de Montpestler*, the subject of the clause. This DP is in turn followed by the scene-setting non-finite clause *estant ella viva* ‘while she was alive’. The subject DP differs significantly informationally from the two Scene Setters, since instead of indicating how the new content introduced in the clause fits within the discourse, it refers to

an entity that has previously been mentioned in the discourse, as shown by the context, and therefore, it is a topic. Hence, in this case, we find the following linear sequence:

(8) [Scene setter] [Topic] [Scene setter] V

Does this indicate that there is a recursive Scene Setting projection in the left periphery? Not necessarily. The position of scene setting elements has been extensively studied in the literature, and it has been suggested that they can be freely adjoined to other projections. In the case of temporal adverbial modifiers such as *puy* ‘then’, it is arguable that, as a temporal adverbial modifier, it is adjoined to TP (Cinque & Rizzi, 2010, p. 103). However, given that the relative order between the adverb and the non-finite clause is the same as in (4), and that in both cases, the temporal adverb has discourse deictic properties and the non-finite clause’s subject is old material that also binds the clause with the discourse, two hypotheses emerge: (i) that these elements are indeed located in the left periphery and not adjoined to TP given their discourse deictic properties (Haegeman, 2013), and (ii) that there exists a hierarchy of constituents that occur in this position. Nevertheless, these two matters do not occupy us at present and this matter has to be left to future research.

(9) Context: E levà·s lo comte d’Ampúries e dix: “Açò us diré yo ans de la resposta que·us deuen fer los vostres nobles: (...)”

And the count of Empúries rose to his feet and said. “I will say this to you, before the other noblemen answer you: (...)”

[si hòmens del món àn mala fama], nós sí la
if men of;the world have.3PL bad reputation we indeed it

havem bona (...)

have good

If there are men with a bad reputation in the world, ours is indeed so good (...)

Fol. 29r, l 10

Example (9) differs from the previous examples presented in this section in that the preverbal scene setter element is neither an adverb, nor a prepositional phrase, or a non-finite clause: it is a finite conditional clause introduced by *si* ‘if’. Haegeman (2003) distinguishes two types of conditional clauses: event-conditional clauses, which modify the clause’s event and are integrated in the speech act of the main clause, and premise conditionals, which structure the discourse by providing a context for the associated clause. Event and premise

conditional clauses differ in terms of internal and external syntax. Clause-internally, event conditionals are derived through movement of the subordinating conjunction, which prevents other constituents from accessing their left periphery (Haegeman, 2006; 2010; 2012), while premise conditionals present a fully fledged left periphery. Clause-externally, event conditionals are adjoined to the TP layer, while premise conditionals are adjoined to the CP layer, in a position where they can create a connection with the discourse by echoing it and providing a ‘context from which the main clause can be inferred, or in which the question raised by the main clause arises, or in which the question raised in the main clause is answered’ (adapted from Haegeman [2003, p. 327]). In example (9), we find a premise-conditional clause: it provides us a context against which the main clause is contrasted, and is thus located in the CP.

There are several candidates in the literature for functional projections hosting premise-conditionals. Gärtner (2001, p. 105) proposes the existence of π P, above the CP layer, Cinque (2008), on the other hand, suggests the existence of an HP (a phrase headed by an empty head) above CP, where Haegeman (2012, p. 171) suggests that premise-conditionals could be adjoined. Other proposals suggest that premise-conditional clauses are located within the left periphery, below CP: Paoli (2003, p. 273) locates them in the same position as CILDED topics, within the Topic field, while Munaro (2005) and Planas (2015, n. 28) suggest that they are located in HypotheticalP, a projection located between the Frame and Topic fields. The latter hypothesis allows us to account for example (9) without having to call upon extra projections above CP: if the premise-conditional clause is located in a projection within the low Frame field, it would naturally follow that in (9), the conditional clause is followed by the pronoun *nós* ‘we’, left dislocated in the Topic field to receive a contrastive reading (whether CILDED or HT, we cannot know), which is in turn followed by a focalised element (find an extended discussion on the nature of *sí* in OCat in Chapter 3, section 3.1.6, and in Chapter 4, section 4.4 for a section demonstrating that it did not have expletive or pronominal uses in OCat). Therefore, this example provides us with evidence that the Frame field links the main clause to the discourse beyond the use of deictic temporal or locative adverbs.

- (10) Context: On, nós vos deïm, sobre·ls ·iii· conseyls que vós nos havets demanats, que metats pau en vostra terra e que us ajudem en guisa que aquest fet se pusca complir a honor de vós e de nós: primerament, que fassats pau e treves per tota Catalunya, e aquels que y volran ésser, que·ls metats tots en vostre scrit. (...)

Where, we tell you, about the three pieces of advice that you have asked us for, that you should pacify your land, and that we will help you in such a way that you may achieve this to your benefit and ours: firstly, you should make peace and truces all around Catalonia, and those who will want to be in it, you should write all their names down (...)

E,	[si	<u>negú</u>	<u>no</u>	<u>y</u>	<u>vol</u>	<u>ésser</u>	<u>de</u>	<u>Cathaluyna</u>],
and	if	anyone	not	there.CL=	want.3SG	be.INF	of	Catalonia
NÓS	li	farem	ésser,	si	li	pesa		
we	to;him.CL =	do.1PL.FUT	be.INF	if	to;him.CL=	burdens		
o	li	plau						
or	to;him.CL=	pleases						

And, if anyone from Catalonia does not want to be there, we will make them be there, whether it pleases them or not.

Fol. 29v, l 21

In example (10), like in example (9), we find a premise conditional clause in the left periphery. In this case, it is followed directly by a focalised pronoun (as exposed in Chapter 3, section 3.1.5, the presence of a focalised constituent in the left periphery triggers proclisis, as in (10)).

In the LFRJ database, the two types of embedded clauses considered are: completive embedded clauses introduced by the conjunction *que* and adverbial clauses of different types (temporal, causal, purpose, finality and conditional clauses). These clause types show parallel distribution of subjects and parallel frequencies in verb position when compared to main clauses, as we have seen in Chapters 4, section 4.3. Completive embedded clauses have thoroughly been described in the literature as having a fully fledged left periphery which allows them to display main clause phenomena. Adverbial clauses also present phenomena that indicate that they have a fully fledged left periphery, as shown below in regard to foci, as well as in regard to their external syntax. This suggests that they are peripheral adverbial clauses, as defined by Haegeman (2003, 2006, 2010), and therefore, as in the case of subjects, we expect parallelisms in the make-up of the left periphery with main clauses.

However, there is a striking lack of scene setters in embedded clauses of either type. The only element that we find that can be identified as being located in the left periphery are

embedded adverbial clauses, but for one case in which we find a prepositional phrase acting as a scene setter:

- (11) car éls li havien promès que, [si aquel
 since they to;him.CL= had.3PL promised that if that
 matrimoni no·s faés], que la **tornassen**
 marriage not=itself.CL= do.3SG.PST.SBJV que her.CL=give back.3PL.SBJV
 a son pare per terra o per mar;
 to her father by land or by sea

(...) since they had promised him that, if that marriage did not take place, that they could give her back to her father by land or by sea (...).

Fol. 10v, 16

- (12) E demanà·ns ·i· penó nostre, per tal que, [si venien
 and asked.3SG=us.CL one standard of ours so that if came.3PL
sos missatgers a la ost], que non los **faessen**
 his messengers to the army that not to;them.CL= do.3PL.PST.SBJV
 mal los nostres, e donam-lo-li de bon grat.
 harm the ours and gave.1PL=it.CL=to;him.CL gladly

And he asked us for a standard of ours, so that, if his messengers came to the army, that our men would not harm them, and we gave it to him gladly.

Fol. 42v, 19

- (13) E tant, que per alongament de les paraules, que·ns
 and such that for lengthening of the words that=ourselves.CL
 avenguem que [al ·v_è· dia] que·n **rendrien**
 agreed.1PL that at;the 5th day that=to;us.CL render.3PL.PRES.COND
 la vila e que començarien d' exir.
 and that and that start.3PL.PRES.COND of= get out.INF

And in the meanwhile, due to the length of the speeches, we agreed that in the 5th day they would render us the village and that they would start to exit it.

Fol. 120r, 112

- (14) “ (...) Responem-li que él nos ho auria a dir,
 answered.1PL=him.CL that he to;us.CL=it.CL=should.3SG to say
 quar, [pus él envià per nós], no li **diríem**
 since since he sent.3SG for us not to;him.CL= say.1PL.COND
 re, si él no deya per què".
 nothing if he not said.3SG why

And we answered him that he should tell us, since, for he had sent for us, we would not tell him anything, if he did not tell us why.

Fol. 44r, l 22

Examples (11-14) show embedded clauses with embedded clauses in their left periphery. In (11) and (12), the embedded clauses in the left periphery are premise conditional clauses, and therefore, are merged in the left periphery of the clause (Haegeman 2006, 2009). Beyond their semantics, a proof of this is the fact that in both (11) and (12), the conditional clause occurs between *que1* and *que2*, the former being located in ForceP and the latter in a projection between the Topic and the Focus field (recomplementation, the occurrence of two *que* ‘that’ within the same clause, is examined in section 5.4.2 within this Chapter). The presence of *que2* marks that the conditional clause is located above the Topic field, where *que2* lexicalises, and we can hypothesise, from the evidence found in main clauses, that they are located within the Frame field. In (13) we find the only example of embedded clause with a prepositional phrase as a Scene Setter. Interestingly, like in (11) and (12), the scene setting prepositional phrase occurs between *que1* and *que2*, marking that it is in the Frame field, above the Topic field.

In (14), the clause located in the left periphery of the *car* ‘since’, a causal adverbial clause is introduced by the conjunction *puy* ‘since’. In this case, there is no *que2* or another element that can contribute to locating this element within the left periphery. However, following Haegeman (2003, 2009, 2010), we suppose that the *car* ‘since’ clause hosting it has a fully fledged left periphery, or at least a left periphery in which the projection associated to external adverbial clauses projects, in the higher left periphery. More data would be required to reach a more specific description of the distribution of scene setters in the left periphery of OCat embedded clauses.

In this subsection we have seen the different elements that can occupy the highest projections of the left periphery of main clauses: finite temporal clauses, non-finite clauses,

adverbs, prepositional phrases and conditional clauses. It has been shown that these elements can occur in the left periphery, where they are base generated (Benincà, 2004) without the need of any other projection being filled (they can be verb-adjacent). They can also be followed by topical and focalised elements, suggesting that OCat's left periphery is similar to that sketched by Benincà (2004, 2006) for Old Romance. In the case of embedded clauses, the scarcity of cases of embedded clauses with scene setters makes it difficult to locate scene setters within their left periphery. Nevertheless, given the parallelisms that we have already established between main and embedded clauses and the clause types found in the LFRJ database, which crosslinguistically behave like main clauses, we can hypothesise that they can host the same scene setters as main clauses.

5.2.2 Hanging Topic Left Dislocation

Cinque (1977) first described HTs, drawing a parallel between them and the Latin construction *nominativus pendens*, by which DPs in the nominative case occur to the left of a clause, without being syntactically connected to it, and which can be retrieved in the clause by an anaphoric clitic pronoun (Rodríguez Molina, 2010, p. 1398). While there is great controversy regarding the syntactic derivation of CILDs, the literature favours a base-generation account for HTs, given their lack of compulsory binding with a resumptive pronoun in the clause core, as phrased by Sturgeon (2008, p. 130): 'There is substantial evidence in support of movement of the left dislocate from a clause-internal position to a position in the left periphery in CILD and a parallel lack of evidence of such movement in Hanging Topic Left Dislocation (HTLD)'. This position is also defended in Frascarelli (2000, p. 170) for MIIt, Bouzouita (2008, 2014) for OSp, Badan & del Globo (2010) for Mandarin Chinese, among others.

The literature on the interpretation of HTLD in the Romance languages consistently evokes two elements: topic shift from the previous sentence's topic, contrasting with it, and introduction of a topic from the common ground Cinque (1997, p. 95); Villalba (2000, p. 97ff).

Villalba (2000) summarises the features that formally distinguish CILD topics from HTLD in MCat (features parallel to those sketched for Old Romance HT in Benincà [2006]):

	CILD	HTLD
i. Category neutral	Yes	DPs only
ii. Iterative	Yes	No
iii. Non-root contexts	Yes	No
iv. Free ordering of the dislocates	Yes	-
v. Obligatory resumptive clitic	Yes	No
vi. Ordering with respect to wh/C	C-CILD-wh	C-HTLD-wh
vii. Connectedness	Yes	No
viii. Island sensitivity	Yes	No

Table 1 – From Villalba (2000, p. 81)

Villalba (2000) notes that in M_{Cat}, as well as in other Modern Romance languages (such as M_{Sp}, M_{It} and M_{Fr}), HTLD does require the presence of a resumptive element within the clause, but it need not be a clitic pronoun:

- (15) [La Maria], en canvi, tothom en parla malament.
the Mary in change everybody of;her=speaks badly

Maria, instead, everybody talks badly of her.

Villalba (2009, p. 90)

- (16) [La Maria], en canvi, tothom parla malament d'ella.
the Mary in change everybody speaks badly of;her

Maria, instead, everybody talks badly of her.

Villalba (2009, p. 91)

As it can be appreciated, the HTLDed DP occurring in the left periphery is not case marked (it is not preceded by a preposition, as expected for the prepositional complement of the verb *parlar* ‘to speak’, which is normally introduced by *de* ‘of’), and while it has a resumptive element in the main clause, it is not a clitic pronoun, but a PP containing the strong pronoun *ella*, ‘she’. These are the two most salient features of HTs.

In the LFRJ database has a few potential cases of HTDL, which include DPs and pronouns:

- (17) Context: E ·i· dia, entre hora nona e vespres, En Ramon de Muntcada havia la guayta dels fenèvolts de dia e de nuyt. E, él estan a la gayta, eren ab él **En Sanço Péreç de Pomar**, fiyl d'En Pero de Pomar, e En Bardoyl, qui era batle seu de Castelserà, e Arnau de Robió, cavaller. E quan viren los hòmens de Balaguer que tan pochts eren, e En Guillem de Cardona, qui era de dins, exiren ab cavals armats per lo mur, que y agren feyt portell, e entraren en lo vayl amagadament ab fayles seques e ontades de sèu. E nós érem en la tenda d'En Guillem de Cervera, que l'haviem vengut veer, e estàvem parlan ab él, e cridaren: "A armes, a armes, que·ls fenèvolts vénen cremar, que fayles aduen encesas!" e havia En Guillem de Cardona ab sí tro a ·xxv· cavallers armats e ·cc· hòmens a peu entre aquels qui portaven les fayles e·ls altres. E exí ab él Sire Guilleumes, fiyl del rey de Navarra, que havia haüt d'una dona.

*And one day, between the ninth hour and the evening mass, Sir Ramon of Muntcada had the catapult watch day and night. And while he was doing the watch, **Sir Sanço Pére of Pomar**, son of Sir Pero of Pomar, and Sir Bardoyl, who was his officer in Castelserà, and the knight Arnau of Robió were with him. And when they saw that there were so few men from Balaguer, and with Sir Guillem of Cardona, who was in, they went out through the walls with horses, and they reached the gate, and entered the valley discretely with torches soaked in fat. And We were in the tent of Sir Guillem of Cervera, whom we had come to see, and we were talking to him, and they shouted: 'to arms, to arms, that they come here to burn the catapults, that they bring lit torches!', and Sir Guillem of Cardona had with him more than 25 armed knights and 200 men on foot, among whom the torches were distributed. And Sir Guilleumes, son of the king of Navarre, whom he had had from a woman.*

E	[<u>Don Sanç</u>	<u>Péreç</u>	<u>de</u>	<u>Pomar</u>];	no	li _i	pòch
and	Sir Sanç	Péreç	of	Pomar	not	him.CL=can.3SG	
soffrir lo	cor	que	ss'	aturàs	e	començà	
suffer the	heart	that	himself.CL=	stopped.3SG	and	started.3SG	
a fugir	e	vench	-se	a	la	host (...)	
to scape.INF	and	came.3SG=REFL.3SG.CL		to	the	army	

And Sir Sanç Péreç de Pomar’s heart could not endure this much and it stopped, and he started scaping and came to the army.

Fol. 40v, l 20

- (18) Context: “Seyor, nós havem demanat a·N Pere Martel de ço que creem que a vós plaurà, d’una yla que ha nom Maylorques. E en aquela yla ha rey, e dejús aquel regne ha altres iles, Manorques e Iviça, e aquestes són subjugades al rey de Maylorques.”

“Sir, we have asked Sir Pere Martel about something that we think will please you, about an island that is called Majorca. In this island there is a king, and under this kingdom there are islands, Majorca and Ibiza, and they are subjugated to the king of Majorca.”

[Ço que	déus	vol]	no	pot	negú	desviar	ni
that which	God	wants	not	can.3SG	no one	tweak.INF	nor

tolre (...)

take away.INF

What God wants, no one can tweak or take away (...).

Fol. 27v, l 15

In (17), the HT is *Don Sanç Péreç de Pomar* ‘Sir Sanç Péreç of Pomar’. The DP is not case-marked, however, it is coreferent with the dative clitic pronoun *li* ‘to him’, corresponding to the description of HT sketched above. As shown in the context of (17), Sir Sanç Péreç of Pomar is introduced in the discourse prior to the sentence at hand, in a fragment whose discourse topic is Ramon of Muntcada’s catapult-watch, during which they are attacked. Sir Sanç and other characters are introduced at the beginning of the fragment, as companions of Guillem of Muntcada. After this, the attack from those from Balaguer is described, and the intra-homodiegetic narrator specifies its location during the action. Immediately afterwards, the forces of the attackers are described: they counted with many men on foot carrying torches, several knights, and a bastard son of the king of Navarre. It is after this description of the fearsome enemies, we find the clause containing ‘Sir Sanç’, who, scared by the sight, deserts his post. The re-introduction of ‘Sir Sanç’, shifts the sentential topic from the attackers to the defenders of the catapult. Therefore, (17) parallels analyses that link M_{Cat} HTLD to topic shift.

In (18) the DO relative clause headed by the demonstrative pronoun *ço* ‘this’ occurs in the left periphery. In spite of being headed by a seemingly demonstrative element, the pronoun serves as the neuter antecedent the relative clause that follows (see Hinzelin [2006, 2010] for an analysis of OCat neuter pronouns). The neuter antecedent of the clause can play a role within the main clause, in this case, direct object. Two informational analyses of this constituent can be proposed. Firstly, it can be hypothesised that the clause introduced by *ço* is focalized, given that it lacks a clitic copy in the core of the clause, and that it has not been mentioned in the preceding discourse. However, an alternative reading is available: if *ço que Déus vol*, which can be idiomatically translated as ‘God’s will’ exists as a referent in the speaker’s and hearer’s common ground, it can be evoked from it without having been overtly mentioned, carrying out topic shift from the topic of the preceding clause, *Menorques e Iviça* ‘Minorca and Ibiza’. The fact that there is no resumptive element in the core of the clause should not be an impediment for us to analyse this clause as a HT, as noted by Villalba (2000) for MCat and Cinque (1997) for MIIt. Nevertheless, grammaticality judgements of HT without resumptive elements show a great inter-speaker level of variation in the modern varieties, which inclines us to favour a Focus analysis of this constituent.

Other candidates to be HT in OCat are preverbal subjects. As it has been shown in Chapter 4, the absence of subject clitics in OCat makes it impossible to formally distinguish whether subjects occurring in the left periphery are CILDed or HTs.

In conclusion, (17) suggests that HTLD in the LFRJ database is akin to HTLD in MCat and the other Old Romance languages in that it presents a resumptive element in the core of the clause. However, more evidence would be needed to offer a thorough description of this syntactic structure in OCat. Regarding a comparison between the distribution of HTs between MCat and the LFRJ data, the scarcity of HT in the LFRJ database makes it impossible for us to add an OCat column to Villalba’s (2000) table and to establish the order of constituents relative to HTLD.

In terms of embedded clauses, there are no examples of HTLD.

5.2.3 Clitic Left Dislocated Topics

As shown in Table 1, extracted from Villalba (2000, p. 81), CILDed topics differ significantly structurally from HT in MCat. Benincà (2004) proposes parallel features for CILD in Old Romance. Nevertheless, the syntactic nature of CILDed constituents is still a matter of debate in the literature, with proposals arguing for them being base-generated within the clause,

moved within the clause, or base generated within an adjacent clause that is later deleted in Phonetic Form, as it has been sketched above. Also as mentioned above, we take them to be base generated.

Prepositional phrases are the clearest cases of clitic left dislocation, since they are overtly case-marked. Subject DPs and most [+human] DO, however, are not usually overtly case-marked (Differential Object Marking, henceforth DOM), and therefore, it is difficult to distinguish whether they are HT or CLDed constituents. This is shown in examples (19-21): in (19) and (20), we have two cases in which a [+human] DO is marked with the preposition *a*, the differential object marker that is also found in M_{Cat} and M_{Sp}:

(19) (...) on tenia pres Don Rodrigo [a Don Lop
 where had.3SG taken Sir Rodrigo to Sir Lop

d'Alvero]

of;Alvero

(...) where Sir Rodrigo had Sir taken Sir Lop d'Alvero.

Fol. 7v 1 6

(20) (...) que yo us dic que, si [a Don Pero
 that I to;you.CL= say.1SG that if to Sir Pero

Aonés ferits (...)]

Aonés wound.2PL

And I say to you that, if you wound Sir Pero Aonés (...).

Fol. 15v, 1 24

(21) (...) que li enviàssem [Don Nuno].
 that to;him.CL= sent.1PL Sir Nuno

(...) that we sent him Sir Nuno.

Fol. 44v, 1 20

In (19), the appearance of the case marker *a* is linked to the fact that the subject and the DO are both postverbal and follow the verb. To avoid confusion between the subject and the DO, the latter is overtly marked. In (20), it could be argued that the fronting of the DO might be linked to the appearance of the differential object marker, given that due to its preverbal position, it could easily be interpreted as a subject. In (21), like in most cases of

[+human] DO in the LFRJ database, we find a [+human] DO, Don Nuno ‘Sir Nuno’, not case-marked. Note that examples (19-21) come uniquely from embedded clauses: (19) is an adverbial relative clause, (20) a conditional clause, and (21) a completive clause. No instances of DOM are found in main clauses. In main clauses, [+human] DO are not generally case-marked, as in (21).

Consequently, cases of CILDed constituents generally come from PPs, as is the case in (22) and (23), two examples of CILD in main clauses. In both, the CILDed constituent counts with a pronominal clitic copy in the core of the clause. In both cases, the CILDed constituent is the experiencer of the predicate *plaire* ‘to please’ and *pesar* ‘to burden’ respectively.⁵⁵

- (22) Context: E quant nós fom de edat de ·ix· ayns, e que no·ns podien aturar en Montsó, a nós ni al comte de Proença, tant ne volíem exir, (...) foren-se bandos e partides entre·ls richs hòmens d’Aragó. Don Pero Ahonés, (...) e altres qui a nós no memhren, e de richs hòmens e de cavallers, faeren bando e partida ab lo comte de Roselló Sanxo, qui era lur cap, e seguien la sua carrera. (...)

E quan los frares viren que·l comte de Proença se n’era anat, sí que no·ls ho féu saber, enteseren que la nostra estada no·ls era bona.

And when we were 9, they could not retain us in Montsó, us or the county of Provence, (...) and prominent men divided in different sides and groups. Sir Pero Ahonés, and others whom we don’t remember, and prominent men and knights, took the side and joined the group of the count of Rosselló Sanxo, who was their leader. (...)

And when the friars realised that the count of Provence had left, although he had not told them, they understood that our presence there was not beneficial for them.

E	[al	comte don	Sanxo];	quan	ho	hoí,	
And	to;the	count sir	Sanxo	when	it.CL=	heard.3SG	
pesà-li;		molt	la	anada	del	comte de	Proença.
burdened.3SG=him.CL		much	the	going	of;the	count of	Provence

⁵⁵ Montserrat Batllori (personal correspondence) points out that in the case of *plaire* ‘to please’ and *pesar* ‘to burden’ are psychological verbs, and that the experiencer may be an applicative, instead of a CILD topic. For the sake of space, this line of research is not pursued in this work.

And count Sir Sanxo, when he heard it, the leaving of the count of Provence saddened him very much.

Fol. 6v, l 22

- (23) Context: E En Ramon de Cardona e·ls richs hòmens enviaren-nos En Guillem de Castellaulí e En Guillem de Rajadell e dixeren-nos que volien parlar ab nós denant nostra cort;

And Sir Ramon of Cardona and the rich men sent us Sir Guillem of Castellaulí and Sir Guillem of Rajadell, and they told us that they wanted to talk to us in front of our court;

E [a nós]_i plach -nos_i molt (...)
and to us pleased.3SG =us.CL greatly

And it pleased us greatly (...)

Fol. 195r, l 12

In both examples, predicates are of emotional experience: *pesar* ‘to sadden’, and *ploure* ‘to please’, and the CILDed PP encodes the experiencer, and is thus argumental. In (22), the PP makes reference to *el comte Don Sanxo*, ‘count Sir Sanxo’, a character that had already been introduced in the common ground when mentioned as the leader of a group of knights in charge of getting King James I and the count of Provence out from the castle of Montsó. Therefore, in this case, CILD is performing topic shift, but at the same time, it has a contrastive flavour, because from all the knights who could have been upset by the departure of the count of Provence, it is him who is. In (23), on the other hand, the CILDed constituent carries out topic shift from the preceding clause and it has no contrastive value.

No more examples of CILD are found in main clauses in the LFRJ database. The scarcity of data does not allow us to establish whether there is an actual correlation, or it is merely chance that the two cases that we find occur with psychological verbs that select a theme subject and an experiencer introduced by a PP. In addition, since we only count with two examples of clauses with CILDed elements, we cannot establish the relative order between constituents in the left periphery of main clauses with respect to them.

CILD is also found in embedded clauses, as illustrated by examples (24-26). In all cases, the CILDed constituent is the DO of the clause, contrasting with the examples found in main clauses, all argumental PPs.

- (24) Context: “A, séyer, més havets perdut que no us cuydats, que En Guillem de Muntcada e·N Ramon són morts”. “Con morts?” --dixem nós. E presem-nos a plorar. E puys dixem al bisbe: “No plorem, que ara no és ora de plorar, mas levem-los del camp, pus que morts són”. E estiguem aquí una peça, ploram e puys sobre·N Ramon atre tal. (...) E, quan vench al matí, que fo asseguda la albergada, ajustaren-se los bisbes e els nobles e vengren a la nostra tenda. E dix lo bisbe de Barcelona, En Berenguer de Palou per nom:

‘Ah, Sir, you have lost more than you think, since Sir Guillem of Muntcada and Sir Ramon are dead’: ‘What do you mean, dead?’, we said. And we started crying. (...) And we remained here for a while, we cried, and again on Sir Ramon. And when the morning came, and the camp was set, the bishops and noblemen gathered and came to our tent. And the bishop of Barcelona, Sir Berenguer of Palou, said:

“Seyor, mester **seria** que [aquests corsos qui són
 Sir need be.3SG.COND that these bodies who are.3PL
morts]; que·ls; soterràs hom.”
 dead that;them.CL= buried.3SG.SBJV one

‘Sir, it would be necessary that the dead bodies would be buried.’

Fol. 40r, l 19

- (25) Context: E dix En Guillem de Cardona: “Jo no hic vinch per àls ni per pladejar, mas per dir **ço que·m castigaren**”. E dix En Guillem de Muntcada: “¿Havets procuració neguna vós d’En Guerau?” e dix él:

*And Guillem de Cardona said: “I have not come here to argue or anything else, but to say **what they punished me for.**” And Guillem de Muntcada said: “Have you got anything that proves you can act on Guerau’s behalf?”. And he said:*

“No, mas [ço de què ens castigaren];, jo us ho;
 no but this of which us.CL= punished.3PL I to;you.CL= it.CL=
 e dit (...)
 have.1SG said

‘No, but the reason why we got punished, I have already told you (...)’

Fol. 21v, l 14

- (26) Context: The king takes part in a conflict for the possession of a castle. An attack against the castle takes place.

E	de	la	vila	enviaren-nos	missatge que féssem,
and from	the	village	sent.3PL=to;us.CL	message that	do.1SG.PST.SBJV
que [tot];	ho;	haviem	goayat,	sol	que·l
that all	it.CL=	had.1PL	won	only	that;the
comte isqués		de	fora.		
count	go.out.3SG.PST.SBJV	of	out		

And from the village they sent us a message that we should go on, because we had won everything, but for getting the count to get out.

Fol. 25r, l 25

In (24), the CILDed constituent is the DO [+masculine, +plural] DO *aquests corsos qui són morts* ‘these bodies that are dead’, echoed in the core of the clause by the object clitic pronoun *els* ‘they’, [+ masculine, + plural] and accusative. While the ‘corpses’ had not been overtly mentioned in the preceding discourse, the death of two knights had activated the notion of ‘corpses’ in the interlocutors’ minds. The presence of the demonstrative *aquests* ‘these’ establishes a contrast between those corpses and others: those of the enemy. Therefore, the CILDed topic can be argued to have a contrastive reading. In (25), again, it is the DO of the clause that is CILDed, *çò* ‘this’ and its dependant relative clause. Since the pronoun *çò* ‘this’ is one of the few remnants of the neuter gender in OCat, is linked to the neuter clitic *ho* ‘this’, occurring in the core of the clause. Finally, in (26), the neuter indefinite pronoun *tot* ‘all, everything’ appears in the left periphery, and it is coreferent with the neuter clitic pronoun *ho* ‘this’. This example is problematic, since indefinite pronouns are bound to receive focus reading. However, the presence of the clitic suggests that this is not the case. The pronoun *tot* ‘everything’ makes reference to the various fronts that were open in the attack, and it is contrasted with the following adverbial clause, where the one thing that is yet to be done is mentioned. Thus, (26) can indeed be considered a case of CILD and *tot* a shifting topic with a contrastive flavour.

From this section, therefore, it can be concluded CILD and HTLD were two topicalisation strategies available to the OCat of the LFRJ database. In the case of the latter, unlike it has been described for other Old Romance varieties (Benincà 2004, 2006) but like in

MCat and MSp, it presents a clitic copy in the core of the clause. The informational reading of CILDED topics varies: while in all cases they carry out topic shift, in a few this is combined with contrastiveness.

5.2.4 Foci

Different labels are used for different types of foci in the literature, but here we follow the recent literature in Old Romance word order and IS (Cruschina, 2012; Steiner, 2015; Wolfe, 2015) in identifying three types of foci,⁵⁶ all with different syntactic reflexes: informational foci (IFoc), which convey new information, contrastive foci (CFoc), which convey contrast between a prior element and a new one, and finally, focus fronting (FF) or *verum* focus in (Leonetti & Escandell-Vidal 2009, Sitaridou 2011, Poole 2013), by which an element that can carry old information, but which is read emphatically is fronted to the left periphery, triggering the focalisation of the polarity of the clause. In MCat, MSp and MIIt, CFoc and FF occur in the left periphery, preverbally, while IFoc occurs postverbally. Cruschina (2012), with MSic data, argues that FF and CFoc occupy two different projections in the left periphery, since in the case of FF, verb-adjacency is compulsory, while in the case of CFoc adjacency between the focalised constituent and the verb is not required. This is illustrated in (27) and (28):

- (27) A: Chi cci ricisti a tò niputi?
 what to;him.CL= said.2SG to your nephew
- B: A virità (*a mè niputi) cci rissi.
 the truth to my nephews to;them.CL= said.1SG
- A: What did you say to your nephews?
 B: I told them the truth.

MSic

Cruschina (2012, p. 107, example 41)

- (28) Na Littra, a Pina, cci scrissi (no un pizzinu)
 a letter to Pina her.CL=wrote.1SG not a note

⁵⁶ Cruschina (2012, 2015) puts forward the notion of Mirative Focus, a type of focus that receives a surprise or unexpectedness interpretation, which syntactically and prosodically differs from FFing. The exploration of this structure in regard to Modern or OCat is left for future research.

I wrote a letter to Pina, not a note.

MSic

Cruschina (2012, p. 107, example 42)

Cruschina uses examples like (27) and (28) to argue that FF and CFoc occupy different positions in the left periphery, whereby CFoc is higher in the structure, since it allows for topics to occur to its right, as in (28). In contrast, in (27), a case of FF, verb-adjacency is required. Thus, Cruschina sides with Rizzi (1997) in advocating for topic recursion, contra Benincà & Poletto (2006):

(29) ForceP TopP* CFocP{CFoc} TopP* IFocP{Wh/IFoc} FinP

Adapted from Cruschina (2012, p. example 37)

The same has been claimed for M_{Cat} by Vallduví (1995). Example (30) illustrates how a constituent can intervene between the contrastively focused constituent and the verb:

(30) El mercedes el Jordi no sap qui s'ha comprat.
the Mercedes the Jordi not knows who REFL.CL=has bought

Jordi doesn't know who bought the Mercedes.

Vallduví (1995, p. 131-132)

However, as noted by Quer (2002, ft. 3) many Catalan speakers judge sentences like (31) as highly ungrammatical. In line with Quer, all Central Catalan speakers consulted deemed (32) as highly ungrammatical. Therefore, for the purposes of this work, we take M_{Cat} CFoc to require verb adjacency, as it is the case in M_{Sp} (Zubizarreta, 1998):⁵⁷

(31) ^{??}El Mercedes, no sap en Jordi qui s' ha
the Mercedes not knows the Jordi who REFL.CL= has

comprat.

bought

Jordi does not know who has bought the Mercedes.

M_{Cat}

⁵⁷ See Calabrese (1992) for a prosodic analysis of the adjacency between the verb and the fronted constituent.

(32) a. * Una carta a les nebodes he escrit, no
 a letter to the nieces have.1SG written not
 una nota.
 a note

b. Una carta he escrit a les meves nebodes, no una nota.
 a letter have.1SG written to the my nieces not a note

It is a letter that I have sent to my nieces, not a postcard.

MCat

Verb adjacency is also a feature of MCat FF:

(33) MOLT m' agraden les nespres.
 a lot to;me.CL= please the loquats

I like loquats a lot.

MCat

(34) AIXÒ dic jo!
 this say.1SG I

It is precisely this that I say!

MCat

(33) and (34) are two examples of FF in MCat. In (33), the quantifying adverb *molt* 'a lot', is moved to the left periphery to receive an emphatic reading. In (34), the direct object *això* 'this' occurs to the left of the verb, with verb-subject inversion. In both cases, the fronted element receives an emphatic reading, and according to Leonetti & Escandell-Vidal (2009), their fronting results in the focalization of the polarity of the clause.⁵⁸ Therefore, we have established that in MCat, FF and CFoc are both realized in the left periphery and involve verb adjacency, aligning with MSp and contrasting with MIt and MSic, which according to Cruschina (2012) have two separate projections in the left periphery for CFoc and FF respectively.

⁵⁸ The discussion of verb and subject position in MCat falls outside the scope of this work. However, it is worth noting that FFing yields strict subject-verb inversion, suggesting verb movement to the left periphery.

As mentioned above, information focus (IFoc) in M_{Cat}, is located to the right of the verb (Leonetti, 2010; Leonetti & Escandell-Vidal, 2009a; Zubizarreta, 1998):

- (35) A: Què dinaràs avui?
 what lunch.2SG.FUT today
- B: Avui dinaré [sopa de verdura].
 today lunch.1SG.FUT soup of vegetables
- A: What will you have for lunch?
- B: Today I will have vegetable soup.

M_{Cat}

- (36) A: On anireu de vacances?
 where go.2PL.FUT of holidays
- B: Anirem [a Sicília].
 go.1PL.FUT to Sicily
- A: Where will you go on holiday?
- B: We will go to Sicily.

M_{Cat}

In (35) and (36), we have two cases of narrow information foci in M_{Cat}. In both cases, the constituent that carries new information is located to the right of the verb, coinciding with clause's highest pitch.

Having sketched the distribution of different types of foci in M_{Cat}, we are now going to examine the distribution of this foci types in the LFRJ database. Firstly, we consider FF, followed by a discussion of CFoc, and finally, we address the distribution of IFoc.

Focus Fronting

Focus fronting is productive in O_{Cat}. Several types of constituents can be moved to the left periphery, without a clitic copy in the core of the clause. They include minimal and maximal categories: adverbs and adverbial phrases, adjectives and adjectival phrases, pronouns, determiner phrases acting as direct objects and subjects, and non-finite verbal forms (which are explored in section 5.4.1). The following examples display cases of FF:

- (37) MOLT som pagats d'aquest pensament que nós
 a lot are.1PL payed of;this thought that we

veem fer a vosaltres.
 see.1PL do.INF to you

We are very pleased of what we see you are thinking.

Fol. 27v, 1 2

(38) VERA COSA **és** e certa que nostre avi (...)
 true thing is and certain that our grandfather

It is a TRUE THING and certain that our grandfather (...).

Fol. 2r, 1 3

(39) “Pare sant, **NÓS** nos en **volem** anar, (...)”.
 father holy we ourselves.CL= ADV.CL= want.1PL go.INF

“Holy father, WE want to leave (...)”.

Fol. 194v, 1 6

(40) On **NÓS** vos **deim**, sobre los ·iii· conseyls que
 where we to;you.CL= say.1PL about the 3 advice that
 vós nos havets demanats, que (...)
 you to;us.CL= have.2PL asked-PART that (...)

Where WE tell you, about the three pieces of advice that you have asked from us, that (...).

Fol. 29v, 1 12

(41) ‘Açò us **diré** yo: (...)’

this to;you.CL= say.1SG.FUT I

‘I will tell you THIS: (...)’.

Fol. 29r, 1 8

In example (37), the fronted element receiving an emphatic reading is the adverb *molt* ‘a lot’. This adverb modifies the predicative past participle *pagats* ‘pleased’, and it has undergone movement from the subject complement phrase to the left periphery. In (38), the fronted element is part of the coordinated subject complement *vera cosa e certa* ‘true and certain thing’. Here, the movement is made apparent by the fact that *certa* ‘certain’ remains in situ, in vP. In (39) and (40), it is the pronoun *nós* ‘we’ that is focus fronted. We have argued

in Chapter 4 that pronouns can occupy several positions in the left periphery. In (40) and (41) and using proclisis as a diagnosis for the focalisation of the preverbal constituent, we can safely locate the subject pronoun in the Focus field. In (40), it could be argued that the subject pronoun receives a contrastive reading, being contrasted against the subject of the object subordinate clause *vós* ‘you’. Finally, in (41), it is the DO *açò* ‘this’, that is fronted. While it is a pronoun, it cataphorically refers to what the speaker is going to say. This strategy is used frequently throughout *El Llibre dels Fets* in reported speech fragments, when the speaker wants to draw attention to what they are going to say.

In all these cases, the fronted constituent could be argued to receive a contrastive reading: in (37), *molt* could be fronted to specify that the speaker is very pleased, instead of just a little; in (38), that what is going to be said is certainly true; in (39) that the person who wants to leave is unambiguously the king, as opposed to other knights; in (40) again, the subject of the clause is emphasised to clarify that it is the king himself, and no one else, who utters the words that follow, and finally, in (41), *açò* ‘this’ is necessarily referring to what the speaker will say, and nothing else. This perceived contrastiveness can be connected to the fact that FF yields focalisation of the assertive force of the proposition, cancelling out alternative readings (Leonetti & Escandell-Vidal 2009, p. 189). The fact that this construction rejects a previously accessible reading of a proposition explains why it occurs with highly referential elements, such as personal and indefinite pronouns, and DPs containing demonstratives, and it always has an echoic value.

There are also a few instances of FF that co-occur with the negation adverb *no*:

- (42) Context: *E nós parlam ab En Guillem de Muntcada e dixem-li que aquest pleyt nos volien fer, ço és assaber (...)*

And we spoke with Sir Guillem of Muntcada and told him that they were pleading us the following, that is (...)

‘*AÇÒ no **tench** yo com cosa que vós dejats fer.*’
 this not have.1SG I as thing that you should.2PL do.INF

I do not consider THIS something you should do.

Fol. 25v, l 3

The FFed constituents in (42), like in the positive clauses above, I singled out by being fronted, cancelling any other reading of the proposition. The neuter pronoun *açò* ‘this’, in this

case anaphoric, is FFed, to make clear that it is precisely the plead that is being made that the king should not follow. In Chapter 3, section 3.1.6, we have discussed that emphatic polarity is encoded by means of the interaction of two polarity related heads in the clause: relative polarity features are features that encode the agreement or denial of the polarity of a clause in relation to the previous utterance: [reverse] and [same] in Farkas & Bruce (2010) analysis; and ΣP , hosts absolute polarity features, [+] and [-], corresponding to [+affirmation] and [+negation]. Thus, in (42) the assertive nature of the utterance is highlighted, regardless of its polarity. Since it is the clause's relative polarity that is being highlighted, Leonetti & Escandell-Vidal (2009b, p. 14) propose that no constituent within the clause receives emphatic reading. In M_{Cat}, FFing is restricted to quantified constituents and requires right dislocation of postverbal constituents that frame the utterance (Batllori & Hernanz, 2011). These constraints are not present in FFing in M_{Sp}.

Leonetti & Escandell-Vidal (2009) describe different contexts in which FF is elicited in M_{Sp}: (i) yes/no questions, (ii) occasions in which the propositional content is presented as an intention, duty or belief, (iii) propositional objects depending on verbs of possibility, intention or belief, and (iv) propositions with a contrastive reading that reject the reading of an implicit contextual assumption, as it is the case in (43):

- (43) A: Ahir van posar una multa a en Joan. T' ho
 yesterday go.3PL put.INF a fine to the Joan REFL.2SG=it.CL
 pots creure? Si és un conductor
 can.2SG believe.INF indeed is a driver
 impecable! Què coi?
 impeccable what gosh
- B: Alguna cosa devia fer en Joan, si li
 some thing must.3SG do.INF the Joan if to;him.CL
 van posar una multa.
 go.3PL put.INF a fine
- A: Yesterday Joan got a fine. Can you believe it? He is such a flawless driver!
 What on earth!
- B: 'Joan must have done something, if he got a fine'

M_{Cat}

In (43), A's intervention contains the implicit assumption that it is highly unlikely for Joan to have done something to get a fine. B's answer refutes the assumption, and via FF it highlights that indeed, he must have done something.

Contrastive contexts such as the one sketched in (43) can also explain the cases of FF in the LFRJ database. However, they do not always reverse the polarity of a previous assumption: they can also cancel out possible future assumptions. This is illustrated in (44), which contains example (38) described above:

- (44) Vera cosa és e certa que nostre avi, el rey Don Amfós, féu parlar matrimoni a l'emperador de Contastinoble que li donàs sa filla per muyler. E sobre aqueles paraules que foren tractades e acordades d'amdues les parts, ço és assaber, de nostre avi e de l'emperador, féu matrimoni nostre avi ab la regina Dona Sanxa, que fo fiyla de l'emperador de Castella.

It is a true and certain thing that our grandfather, King Sir Amfós, convinced the emperor of Constantinople so that he would give him his daughter as wife. And on these words that were agreed and discussed from both sides, that is, our grandfather and the emperor, my grandfather married queen Lady Sanxa, who was daughter of the emperor of Castile.

Fol. 2r, l 3

In (44), FF is used to refute the anticipated hearer's reaction to the King's grandfather having been betrothed to a Turkish princess, when it is common knowledge that he married Sanxa of Castile. Similarly, in (45) FF is linked to the refutation of a contextually implicit assumption:

- (45) Context: E nós parlam ab En Guillem de Muntcada e dixem-li que aquest pleyt nos volien fer, ço és assaber (...)

And we spoke with Sir Guillem of Muntcada and told him that they were pleading us the following, that is (...)

'AÇÒ no **tench** yo com cosa que vós dejats fer.'
 this not have.1SG I as thing that you should.2PL do.INF

I do not consider THIS something you should do.

Fol. 25v, l 3

In (45), in contrast with (44), the assumption that is being refuted precedes the FF sentence, as it was the case in (43), the M_{Cat} example of contrastive FF provided above. No other contexts eliciting FF can be identified within the LFRJ database.

FF is tightly linked with the grammaticalisation process of positive polarity particles. In Chapter 3, section 3.1.5, the distribution of *sí* in O_{Cat} is analysed and it is shown that it was already a positive polarity marker in the 13th century, whose reanalysis was tightly linked to it receiving FF in (non)veridical contexts.

In M_{Cat} (see example (43)) and M_{Sp}, FF involves verb-subject adjacency when the subject is overt. This is also the case for the examples (37), (38) and (41) above, when the FF element is not a subject pronoun. The sequence XP V S with strict verb subject adjacency suggests that in these contexts there is verb movement to the left periphery, as suggested by Batllori & Hernanz (2011). Holmberg (2016, §3.2) describes that verb movement is both a feature of questions and answers to alternative questions and yes/no questions (questions that present a disjunctive set of alternatives as answers) in some languages, among which O_{Sp} (Rodríguez Molina, 2014), and M_{Pt} (Martins, 1994, 2004, 2005, 2013), as we have discussed in Chapter 3, section 3.1.6. Data from the LFRJ database suggests that that can also be the case in O_{Cat} in polarity reversal contexts (Holmberg, 2003; Farkas & Bruce, 2009).

Answers that reject the negative alternative of a question crosslinguistically present particular syntactic features. Two examples of this would be M_{Fr}, which has the emphatic positive polarity particle *sí*, which emerges in polarity reversal contexts, or M_{Cat}, in which it is necessary to answer with a full sentence introduced by the cluster *sí que* (for a discussion of *sí* in O_{Cat}, see section 3.1.6 in Chapter 3). We propose that FF in O_{Cat} was one of the mechanisms available to carry out polarity reversal from an implicit negative question, and not just assumption (Leonetti & Escandell-Vidal, 2009a). The implied question asks for confirmation of what the speaker thought to be true, but that is now doubting. This can be paraphrased as: ‘I thought that *x* was the case, but now I am not sure, please confirm that I am right’. Thus, in the LFRJ database, FF anticipates a negative question from the interlocutor and provides a polarity reversal answer.

Formally, FF involves (i) V-to-C, associated with polarity reversal answers, and (ii) the move of a constituent to SpecPolP and the verb to PolP, and subsequently to SpecFocP and FocP respectively, so that it can assign a truth value to PolP and indicate which of the

adverbial embedded clauses. Table 2 shows the frequency in which different constituent types (direct objects, indirect objects, prepositional phrases and adverbial phrases) are fronted in main and embedded clauses:

	Main Clauses	Embedded Clauses
Direct Objects	21	47
Indirect Objects	7	2
Prepositional Phrases	35	47
Adverbial Phrases (XP)	54	38
Total	117 (11,7%/total, 22.7%/V2, V3* clauses)	134 (13,4%/total, 25%/V2, V3* clauses)

Table 2 – Fronted constituents in main and embedded clauses from the LFRJ database

The main difference in the distribution of fronted constituents in main and embedded clauses is the frequency of fronted DO, which is twice as frequent in embedded clauses than in main clauses.

FF in embedded clauses is found in complement clauses of *verba dicendi* and factive verbs, as well as in adverbial clauses.

- (50) Context: E per lo mal ordonament e per lo peccat que era en ells, hac-se a vençre la batayla (...). E aquí morí nostre pare.

And for the bad organisation and for the sin that was in them, they lost the battle. (...)

And here died our father.

Car	AIXÍ	ho	ha	usat	nostre	lynatge	totz	temps,	que
since	thus	it.CL=	has	used	our	lineage	always	that	
en les	batayles	qu'els	àn	feytes	ne	nós	farem,	de	
in the	battles	that;they	have	done	or	we	do.1PL.FUT	of	
vençre		o	morir.						
win.INF		or	die.INF						

Since our lineage has always acted thus: in the battles that they have done and in those that we will do, [our lineage has always] conquered or died.

Fol. 5r, l 27

In (50), the adverb *així* ‘thus’ is FFed to reject the implicit negative question arising from the context (‘Your lineage has not always acted thus, has it?’), like in the FFing examples of main clauses. In this case, the subordinate clause containing the adverb is an adverbial causal clause.

(51) Context: E En Guillem de Muntcada respòs per los nobles e per él e dix

And Sir Guillem of Muntcada answered on behalf of the noblemen and himself

que	molt	graïa	a	Nostre Seyor	lo	bon	propòsit	que
that	much	thanked.3SG	to	our Lord	the	good	purpose	that
él	nos	havia	dat.					
he	to;us.CL=	had	given					

that he greatly thanked our Lord for the good purpose that he had given to us.

Fol. 28v, l 19

In (51), the quantitative adverb *molt* ‘much’ is fronted from its base generated in VP. As in the main clause cases of quantitative adverb fronting, FF reverses the polarity of an implicit negative question regarding whether or not Guillem de Muntcada thanked God for the new purpose that him and the noblemen were given. In this case, the embedded clause containing FFing is a completive clause, direct object of the *verbum dicendi* *dix* ‘he said’.

(52) Context: E açò féu aquel àngel que Déus nos envià; e, quan dich àngel, él era sarraý,

And this is what that angel that God sent us did, and when I say angel, he was a Saracen,

mas	tant	nos	tench	bon	loch,	que	PER	ÀNGEL	lo	prenguem
but	so	to;us	had.3SG	good	place	that	for	angel	him.CL=	took.1SG
e	per	açò	li	faem	con	semblança	d’àngel.			
And	for	this	to;him.CL=	did.1PL	as	resemblance	of;angel			

but he was so considerate to us, that we took him for an angel.

Fol. 42, l 7

And we found Aragon and Catalonia disturbed, the ones fought against the others and did not agree on anything, what one side wanted, the other did not, and you had bad reputation throughout the world because of what had happened.

AQUEST MAL nós no **podem** adobar sinó per dues maneres:
 this evil we not can.1PL mend.INF but for two manners

And this evil, we cannot mend but in two ways: (...)

Fol. 28r, l 25

(55) Context: E nós, quant aguem hoïda la lur paraula e lur conseyl, dixem-los que

And once we had heard their word and their advice, we told them that

ESTA COSA nós no **faríem**.
 this thing we not do.1PL.COND

this is not something that we will do.

Fol. 32v, l 26

(54) and (55), in spite of being a main and an embedded clause, are structurally parallel: both clauses contain the negative particle *no* in ΣP ; both clauses present a fronted DO in the left periphery, followed by the subject pronoun *nós* ‘we’; in both cases, the fronted DO contains an anaphoric demonstrative: *aquest* ‘this’ in (54), referring back to the situation in which Catalonia and Aragon were, and *aquesta* ‘this’ in (55), which refers back to *la lur paraula* ‘their advice’.

In these cases, FF responds to a question with a positive bias: ‘and you could address this issue in many ways, couldn’t you?’ for (54) and ‘you will follow their advice, won’t you?’ for (55).

We propose that in instances where FF co-occurs with negation, the FFed constituent moves to the same projection in the left periphery as in main clauses, but without passing through PolP, since polarity reversal from $[+]>[-]$ is carried out by overt negative particle *no* and not through verb movement. This accounts for the lack of subject verb inversion, which is otherwise found in cases of FF with negative polarity reversal. The lack of verb movement allows for the presence of material between the FFed constituent and the verb: in (54) and (55) the pronoun *nós* ‘we’. In both cases, the subject of the clause with FF changes with respect to the subject of the preceding clause, therefore, it performs topic shift in SpecTP.

And there was Ferran Sànxec de Castre and Sir Artal de Luna and Sir Pero Corneyl and many others. And because of this they did not come into Lleida, since they were afraid, and we sent someone to tell them that we would ward them from all men, and they still did not want to enter there, but they sent here Sir Guillem of Castellaulí and Sir Guillem of Rajadell, representatives of theirs. And here they asked us to provide them with an interlocutor and

nós **donam**-los [En Ramon de Vals, canonge de Leyda, e
 we gave.1PL=them.CL Sir Ramon of Vals canon of Lleida and
En Ramon Gili].
 Sir Ramon Gili

(...) we gave them Sir Ramon Vals, priest of Lleida, and Sir Ramon Gili.

Fol. 196v, 11

- (58) Context: “ (...) E a nós no avendrà així, car nós no vim anch apostoli si vós no, per què volem penitència pendre de vós”. E ell fou-ne alegre molt e pagat

‘And this shall not happen to us in this way, since we have never known a pope but you, because we want to take forgiveness from you’. And he was very happy and satisfied

e dix [que la·ns daria].
 and said.3SG that it.CL;to;us.CL=give.3SG

and he said that he would give it [forgiveness] to us.

Fol. 194v, 15

- (59) Context: E ab aytant manam nostra ost a l’estiu aenant e anam sobre Albarrezí.

And in the meanwhile, we lead our army from the summer onwards and we marched on Albarrasí.

E **mesem** [nostre seti] sobre la torra de l’Andador,
 and set.1PL our seige on the tower of the;Andador
 en una serreta que·s fa sobre aquela torre.
 in a little mountain range that=itself.CL= does above that tower

- (62) Context: ço és, per volentat de Déu que·ns endreç en nostres affers, e que començem tals coses, e a vós e a nós, que a él vinga de plaer e que la cosa sia tan gran e tan bona que la mala fama que és entre vós que·s tolga,

This is, for the will of God who will tidy up our affairs, and that we start such things that agree with you and us and with him, and that this thing will be so good that the bad reputation that is among you, that it should be removed,

car la claror de les bones obres desfà [l'escuredat].
 since the light of the good deeds undoes the;darkness
 since the light of good deeds destroys darkness.

Fol. 28v, 13

(61) and (62) show how new information is located systematically postverbally like in main clauses. Nevertheless, as it was shown in Table 2, embedded clauses present twice as many preverbal direct objects, 47 in total, than main clauses, 21 in total. Most cases can be explained through FF. However, there are a few cases where the fronted constituent provides new information, and does not fall into the several categories identified as prone to FF:

- (63) “Pregam-vos que no us ho tolgats, que
 beg.1PL=to;you.CL that not to;you it.CL= leave.2PL because
 honta hi prendriem”.
 shame in;it.CL= take.1PL.COND

‘We beg you not to give this up, since it would shame us’.

Fol. 121v, 128

- (64) (...) e dixeran-nos que farien encendre una lanterna, mas havien paor que·ls veessen les guardes de Maylorques.

(...) and they told us that they would light up a lantern, but that they were worried that the wards from Mallorca would see them.

E nós dixem-los que conseyl hi podien
 and we said.1PL=them.CL that advice there.CL= could.3PL
 pendre: (...)

take.INF

And we said that they could take some advice:

Fol. 33v, 15

In both (63) and (64), the DO appears preverbally and has not been previously mentioned in the discourse. Nevertheless, in both cases the predicate and the object cooccur and convey a specific idiomatic meaning. The verb *prendre* ‘to take’ often is often used as a light verb, that is, a verb that occurs in a complex predicate where the main semantic content is provided by its direct object. This is the case in (63), with *prendre honta* ‘to be ashamed by’ and in (64) *prendre conseyl* ‘to take advice’.⁵⁹ The relationship between light verbs and their objects is tighter than that of regular transitive verbs and their objects or even that of transitive verbs and their prototypical objects (Forcadell 2013, p. 59). Therefore, while these DO had not been previously mentioned in the discourse, they are predictable within the context. Their fronting cancels out other possible interpretations of the light verb, discarding the possibility of these direct objects receiving narrow focus in the left periphery.

Thus, information focus in OCat main and embedded clauses presents the same distribution as in MCat, where Information Focus is invariably encoded postverbally (also with unaccusative predicates, when the subject introduces new information into the discourse, as shown in Chapter 4, section 4.2).

5.3 Interim summary

Up to this point, it can be established that the broad structure of OCat left periphery is parallel to that sketched by Benincà (2004, 2006) for Old Romance: it has a Frame field, followed by a Topic field, which is in turned followed by a Focus field. We have not been able to establish whether different types of scene setters, topics and foci display a relative order within their fields given the lack of data. Nevertheless, we have been able to show that main and embedded clauses have the same positions available in the left periphery, strengthening the idea already defended in Chapter 4 that there is no significant asymmetry in their syntax.

We have also been able to show that there is a significant difference between the distribution of IFoc in OCat and in the Old Romance languages. Wolfe (2015a) finds that Old

⁵⁹ Other collocations with the verb *prendre* ‘to take’ include *prendre la paraula* ‘to take the floor’, *prendre mida* ‘to measure’, *prendre consciència* ‘to become aware of something’, among many others, as *prendre* + DO was a very productive pattern.

Romance Fin V2 varieties (those where the V2 constraint is located in FinP, mainly OOc, OSic and OSard), have IFoc located in the left periphery, in line with Benincà (2004, 2006) description of the Old Romance left periphery. Sitaridou (2011, 2015, 2019), in her non-V2 descriptions of the OSp left periphery, also identifies an IFoc projection in the left periphery. In OCat, or at least in the data of the LFRJ database, IFoc was unequivocally encoded postverbally. Therefore, OCat sets itself aside from the other Old Romance languages once more, aligning itself with the modern varieties. It is worth noting that FF, linked to polarity reversal, and therefore, the notions of contrast (Rodríguez Molina, 2014) and emphatic polarity, is indeed encoded in the left periphery, like in MCat.

The connection between FF and polarity has also been explored at length, showing that all FFing cases can be explained by means of polarity reversal, and that this mechanism is intimately linked with the grammaticalisation of *sí* as an EPPA. The fact that FFing is linked to polarity reversal adds one more argument to the rejection of the existence of Germanic-style inversion in OCat, since all cases of fronted DOs are FFed, and not the result of unmarked verb movement to the left periphery.

5.4 Information Structure Related Phenomena

5.4.1 Anteposition of Non-Finite Forms

The fronting of a non-finite form from a verbal periphrasis, which has been referred to as long head movement (LHM), stylistic fronting (SF) or anteposition of a non-finite form,⁶⁰ is a well studied phenomenon in Scandinavian and Germanic languages (Diesing, 1990; Holmberg, 2006; Rögnvaldsson & Thráinsson, 1990; Santorini, 1989; Vikner, 1995), as well as in the Old Romance languages (Alboiu, Hill, & Sitaridou, 2015; Batllori, 2015; Fischer, 2005; Fontana, 1993; Ribeiro, 1995; Rivero, 1993; Sitaridou, 2015).

Analyses of this phenomenon fall into two sides: some authors assume that the possibility of fronting non-finite forms derives from the Medieval Romance having V2 syntax (Fontana, 1993; Ribeiro, 1995; Rivero 1993; Franco, 2009; Poletto, 2014), while others have

⁶⁰ These three labels are linked to different analyses of the fronting of non-finite forms. Stylistic Fronting is used to refer to the movement of maximal categories, including non-finite verb forms, to the left periphery for IS motivations. LHM, on the other hand, its used to refer to movement of the non-finite form to the left periphery triggered by formal requirements (Fontana, 1993; Rivero, 1993); or as a label to describe IS-triggered movement of the non-finite form, taken to be a head and not a phrase (Alboiu et al., 2015; Sitaridou, 2015) Finally, anteposition of non-finite forms is used neutrally as a descriptive label of the phenomenon (Batllori, 2015). It is for the neutrality attached to the latter label that we choose to use it here.

linked the availability of such structures to IS and the make-up of the left periphery. In this section, I explore the anteposition of non-finite forms (henceforth ANF) in *El Llibre dels Fets*, using a database that was created for my MPhil thesis (Pujol i Campeny, 2014). Firstly, I describe the distribution of this phenomenon. This is followed by an assessment of this phenomenon against analyses that link it to a V2 grammar and those that link it to IS. It is finally concluded that in OCat, ANF was another instantiation of FFing.

Before starting, it is necessary to distinguish ANF from mesoclitic constructions. Mesoclitic constructions consist of conditional or future tense verbs ($V_{\text{NON FINITE (INF) AUX}_{\text{FUT/COND}}}$) where a clitic pronoun intervenes between the infinitive and the auxiliary. I refer the reader to Batllori (2016) de Toledo y Huerta (2015) and references therein. Mesoclisism is illustrated in (65):

(65) ‘e la torra, quan serà canada, falir-li
 and the tower when be.3SG.FUT scaffolded not need.INF=to;it.CL=
 àn los estolons’.
 have.3PL the supports

‘and the tower, once it will be scaffolded, it will not need the supports any longer’.

Fol. 43r, l 17

Table 3 shows the distribution of ANF in main and embedded clauses, displaying whether they also present pro-drop:

Clause type		Non pro-drop	Pro-drop	Total
Main clauses		4	2	6 (6.7%)
Embedded clauses	Adverbial	4	14	18 (19.8%)
	Conditional	1	3	4 (4.4%)
	Complement	3	11	14 (15.3%)
	Relative	11	38	49 (53.8%)
Total		23 (25.3%)	68 (74.7%)	91 (100%)

Table 3 – Main and embedded clauses with ANF

Table 3 shows that ANF presents certain distributional tendencies: it overwhelmingly tends to occur in embedded clauses (93.3% of cases, vs. 6.7% occurring in main clauses), and within embedded clauses, relative clauses are the most frequent environment in which ANF occurs. Another clear tendency is for ANF clauses to display pro-drop.

ANF occurs with tense auxiliaries (*ser* ‘to be’ and *haver* ‘to have’) and with modal verbs (*poder* ‘to be able to’, *deure* ‘ought’ and *voler* ‘want’). The distribution of these predicates does not show a significant pattern between main and embedded clauses. Therefore, differences between tense auxiliaries and modal verbs will not be considered here and will be left for future research.

In early literature on the anteposition of non-finite forms in the Romance languages, it was suggested that anteposition was triggered to provide a phonological host for clitic pronouns. Nevertheless, as has been discussed in Chapter 3, section 3.1.5, OCat clitic placement did not determine whether a constituent occurred pre or postverbally. Rather, the position of clitics was determined by whether FocP was saturated or not (Batllori, et al., 2005). This is further confirmed by the multiple cases of non-finite form anteposition without clitics occurring in the clause:

- (66) E calaren e hagren vergonya de ço que dit
 and shut.up.3PL and had.3PL shame of that which said
havien.
 had.3PL

And they remained quiet and were ashamed of what they had said.

Fol. 47r, l 24

- (67) E, (...) començam de partir les cases ab l’arquibisbe
 and started.1PL of leave.INF the houses with the;archbishop
 de Narbona e·ls bisbes e·ls nobles qui estat
 of Narbonne and;the bishops and;the noblemen that been
havien ab nós (...)
 had.3PL with us

And we started to leave the village with the archbishop of Narbonne and the bishops and noblemen that had been with us.

while this was not possible in embedded clauses, where no V-to-C movement is required (see Roberts [1993] for more on asymmetric pro-drop and null subject licensing). The high frequency of pro-drop in clauses displaying ANF could thus be linked to the reported need for a subject gap in clauses with this structure (Maling, 1980; Franco, 2009). Whether this is the case or not in OCat is considered in the following paragraphs.

Therefore, it can be established that V2 analyses of ANF in Old Romance fall in two categories: those that propose that ANF is a last resort mechanism to satisfy the V2 parameter by moving the non-finite form to the specifier of the projection to which the verb raises (Fontana, 1993; Ribeiro, 1995; Rivero 1993) and those that describe it as one strategy to allow for subject extraction in contexts where this would otherwise be impossible (Franco, 2009).

We discard the formal satisfaction of this syntactic requirement as a motivation for ANF given that OCat did not display V2 in either main or embedded clauses (see Chapter 3 and Chapter 4 for an SVO analysis of OCat). We also dismiss ANF as a strategy for allowing subject extraction on the basis that OCat was a symmetric pro-drop language in which the subject criterion did not involve the freezing of the subject in SpecTP, given the availability of *pro*. This is further developed in what follows.

As already described in Chapter 4, SpecTP (or a projection therein) is an A-movement landing site for an argument about which the clause's event is predicated, a property that makes this position criterial (linked to IS). In pro-drop languages such as MIIt, this position can be filled with *pro* when a subject binding *pro* is topicalised (Rizzi, 2018, p. 527). In asymmetric V2 languages, such as OIt as described by Franco (2009) and Poletto (2014), the licensing of *pro* in embedded clauses is not readily available, since null subjects are only licensed in these varieties when the verb reaches the left periphery (Roberts, 1993; Wolfe, 2015a). In these contexts, SVO orders, with the verb remaining in TP and the subject in its specifier, are favoured, and no element (apart from certain adverbs and negation) is expected to occur between subject and verb.

The relationship between subject position and ANF was already pointed out by Maling (1980, p. 76), who describes that Modern Icelandic ANF occurs mostly in embedded clauses that necessarily present a subject gap, resulting from either (i) subject extraction, (ii) the presence of an impersonal predicate, or (iii) indefinite NP postposing with *there* insertion.

Franco (2009) also describes the need for a subject gap in ANF clauses in OIt. Nevertheless, this author proposes that ANF is in fact a strategy to extract the subject in contexts where this would have been otherwise impossible, given the unavailability of the generation of *pro* in embedded clauses, which translates into the halting of the subject in SpecTP to satisfy the subject criterion.

In clauses with ANF where the linear word order pattern is S V_{NON-FIN} V_{FIN}, Franco (2009) takes the subject to be topicalised in the left periphery thanks to the non-finite form climbing up to FinP (the lowest head of the left periphery), where it satisfies its featural make-up. Since FinP locally c-commands SpecTP, it can check its features (Franco, 2009, p. 107), allowing for the subject to avoid criterial freezing in this position and to reach another position in the left periphery. This explains the adjacency between the non-finite form and the verb in OIt clauses with ANF, given that the former is in FinP while the latter is found in a high projection within TP.

OCat clauses displaying ANF with an overt subject fall into two groups: those with a preverbal subject, and those with a postverbal subject (not unsurprising for a symmetric pro-drop language). Examples (70-74) illustrate both possibilities:

- (70) (...) e dix-nos que mort era [lo rey
 and said3.SG=to;us.CL that dead was.3SG the king
 de Leó].
 of Leon.

And he told us that the king of Leon had died.

Fol. 58r, 127

- (71) “Seyor, ferit és [Don Bernat Guillem] d’una sageta per la cama”
 Sir wounded is Sir Bernat Guillem of;one arrow through the leg
 ‘Sir, Sir Bernat Guillem has been wounded by an arrow in the leg’.

Fol. 83r, 14

- (72) (...) però gens per la missió que [nós e
 but not at all for the mission that we and
 nostres hòmens] feyta **havem**
 our men done have.1PL

(...) but not at all by the mission that we and our men have done.

(73) E no·ns volch creure d'aquel conseyl
 and not=to;us.CL wanted.3SG believe.INF of;that advice

que [nós] donat li **haviem**.
 that we given to;him.CL= had.1PL

And he did not want to abide by the advice that we had given him.

(74) E éls enteneren que nós ho preàvem poch
 and they understood.3PL that we it value.1PL little

ço que [éls] feyt **haviem**.
 that which they done had

And they understood that we valued very little that which they had done.

One of the motivations for Franco's analysis was the constant adjacency between the non-finite form and the finite verb, which is also found in OCat. Following Fischer's (2005) and Batllori's (2015) proposals, we analyse data from *El Llibre dels as* FF, and contribute to the refinement of Batllori & Hernanz (2013) and Batllori (2015) proposal of the existence of two competing grammars within OCat (in the sense of Kroch [1989]): a more archaising one, allowing for FF of elements of several categories (NPs, PPs, AdjPs, AdvPs and non-finite verbal forms) to the left periphery, and a more modern one, where this option had already been lost. Note that elements that can undergo FFing are the same that are susceptible to undergo SF in the accounts of Franco (2009) and Maling (1980). If examples (70-74) above are analysed with this in mind, the possibility to have preverbal and postverbal subjects with ANF is easily derived. In cases where the subject precedes the non-finite form, like (72-74), we agree with Franco (2009) that they are topicalised in the left periphery, and in the case of OCat, coreferent with *pro*. When the subject is postverbal, like in (70) and (71), two options emerge: the subject can either be in SpecTP, as expected in FFed constructions, or, alternatively, it could remain in its base generated position or in a low position within the left periphery. As we have seen in Chapter 4, section 4.3.1, without the presence of other elements in the clause, this is impossible to tell.

Batllori & Hernanz (2012) and Batllori (2015) analyse non-finite form anteposition in terms of FF, which is in line with Fischer (2005) proposal that the construction is linked to IS rather than structural factors.⁶¹ This analysis can account for fronted infinitives and participles, which, by virtue of being fronted, cancel out the negative counterpart of their presupposition, as illustrated with examples (75) and (76):

- (75) ‘Don Nuno, pugem ab aquesta companya, que ara van, Sir Nuno climb.1PL with this company that now go.3PL que vençuts són, que tuyt van brescan, that defeated are.3PL that all go.3PL breaking.order-GER e companya que va brescan en batayla, no ha and company that goes breaking.order-GER in battle not has qui·ls escometa, que, si·ls escometien, bé who=them.CL assault.3SG.SBJV that if=them.CL= assaulted.3PL well vençuts són’
defeated are.3PL

‘Sir Nuno, let’s go with this company who are living now, who are already defeated, since they are all advancing disorderly, and when a company advances disorderly, it means that there is no one to assault them, because, if someone attacked them, they are really defeated’.

Fol. 38v, l 24

In (75), there are two instances of participial anteposition, both with the verb *vèncer* ‘to conquer’. In the first case, participial fronting confirms that the company is indeed advancing disorderly, because they are really *conquered* (the negative counterpart of the proposition is cancelled out), therefore, Batllori & Hernanz analysis is derived straightforwardly. In the second case, the non-finite form is preceded by the adverb *bé* ‘well’. Since ANF is XP movement given that it can co-occur with negation (Batllori, 2015; Franco, 2009, p. 109), the participle must undergo remnant movement of vP, by which the non-finite form and its specifier are moved to the left periphery. In periphrastic configurations, *bé* ‘well’

⁶¹ Franco (2009) states that in Fischer & Alexiadou (2001) SF in Old Romance is linked to the focalisation of the clause’s polarity, in line with Fischer (2005). Unfortunately, I have not been able to access this work.

Narrative body		28 (30.8%)	28 (30.8%)
Reported Speech	Direct Speech	38 (41.7%)	63 (69.2%)
	Indirect Speech	25 (27.5%)	
Total			91 (100%)

Table 4 – Distribution of clauses with ANF in the narrative body and reported speech

Interestingly, a similar distributional pattern is found with FFed DOs containing old information, which also prominently occur in reported speech (see section 5.2.4 in this Chapter), while other types of fronted constituents are more prominent in the narrative body.

While FFing could target a variety of constituents in OCat, including non-finite verbal forms, in MCat it is restricted to quantified constituents. The distributional pattern of FFed elements in the LFRJ database might be indicative of beginning of the end of FFing in Catalan, starting by reserving the FFing of certain elements to environments that favour archaising constructions. This restricted distribution of the phenomenon connects with Batllori & Hernanz (2012) proposal of the existence of two grammars in OCat: one which allows FF in polarity reversal contexts and one that does not. The former is more archaising and connects with Latin word order, where focalised constituents were located to the left of the verb (Devine & Stephens, 2006; Sitaridou, 2015, 2016). This option is available more freely in other Old Romance languages (Alboiu et al., 2015; Benincà, 2004; Wolfe, 2015a, Leonetti, 2017), but in the Catalan of the LFRJ database, its distribution is already constrained and started being relegated to embedded clauses and reported speech, two environments that have been noted to be syntactically more archaising (Bybee, 2001; Fleischman, 1991; Marchello-Nizia, 1985; Romaine, 1982; Steiner, 2014). If this holds for OCat, the study of reported speech and embedded clauses might allow us to get a glimpse of two different stages of the language: one, closer to Latin, with the availability of FF, and another, closer to MCat, where only contrastive foci and FF occur preverbally, and IFoc occurs postverbally. *El Llibre dels Fets* displays the first type of grammar, allowing for extensive FFing. However, the distribution of FFing and its frequency rates (higher in embedded clauses and reported speech in the case of ANF and DO fronting) suggests that it is already becoming restricted.

In this section, we have described the distribution of non-finite form anteposition in OCat, establishing that while it can occur in main and embedded clauses, it is more common in the latter. It has also been shown that there is a tendency for non-finite form anteposition to

occur in reported speech (be it direct or indirect), rather than in the narrative body. This distribution could be indicative of the existence of two competing grammars, a more archaising one, which would allow for non-finite form fronting, and a more innovative one, which would disfavour it. This line of inquiry (and which elements lost the availability to be FFed first) is left for future research. Furthermore, it has been established that the anteposition of non-finite forms is not linked to structural requirements, be them the existence of a subject gap, a V2 requirement or clitic placement. Finally, those analyses that link non-finite form anteposition with IS and FF have been confirmed.

5.4.2 Recomplementation

To conclude this overview of IS in the LFRJ database, a further construction is considered: recomplementation. Recomplementation is a phenomenon that has been widely studied in several modern and old Romance varieties, but to my knowledge, never in OCat. Here, I describe and analyse the appearance of two *que* in the same embedded clause found in the LFRJ database and analyse them in line with analyses proposed for this phenomenon for other varieties, such as Modern and OSp (García Cornejo, 2006a; Villa-García, 2015) and MPt (Mascarenhas, 2015) and Italian varieties (Paoli, 2006). However, it is shown that despite displaying some parallelisms with Old Ibero-Romance varieties, recomplementation was much less abundant in OCat than in the other Old Romance languages, even though specific data about its frequency is not available in the literature.⁶²

Here, we use Villa-García's (2015) definition of recomplementation, a term that was first used by Higgins (1988) to describe this phenomenon in Old English:⁶³

- (79) 'The phenomenon of recomplementation consists of one or more left dislocated phrases sandwiched between overt (homophonous) *que* /ke/ complementisers, the second of which is optional in most cases.'

Villa-García (2015, p. 18)

⁶² Fontana (1993, p. 163) states that in OSp: 'Examples like the ones below [containing recomplementation] are by no means exceptional. In fact, in the texts from the 13th to the 16th century that I have examined, the presence of double complementisers in this class of environments appears to be the unmarked option'.

⁶³ Within Ibero-Romance linguistics, other terms have been used for the same phenomenon, like Menéndez Pidal (1908) *que pleonástico*, 'pleonastic que', picked up by García Cornejo (2006a, 2006b), or Wanner (1998) *soubordonnée à double complementeur* 'subordinate clause with two complementisers'.

The first complementiser that Villa-García (2015, p. 18) refers to, corresponds to the subordinating conjunction ‘that’ in English, and following Paoli (2003, 2006), I refer to it as *que*₁. The second complementiser, whose status is not clear at this point, is labelled as *que*₂.

Example (80) illustrates this phenomenon in M_{Cat}⁶⁴, where it is found sparsely in colloquial registers (deemed by Uriagereka [1995] for M_{Sp} as archaising):

(80) Diu la mare que tu, que no vindrà a la festa.
 says the mother that you that not come.3SG.FUT to the party

Mother says that you will not come to the party.

González i Planas (2011, p. 3)

Analyses proposed for this phenomenon widely fall in three groups: firstly, those who defend that *que*₂ is hosted by a recursive ForceP projection (Fontana [1993] for O_{Sp}, Gupton [2010] for M_{Gal}), secondly, those that argue that *que*₁ and *que*₂ are projected in the two extremes of the CP layer, ForceP and FinP (Ledgeway [2005] for SIDs, Paoli [2003] for Tiguise and Ligurian, Dagnac [2012] for Ternois, an oil dialect, and García Cornejo [2006] for O_{Sp}), and thirdly, those approaches that argue that *que* is an overt realization of the head of TopP (Rodríguez Ramalle [2003] and Villa-García (2010, 2012, 2015) for M_{Sp}, González i Planas [2010, 2014] for M_{Sp} and M_{Cat}, and Paoli [2006] for Medieval Romance). First, I

⁶⁴ Despite the evidence provided by González i Planas (2010, 2014) and Montserrat Batllori’s comments on the matter (personal correspondence), I have to argue that recomplementation’s acceptability in M_{Cat} is a matter of debate. Neither I nor the four native speakers that I have consulted accept recomplementation in cases like that of example (80), and only marginally accept recomplementation when the sandwiched left dislocated element is undoubtedly heavy:

- i. *El pare diu que₁ el veí, que₂ es trasllada.
 the father says that the neighbour that REFL.CL=moves
 Father says that the neighbour is moving out.
- ii. ??El pare diu que₁ el veí del sisè que sempre
 the father says that the neighbour from;the sixth that always
 deixa que el gos es pixi al portal, que₂
 lets that the dog REFL.CL=pees.SBJV on;the doorway that
 es trasllada
 REFL.CL= moves
 Father says that the neighbour from the sixth floor who always lets his dog pee in the doorway is moving out.

Example (i), with a light constituent occurring between the two *que*, is deemed as totally ungrammatical by the consulted speakers. However, example (ii), where the sandwiched constituent is significantly heavier, is marginally accepted by some speakers, while it is deemed as ungrammatical by others. Those speakers who marginally accept recomplementation, associate it with oral speech and colloquial registers, in line with Uriagereka (1995) for M_{Sp}.

present the LFRJ data. This is followed by a brief presentation of each of the approaches proposed for this phenomenon, which are assessed against the OCat data. Finally, it is concluded that OCat fits within the analyses proposed for Old and MSp that locate *que2* in TopP, but that this construction did not enjoy the same frequency that it had in OSp (see footnote 60).

In the LFRJ database, there are 419 complement clauses introduced by the subordinating conjunction *que*. Only 16 (3.8%) of them present recomplementation. The 15 cases of recomplementation in the LFRJ database show the following features, which are illustrated below:

- (i) The constituent sandwiched between *que1* and *que2* can only be of topical nature.
- (ii) The heaviness of the sandwiched constituent is not connected to the appearance of *que2*.
- (iii) The presence of *que2* is not linked to the mood of the embedded clause.
- (iv) Recomplementation can only occur with certain predicate types.
- (v) CPs introduced by recomplementation structures are non-referential (they are speech acts of their own).

Firstly, in OCat, the appearance of *que2* required the presence of a topical constituent between the two complementisers regardless of it being light or heavy, as illustrated by examples (81) and (82) (points (i) and (ii)):

- (81) On nós vos pregam molt carament, per ·ii· raons,
 where we to;you.CL= beg.1PL very greatly for 2 reasons
que [vós] que·ns donets conseyl e ajuda.
 that you that=to;us.CL give.2PL.SBJV advice and help

Where we greatly beg you, for two reasons, that you advise and help us.

Fol. 28v, l 2

- (82) E en semblan, que [aquest feyt de què vós nos havets
 and in seeming that this fact of which you to;us.CL= have.2PL
 parlat, de conquerir lo regne de Maylorques, qui és
 spoken of conquer the kingdom of Mallorca who is

dins la mar], que·ns serà major honrament (...).
 Inside the sea that=to;us.CL be.3SG.FUT greater honour

And in a similar way, that this plan that you have exposed to us, of conquering the kingdom of Mallorca, which is in the sea, which will be a greater honour to us (...).

Fol. 29v, 18

In example (81), the element between *que1* and *que2* is the personal pronoun *vós* ‘you’, the addressee of the speaker who is clearly recoverable from the discourse (there is a coreferent 2nd person dative clitic in the main clause). The pronoun receives topic shift reading, since there is a change of subject in respect to the main clause. In (82), the sandwiched element is also the subject of the embedded clause, also previously introduced in the discourse, as the presence of the demonstrative *aquest* ‘this’ shows. However, it is a heavy constituent that contains embedded clauses within.

Different constituents can occur between *que1* and *que2*, all occur either in the Frame or the Topic fields (see section 5.2 in this Chapter for a description of the constituents that can occur in these fields and their value). Table 5 lists them and their frequency:

Location	Constituent type	Number of occurrences
Frame field	Adverbial clause - see example (84)	1
	Conditional clause - see example (85)	3
	Prepositional phrase – see example (83)	1
Topic field	DP – see example (82)	4
	Pron – see example (86)	5
	DO (CILD) – see example (81)	2
Total		16

Table 5 – Sandwiched constituents in clauses with recomplementation

As discussed in section 5.2 in this Chapter, the Frame field is immediately located below Force, the projection where complementisers are hosted. It is in turn followed by the

Topic field. Therefore, the appearance of *que2* is linked to material occurring in these fields. However, this material need not be heavy. Out of the 16 cases of recomplementation found in the LFRJ database, 7 (44%) have light constituents between *que1* and *que2*, and 9 (56%) heavy constituents that themselves are or contain embedded clauses.

Up to this point, we have established that constituents occurring between *que1* and *que2* are either located in the Frame field or the Topic field, and that they need not be heavy. This suggests that *que2* must be located either in a low projection within the Topic field or below it, either in the Focus field or in FinP. Nevertheless, example (83) provides a piece of evidence that clearly advocates for the former analyses:

- (83) E man -vos per la seyoria que hé sobre
 and order.1SG=to;you for the lordship that have.1SG over
 vós que negú no plor ni·n
 you that nobody not cry.3SG.PRS.SBJV nor=of;it.CL=
 faça dol, car nós vos serem seyor, que
 do.3SG.PRS.SBJV greif since we to;you be.1PL.FUT lord that
 [aque] loch que éls vos devien tenir en fer
 that place that they to;you.CL= ought.3PL have.INF in do.INF
 bé];, que nós lo; us farem.
 well that we it.CL= to;you.CL= do.1PL.FUT

And I order you, by the lordship that I have over you, that no one should cry or mourn them, since we will be a lord to you, that that place that they held in your heart in doing good, that we will take that place.

Fol. 40v, 18

In example (83), the sandwiched constituent between *que1* and *que2* is a heavy CILDed DO, *aque] loch que (...)* ‘that place that’. In this case, *que2* is followed by a personal pronoun *nós* ‘we’. Since recomplementation, by definition, occurs in embedded clauses, the fact that clitic pronouns occur proclitically is not a reliable diagnostic. However, in this case it is clear that the pronoun does not appear in SpecTP, given that it is clearly contrasted with the pronoun *éls* ‘they’, located within the heavy CILDed DO. In section 5.2, it has already been shown that in the LFRJ database, there are no clear cases of CTop, but there are instances of CFoc. Hence, we take *nós* ‘we’ to be focalised, and not in SpecTP. Hence, we establish that

*que*₂ occurs above focalised constituents, and its presence is linked to the Frame or Topic fields being activated, and not to the presence or absence of Foci (points (i) and (ii)). In addition, the fact that recomplementation is linked to the activation of the Frame or Topic field allows us to establish that clauses displaying it have a fully fledged left periphery (point (v)).

Furthermore, OCat recomplementation is not linked to the appearance of a specific mood in the embedded clause (point (iii)), unlike in Tiguere and Ligurian (Paoli, 2003). The verb of the embedded clause can appear in either the subjunctive or the indicative, and the mode is determined by the main clause verb. This is illustrated by examples (84-86), and in the rest of examples provided in this section:

- (84) E, enans que moguésem, ordonam l' estol en qual
 and before that move.1PL.SBJV ordered.1PL the navy in what
 manera iria: primerament, que la nau d'En Bovet,
 way go.3SG.COND firstly that the boat of;Sir Bovet
 en què anava En Guillem de Muntcada, que guiàs e
 in which went.3SG Sir Guillem of Muntcada that guide.3SG.PST.SBJV and
que portàs ·i· faró de lanterna, e la d' En Carrós que
 that bring.3SG.PST.SBJV one light of lantern and the of Sir Carrós that
 tingués la reraguarda e que levàs altre
 have.3SG.PST.SBJV the rear-guard and that bring.3SG.PST.SBJV other
 faró de lanterna.
 light of lantern

And before we would move, we told the navy in what way they had to go: firstly, that the boat of Sir Bovet, where Guillem de Muntcada was, that [that boat] would lead and that it would carry a lantern, and that Sir Carró's boat would hold the rear-guard, and that he would carry another lantern.

Fol. 32r, l 28

- (85) 'Seyor, ara hauríem ops a pendre conseyl que en
 Sir, now have.1SG.COND need to take.INF advice that in
 esta nuyt que us gordàssets'.
 this night that yourself.CL= kept.2PL.SBJV

Sir, now we need to take advice that tonight you should keep yourself safe.

Fol. 36v, l 20

(86) E sobre això dixeren que pus que vehiem la yla, que
 and about this said.3PL that since that saw.1PL the island that
 tenien per bo que faéssem calar.
 had.3PL for good that made.1PL.SBJV moor.INF

Fol. 33r, l 27

And about this, they said that since we could see the island, that they deemed it good to moor.

In examples (84-86), the verbs in the embedded clauses appears in the subjunctive mood. This is expected given the predicates found in the main clause of those examples are of directive nature, and select complements in the subjunctive: in (84), the main verb is *ordonam*, ‘we ordered’, and in (85), the directive predicate is expressed by means of a set phrase: *haver ops*, ‘to be necessary’.⁶⁵ In contrast, example (86), where the verb in the main clause is of an assertive nature, the verb of the complement clause is in the indicative, as in the remaining 9 cases. Therefore, it can be established that mood does not interact with recomplementation.

The predicates that take complement clauses with recomplementation fall in three categories (point (iv)): assertive verbs (epistemic verbs and *verba dicendi*), conative verbs, and directive predicates. Furthermore, recomplementation can also occur in complement clauses selected by DPs. This is summarized in Table 6:

Type of head	Verb or expression	Number of cases
Assertive verbs	<i>Dir</i> , ‘to say’	4
	<i>Prometre</i> , ‘to promise’	1
	<i>Fer comptes</i> , ‘to consider’	1
Conative verbs	<i>Pregar</i> , ‘to beg’	2
Directive predicate	<i>Ordenar</i> , ‘to order’	2

⁶⁵ OCat counts with several set phrases to express need or obligation, among which *ésser/haver mester*, *ésser/haver ops*. It also has an impersonal verb, *caldre*, that expresses the same meaning.

	<i>Manar</i>	1
	<i>Prendre conseil</i> , ‘to take advice’	3
	<i>Ser mester</i> , ‘to be necessary’	2
Total		16

Table 6 – Verbs and DPs that take recomplementation in the LFRJ database

These contexts have in common that they take non-referential CPs as complements, that is to say, CPs that contain new information that does not have a referent or antecedent in the discourse. Examples (87) and (88) illustrate the different types of predicates and the one attested DP that can take complements with recomplementation. In addition, each example is framed within its broader context, so as to make evident the non-referentiality of the recomplemented CP.

(87) Assertive verbs

E els dixem que [si neguna re deÿen a
and to;them.CL= said.1PL that if none thing said.3PL to
Don Nuno]; que ho; comprarien.
Sir Nuno that it.CL= buy.3SG.COND

And we told them that, should they say anything to Don Nuno, that they would pay for it.

(88) Conative verb: *pregar* ‘to beg’

e que·ns pregava, con a pare e a senyor en
and that=to;us begged.3SG that as father and as lord in
qui ella havia fiança e sa esperança, que [nós] que
which she has trust and her hope that we that
li ajudàssem.
to;her.CL= help.1PL.PRS.SBJV

And that she was begging us, as a father and as a lord in whom she had trust and hope, that [we] that we helped her.

In example (88), the verb taking the recomplemented clause is *pregar*, ‘to beg’, and the sandwiched element, once more, is the subject of the embedded clause, which in this case, is a light constituent: a personal pronoun. The new content added by the CP is what the countess of Urgell begs to the king: that he will assist her.

(89) Directive predicate:

E	man	-vos	per	la	seyoria	que	hé	sobre
and	order.1SG=to;you		for	the	lordship	that	have.1SG	over
vós	que	negú	no	plor		ni·n		
you	that	nobody	not	cry.3SG.PRS.SBJV		nor=of;it.CL=		
faça		dol,	car	nós	vos	serem	seyor,	
do.3SG.PRS.SBJV		greif	since	we	to;you.CL=	be.1PL.FUT	lord	
<u>que</u>	[aque	loch	que	éls	vos	devien	tenir	en
that	that	place	that	they	to;you.CL=	ought.3PL	have.INF	in
fer	bé];	<u>que</u>	nós	lo;	us	farem.		
do.INF	well	that	we	it.CL=	to;you.CL=	do.1PL.FUT		

And I order you, by the lordship that I have over you, that no one should cry or mourn them, since we will be a lord to you, that that place that they held in your heart in doing good, that we will take that place.

Fol. 40v, 18

In example (89), the verb *manar* ‘to order’ takes two coordinated complement clauses: the former is a regular complement clause, while the latter presents recomplementation. In this case, the sandwiched element is the direct object of the embedded clause. It is certain that it is topical in nature, since it is coreferent with the masculine singular object clitic, thus being a clear case of CILD. The new material introduced by this CP is the assurance by king to his men that he will stand by them and hold them as dear as those men who perished.

(90) DP complement

‘Seyor,	mester	seria	que	[aquests	corsos	qui	són
Sir	need	be.3SG.COND	that	these	bodies	who	are.3PL
morts];	que	els;	soterràs		hom.’		
dead	that	them.CL=	buried.3SG.SBJV	one			

‘Sir, it would be necessary that the dead bodies would be buried.’

Fol. 40r, l 19

The predicate taking the complement clause with recomplementation in (90) is a periphrastic directive predicate compounded by the noun *mester* ‘need’, and the verb *ser* ‘to be’, meaning ‘to be needed, required’. In this case, the sandwiched element is the direct object of the embedded clause, coreferent with the masculine plural object clitic, signalling that this is another instance of CILD. The new content introduced by the CP is the need to bury the corpses, which were already present in the discourse.

So far, it has been established that (i) recomplementation in OCat is linked to the presence of a constituent in the Frame or Topic fields; (ii) that the heaviness of the said constituent does not determine whether *que2* occurs or not, (iii) that the presence of *que2* is not linked to the mood of the embedded clause; (iv) that recomplementation is not linked to a specific predicate type, as it appears with assertive, conative and directive predicates that take (v) complement clauses with a fully fledged left periphery. Therefore, OCat recomplementation fits into Villa-García (2015, p. 24) description of recomplementation in MSp:

(91) ForceP_{primary que} > TopP_{recomplementation/secondary que} > Interrogative P > TopP > FocP
>FinitenessP_{jussive/optative que} ...

From Villa-García (2015, p. 24)

Villa-García identifies two projections that can host *que2* in MSp: one in the Topic field, which he associates with recomplementation, and one associated with jussive and optative predicates (which have been labelled ‘conative’ in this section). He offers ample evidence for the need for this divide to account for OSp data. In our limited pool of clauses with recomplementation from the LFRJ database, we have treated cases of recomplementation with conative predicates like *manar* ‘to order’ in example (89) (jussive) or *pregar* ‘to beg’ in example (88) (optative) as instances of the same phenomena. Unfortunately, OCat data does not allow us to establish whether instances of recomplementation following assertive verbs and those following conative verbs correspond to the same syntactic structure. In the examples examined for this thesis, there are no observable significant differences between the elements found sandwiched in recomplementation cases with assertive or conative predicates, apart from the fact that CILDed DOs are found solely with jussive predicates, like the one case of sandwiched PP, and that the one sandwiched conditional clause occurs with an

assertive verb. However, given the scant amount of evidence, no conclusion can be drawn from this potentially arbitrary distribution.

Therefore, and without discarding the possibility of the existence of two projections capable to host *que2*, a higher one only linked to the activation of the Frame and Topic fields, and a lower one linked to jussive or optative predicates, we suggest that in OCat, *que2* is an optional lexicalisation of the lowest head of the Topic field, below projections hosting CILDed topics and shifting topics. It is worth noting that this construction is marked, as its frequency is very low, and that it is fully optional.

5.5 SUMMARY

In this Chapter, the left periphery of the LFRJ database's clauses has been described. As noted in 5.3, one of its main findings is that IFoc is not located in the left periphery, but in the core of the clause. The connection between the expression of positive polarity and the left periphery has also been explored at length. Section 5.4.1 strengthens this point by connecting the anteposition of non-finite forms to polarity too. Finally, in section 5.4.2, the distribution of recomplementation in the LFRJ database has been explored, concluding that recomplementation in OCat is akin to the same phenomenon in MSp, and by no means pervasive.

6 REGISTER VARIATION AND WORD ORDER

Linguistic variation is inherently linked to registers, speech varieties associated to different usage contexts (Castellà Lidon, 2002, 2004). In Romance historical linguistics, our data sources are limited to written texts. Writing is a secondary coding of speech (Schneider, 2008) and simply by virtue of being written, is subject to constraints that set it apart from oral speech, a fact that should always be considered when producing historical accounts of Old Romance data. In this chapter, we explore the relation between register and word order in OCat and MCat by comparing the findings of Chapters 3 and 4 with data from two MCat texts. It is shown that register plays a crucial role in determining the frequency of syntactic patterns, and that descriptions of Old Romance syntax should always take into consideration the register of the source texts and their wider philological context so as to control for their representativeness and validity.

The Chapter is structured as follows: in section 6.1, the nature of syntactic register variation and the notion of representativeness are considered in regard to *El Llibre dels Fets*. In section 6.2, data from the LFRJ database is compared to that of two MCat databases: a historical novel and a recorded conversation, with especial focus on the impact of mode and register on word order. Finally, in section 6.3, we compare data from an OCat legal text with that of the LFRJ database, exploring variation in word order patterns within OCat.

6.1 *EL LLIBRE DELS FETS*: A REPRESENTATIVE TEXT FOR THE STUDY OF OLD CATALAN SYNTAX?

The study of syntactic variation linked to register does not differ from the study of linguistic variation in any other axis if we depart from the premise that registers, like dialects, have a grammar of their own (Stowell & Massam, 2018, p. 150). Registers vary in terms of channel, and mode, which can be oral or written. Register variation is also linked to the situational context of the production and the reception of the text, which, according to (Castellà Lidon, 2002, p. 50) are the following:

- (1) a. Planification vs. non-planification.
- b. Monologue vs. dialogue.
- c. Unidirectionality vs. multidirectionality.
- d. Shared space vs. non-shared space.
- e. Live reception vs. differed reception.

A further element that intervenes in register variation is genre, which can range from a news headline to a recipe or instructions manual and introduces expectations regarding the abiding of the text onto a set of pre-established conventions.

A divide often used to describe registers is whether they are standard or non-standard, a contrast that dichotomises linguistic variation and simplifies the linguistic space leading to a diglossic view of it, when it may not be the case (Dorian, 2002; Rutten, 2016). Nevertheless, they can also be classified in terms of mode, and in turn, in terms of their degree of colloquiality (which correlates to the standard/non-standard divide). Castellà Lidon (2002, 2004) puts emphasis on the fact that there is a gradual scale of transition between prototypically oral texts (which he identifies with informal conversations), and prototypically written texts (which he identifies with expositive prose). Several attempts of classification of inter-modal texts have been made in the literature. For the purposes of this work, that of Munby (1978) is especially relevant, as he distinguishes written texts in regard to whether they were *written to be heard* or *written to be spoken as if not written*.

Historical linguistics often have a limited pool of texts from which data can be extracted, and they may not always be representative of the distribution of syntactic features within a language, given that medieval texts tend to have been written by powerful noble or ecclesiastic educated men that represent a privileged segment of the population (Elspeß, 2007, p. 5). While this bias is at times unavoidable, considering the representativeness of a text is key for understanding how the data that we are analysing fits the broader picture. The validity of the data (the quality of the recording and contextualisation) is also necessary in order to gain a better understanding of which piece of the puzzle is provided by the data analysed (Schneider, 2008, p. 83).

In Chapter 2, the reasons why *El Llibre dels Fets* was chosen to be the data source for this thesis have been presented: (i) it is the first of the four Catalan chronicles, and therefore, it does not have a direct model; (ii) there is certainty of the King's involvement with its production; and (iii) the manuscript used for the edition used to produce the LFRJ database (Bruguera, 1991) has been convincingly argued to be the closest to the archetype on philological and linguistic grounds.

While (iii) ensures the validity of the data in terms of Schneider (2008, p. 83), its degree of representativeness is tightly linked to the register, formal, and the genre, historiographical chronicle with an intrahomodiegetic narrator. Both register and genre are

arguments that have been brought forward to defend the high degree of orality of the text. Some authors, among which Bruguera (2012) and Ferrando i Francès, (2012, p. 66) defend that the King was the direct author of the text, and that the text can be taken to represent his language, in spite of scribal variations being introduced by scribes (Ferrando Francès, 2012, p. 67). Therefore, is it representative? If we follow Bruguera and Ferrando in assuming that the speech of the King is respected and preserved through the process of writing it down and through the successive copies due to its *holiness*, this text represents the language of a powerful educated man of noble upbringing that had a very particular life, unlike any other man. However, it is necessary to consider that the way in which books are read has radically changed over the last 800 years: *El Llibre dels Fets* was conceived to be read aloud, as Argenter (2006); Casanova (2012, p. 87) point out, and therefore, conceived not as an ego-document reflecting the King's language, but as a document to reach a wider audience, being read in public, as (1) and (2) illustrate:

(2) E, per tal que sàpien aquels qui hoiran
 and so that know.3PL.SBJV those who hear.3PL.FUT
 aquest libre (...)
 this book

And, so that those who will hear this book may know (...)

Fol. 41v, l 15

(3) E, per tal que sàpien aquels qui aquest libre
 and so that know.3PL.SBJV those who this book
 veuran (...)
 see.3PL.FUT

And, so that those that will see the book will know (...)

Fol. 42v, l 24

Note that in (1), the verb used in reference to the book is *hoir* 'to hear', while in (2), the verb is *veure* 'to see', referring to two ways in which the text could be consumed: by being heard or by being read. Therefore, the book was produced with its audience in mind.

On the basis of such evidence, it can be safely assumed that, while the *El Llibre dels Fets* was dictated by King James, and even though his language may not have been the most representative example of 13th century OCat, he had a wider audience in mind, when dictating

El Llibre dels Fets and tailored his language to it. Hence, while not a direct account of vernacular OCat, it provides us with evidence of a narrative text that departs from the well-established *chanson de gèste*, in being written in prose, not being constrained by metric rules, and in that, unlike contemporary and older texts found in OCat, it did not have a pre-established antecedent influencing its syntax through prescriptivism, as is the case with legal (translations of the *Liber Iudicum*, 12th century; *Els Usatges de Barcelona*, 12th century), religious (*Homilies d'Organyà*, end of the 12th or beginning of the 13th century) and lyrical texts (written in OOC). Therefore, *El Llibre dels Fets* is the most likely source of representative data of 13th century Catalan available to us.

6.2 EL LLIBRE DELS FETS AND REGISTER VARIATION IN MODERN CATALAN

To my knowledge, there are no studies of syntactic variation in OCat texts. Here, by comparing the findings presented in the Chapters above with two MCat texts, I aim at illustrating the importance of taking the connection between register and syntax into account when studying Old Romance syntax.

The two databases used for this purpose are built on MCat texts. On the one hand, the *Victus* database is built on a historiographical novel that shares narrative features with *El Llibre dels Fets*: the story is narrated by an intrahodiegetic narrator that identifies with the main character in the 1st person singular. Like *El Llibre dels Fets*, it is narrated retrospectively. Nevertheless, a key difference separates both texts: one was written as a novel, while the other, as discussed above, was probably dictated. On the other hand, *Sobretaula* was built on the record and transcription of a family conversation while and after having dinner. Therefore, it is an unplanned oral text with multiple characters taking the floor.

Table 2 compares the frequency of linear verb position in the three databases:

	V1	V2	V3<
<i>LFRJ</i>	47.4% (474/1000)	44.1% (441/1000)	8.5% (85/1000)
<i>Sobretaula</i>	52.4% (109/208)	37.5% (78/208)	10.1% (21/208)
<i>Victus</i>	34.5% (47/136)	53.9% (72/136)	12.5% (17/136)

Table 2 – Comparison of verb position frequencies in main clauses from the LFRJ database, *Victus* and *Sobretaula*

As shown in Table 2, the two MCat databases show clear differences in their

frequency of V1 and V2 orders: while in *Sobretaula* V1 clauses account for 52.4% of main clauses, in *Victus* the frequency drops to 34.5%, with a difference of 17.9%. In regards to V2 clauses, they represent 37.5% of main clauses in *Sobretaula*, and 53.9% in *Victus*, with a difference of 16.4%. In both cases, V3< clauses are the least attested, counting for 10.1% of main clauses in *Sobretaula* and 12.5% in *Victus*. When compared to the LFRJ database, *Sobretaula* patterns with it in having V1 clauses as the most frequent, with *Sobretaula* presenting 5% more V1 clauses than the LFRJ database. V2 clauses are the second most frequent in both databases, being 6.6% more frequent in the LFRJ database than in *Sobretaula*. In both cases, V3< is the least frequent order.

Linear verb position is not the only area in which *Sobretaula* and the LFRJ pattern together. In Tables 3, 4 and 5, we present the frequency of different types of preverbal constituents in main clauses in the three databases (Table 5 being an adaptation of Table 4 in Chapter 4).

	Non Pro-Drop		Total
	Structure	Occurrences	
Preverbal Subjects	S _{LEX} V	54 (73% of preverbal subjects)	93.7% (74/79)
	S _{PRO} V	20 (27% of preverbal subjects)	
Postverbal subjects	ADV V S _{LEX}	2	6.3% (5/79)
	ADV V S _{PRON}	-	
	DO V S _{LEX}	-	
	DO V S _{PRON}	-	
	PP V S _{LEX}	-	
	PP V S _{PRON}	-	
	SC V S _{LEX}	1	
	SC V S _{PRO}	2	
Total		79/79	100%

Table 3 – Preverbal constituents in *Victus* non-prodrop V2 main clauses

	Non Pro-Drop		Total
	Structure	Occurrences	
Preverbal Subjects	S _{LEX} V	16 (42.1% of preverbal subjects)	82.6% (38/46)
	S _{PRO} V	22 (57.9% of preverbal subjects)	
Postverbal subjects with verb – subject adjacency	ADV V S _{LEX}	1	17.4% (8/46)
	ADV V S _{PRON}	4	
	DO V S _{PRON}	1	
	SC V S _{PRO}	1	
	ADV V SC S _{PRO}	1	
Total		46 (100%)	100%

Table 4 – Preverbal constituents in *Sobretaula* non-prodrop V2 main clauses

	Non Pro-Drop		Total
	Structure	Occurrences	
Preverbal Subjects	S _{LEX} V	110 (42.5% of preverbal subjects)	72.1% (259/359)
	S _{PRO} V	149 (57.3% of preverbal subjects)	
Postverbal subjects with verb – subject adjacency	ADV V S _{LEX}	36	23.2% (83/359)
	ADV V S _{PRON}	8	
	DO V S _{LEX}	2	
	DO V S _{PRON}	10	
	PP V S _{LEX}	16	
	PP V S _{PRON}	7	
	SC V S _{LEX}	4	
Non-verb adjacent postverbal subjects	DO V PP S _{LEX}	1	4.7% (17/359)
	PP V PP S _{LEX}	1	
	PP V PP S _{PRO}	1	
	ADV V PP S _{LEX}	8	
	ADV V PP S _{PRON}	1	
	ADV V DO S _{LEX}	2	
	PP V DO S _{LEX}	1	
	ADV V SC S _{LEX}	3	
	SC V SC SN	1	
Total		359	100%

Table 5 – Preverbal constituents in non-prodrop V2 main clauses in the LFRJ database

SV orders are the most frequent in clauses with overt subjects across all databases, with *Victus* showing the highest proportion, with 93.7%. In *Sobretaula* and the LFRJ database, the percentage of SV clauses is lower: 82.6% and 72.1% respectively. Interestingly, *Victus* also differs from the other two databases in that lexical preverbal subjects are more frequent than pronominal, representing 57.5% and 42.5% of the total of preverbal subjects respectively. In *Sobretaula* and the LFRJ pronouns are the most frequent preverbal subjects, being 57.9% of preverbal subjects in the case of *Sobretaula* and 57.3% in the case of the LFRJ database, roughly 15% more common than in the *Victus* database.

Therefore, up to this point, we have identified parallelisms in the frequency of V1 and V2 clauses between the LFRJ database and the *Sobretaula* database, as well as the frequency of preverbal subjects and the prevalence of pronominal subjects over lexical ones.

The reason for these parallelisms between these two databases is clearly their oral nature: *El Llibre dels Fets* was composed with the listeners of the book in mind and in all probability, dictated by the king, while the data on which *Sobretaula* is built is extracted from a recorded conversation. Nevertheless, their mode is clearly different: while *Sobretaula* is extracted from a prototypical oral text, an informal conversation, *El Llibre dels Fets* is a written text. However, as pointed out by Castellà Lidon (2002, 2004), the mode of texts is not simply divided in the written/oral dichotomy: there are oral texts that present features usually associated to written texts, such as conference presentations, and written texts that tend towards orality, such as monologues. Here, I will argue that *El Llibre dels Fets* was a monologue written to be heard (Munby, 1978). This characterisation fits with the narrative and philological features described above: mainly the fact that *El Llibre dels Fets* was dictated by King James I, and that it had an intrahodiegetic narrator identified with the main character of the story recounted.

Castellà Lidon (2002, p. 72) offers the following characterisation of oral texts according to their degree of formality:

	Colloquial oral text	Non-colloquial oral text
Theme	General	Specific
Mode	Spontaneous	Planned
	Dialogue	Monologue
	Multidirectional	Unidirectional
	Interactive	Informative
	Informal	Formal

Table 4 – Characterisation of oral colloquial and non-colloquial texts, adapted from Castellà Lidon (2002, p. 72)

Interestingly, *Conversa al menjador 09*, the text on which the *Sobretaula* database is based, and *El Llibre dels Fets* match each of these categories perfectly. *Conversa al menjador 09* is an informal conversation taking place at dinner time with no fixed topic. Discourse is built spontaneously, and each of the participants intervene in the dialogue, rendering it multidirectional and interactive. Therefore, it matches the characterisation of colloquial oral texts. On the other hand, *El Llibre dels Fets* matches that of non-colloquial oral texts: it has a specific topic, the life and work of King James I, while dictated, the story line was planned (Ferrando Francès, 2012, p. 67), and the text is a monologue, narrated by the intrahomodiegetic narrator that identifies with the main character. As a monologue, it is conceived to be a unidirectional text: the receptor cannot answer back to it, and it is thus informative.

Therefore, *El Llibre dels Fets* is a good example of how the written/oral mode dichotomy of registers should be avoided in the benefit of a more nuanced classification in register, which should include written texts of oral nature. The orality of *El Llibre dels Fets* goes beyond ticking boxes in Castellà's characterisation. It can account for the parallelism between the *Sobretaula* and the LFRJ databases in terms of verb position. As Castellà Lidon (2004, p. 33) points out, oral texts tend to be verbal, while written texts tend to be nominal. That is to say, the density of nouns and verbs varies in either kind of text: while in oral texts they are much more frequent, written texts favour nominalisations and other strategies, rather than the use of verb (and, as a consequence, the formation of a new clause). This observation could be connected to the reason why V1 orders are more prevalent in the two databases stemming from texts of oral nature, while they are unpreferred in *Victus*. Discourse topic

continuity in oral texts might be to blame for it: the need for Scene Setters decreases when the discourse topic remains stable throughout a string of speech. Other consequences of discourse topic continuity are the pervasive use of discourse markers or polysemous conjunctions, such as *e* ‘and’ (see Chapter 3, section 3.1.2.1) or *que* ‘that, because, since’, among others. Castellà Lidon suggests that the use of these discourse connectors is connected to an extra-clausal position which could be identified with the Speech Act layer proposed by Haegeman (2014). Orality is also the reason behind the higher prevalence of pronominal subjects in the oral databases: repetition is a feature of oral texts, and deemed necessary for the receptor (Castellà Lidon, 2004): overtly repeating pronouns to express not only Topic Shift, but also Topic Continuity is key in the avoidance of ambiguity in oral texts.

In this section, I aimed at illustrating the importance of bridging the study of syntactic variation linked to register variation, so as to be able to better account for word order patterns found in different types of texts and understand which registers are more innovative and closer to vernacular language. The data provided above has also further confirmed the oral nature of *El Llibre dels Fets*, and thus, of the LFRJ database, reaffirming us in choosing this text for the study of OCat word order.

6.3 *EL LLIBRE DELS FETS* AND THE LEGAL REGISTER

As mentioned in section 6.2, contemporary or earlier OCat texts to *El Llibre dels Fets* present the handicap of belonging to highly codified genres that translate into prescriptivism. The legal register was highly constrained by the format imposed by the Roman Law of the *Liber Iudicum*, a legal compendium that collected the laws that affected ‘Roman’ citizens after the Visigothic invasion, and that remained in place during and after the Reconquest. Phrasing directly extracted from the compendium’s legal formulae and pre-established forms are found pervasively in Catalan legal documents up to the 14th century, when the Cancelleria, the scribal body of the Crown of Aragon, intensifies its activity and starts homogenising written Catalan.

The syntax of legal texts was, therefore, conditioned by the existing pre-established models: the purpose and subject of the document also interact with it. Most early documents in OCat (up to the 13th century) that have reached us fall in four categories: (i) feudal oaths and agreements; (ii) wills and donations; (iii) grievances and complaints; and (iv) religious homilies. Feudal oaths and wills are highly codified documents, where there is little space for improvisation. There is only one attestation of a homily, the *Homilies d’Organyà*, which is

suspected to be a translation of an Occitan text, and therefore, not the best source of data for the study of OCat. Grievances, on the other hand, are characterised for being accounts of personal matters in the first person. The first attested grievances are the *Greuges de Guitard Isarn, Senyor de Caboet*, ‘the Grievances of Guitard Isarn, Lord of Caboet’ (henceforth ‘the *Greuges*’), dated between 1080 and 1095. In them, Guitard Isarn states the offences that he has received from Mir Arnau, son of Guillem Arnau. Table 5 shows the frequency of linear verb positions in the *Greuges*:

Verb Position in main clauses	V1		V2	V3<	Total
	Greuges	Total	82.3% (28/34)	14.7% (5/34)	3% (1/34)
eV1		67.6% (23/34)			
e-less V1		14.7% (5/34)			
LFRJ		47.4% (474/1000)	44.1% (441/1000)	8.5% (85/1000)	100% (1000/1000)

Table 5 – Verb linear verb position frequencies in the *Greuges*

As shown in Table 5, the predominant linear position of the verb in the *Greuges* is clearly V1, V1 clauses being 35% more frequent than in the LFRJ database. *E*-less V1 clauses are significantly present, making up 14.7% out of the total, a higher rate than that of the LFRJ database. V2 clauses are much less frequent, accounting only for 14.7% of the total, 29.4% less than in the LFRJ database. Finally, V3 clauses are the least attested, only making up to 3% of clauses, 5.5% less frequent than in the LFRJ database.

Like in the case of LFRJ, clauses tend to be preceded by the coordinating conjunction or discourse cohesion marker *e*, be them V1 or V2. This is illustrated in (3), where a string of sentences from the *Greuges* is reproduced:

- (4) [Et *dixit* mihi Guilelm Arnall]₁ et [coveng
and said.3SG to;me.DAT Guillem Arnau and accorded.3SG

m'ò que (...)]₂. *Et* [*ego dixi ad* Mir Arnal in
 =to;me=it.CL that and I said.1SG to Mir Arnal in
 presencia de Guilelm Arnall que]₃ (...). *Et* [**fuit**
 presence of Guillem Arnau that and was.3SG
 rancuros de Mir Arnall che (...)]₄. *Et* [**rancur-me·n**
 displeased of Mir Arnall that and displease.1SG=me=of;it.CL
 del castel de Cabodet (...)]₅, e [**rredí·l** a Guilelm
 of;the castle of Caboet and bequeathed.1SG=it.CL to Guillem
 Arnall (...)]₆. *Et* [*el redé·l a Mir Arnal*]₇ et
 Arnau and he bequeathed.3SG=it.CL to Mir Arnau and
 [**fed-le·n** fer convenença (...)]₈. *Et* [*illum*
 made.3SG=him.CL=of;it.CL do.INF agreement and this
tortum che el me·n fed et dreçar
 offence that he to;me=of;it.CL=made.3SG and amend.INF
 no em volg]_i [**jachí·l** a *ssuo filio*]₉,
 not to;me.CL=wanted.3SG bequeathed.3SG=it.CL to his son
 et [*jo fuit rancuros del fil*]₁₀; et [*fforon-ne*
 and I was.1SG displeased of;the son and were.3PL=of;it.CL
 iudicis donads (...)]₁₁.
 trials given

And Guillem Arnau said to me and accorded to me that (...) And I said to Mir Arnal
 in the presence of Guillem Arnau that (...). And I was very displeased that Mir Arnall
 (...). And I was displeased about the Castle of Caboet, and I bequeathed it to Guillem
 Arnau (...). And he bequeathed it to Mir Arnau and made him enter an agreement
 (...). And this offence that he committed against me and that he does not want to
 amend, he gave to his son, and I was very displeased with the son and trials took place
 (...)

Greuges, 1 13-27

(3) presents a string of clauses linked by *e*. As expected in a legal text of this period, we see the pressure for a legal document to be written in Latin in the usage of Latin verbal

forms of the verb DICERE ‘to say’ in the first and third clause, or Latin legal phrasing like ILLUM TORTUM ‘this offence’, in sentence 9. Code-switching into Latin occurs only at phrase level.⁶⁶ The first clause presents VS order, a feature that is not unusual with *verba dicendi*, as shown in Chapter 3, section 3.1.2.2. The second clause is coordinated with the first by *e* and shares the DO with the first. The third clause is a V2 clause, with the preverbal 1st person singular Latin pronoun EGO occurring preverbally, signalling topic shift from the preceding clauses, where the subject was Guillem Arnau, to Guitard Isarn, the plaintiff. The fourth clause is a V1 that displays subject continuity with the third and expectedly displays V1. The same is true of clause 5 and 6. Clause 7 is a V2 clause with a preverbal subject, *el* ‘he’, carrying out, once more, topic shift from the preceding clause and going back to Guillem Arnau as subject, coordinated with clause 9, with which there is subject continuity. Clause 9 is a V2 clause with a preverbal CILD direct object, coreferent with a clitic pronoun in the core of the clause and a null subject coreferent with that of clauses 7 and 8. In clause 10, once more we find an overt subject pronoun carrying out topic shift. Finally, clause 11 is a V1 passive clause, with the subject occurring between the auxiliary verb *fforon* ‘they were’ and the past participle *donads* ‘given’. In this case, in spite of there being a subject change from the preceding clause, the subject does not appear preverbally.

The fragment reproduced in (3) that has just been described shows that the triggers for the prevalence of V1 clauses in *Greuges* are the same as those identified for LFRJ: it can be linked to predicate type, as illustrated by clause 1; it can be linked to subject continuity, but it need not to; and it can be linked to discourse topic continuity. Preverbal constituents in V2 clauses in (3) fall in two categories: subject pronouns and DOs. Subject pronouns invariably carry out topic shift. Their location in the Topic field is further confirmed in clause 7, where the presence of the preverbal pronoun does not trigger proclisis (see Chapter 3, section 3.1.5 for a discussion on clitic placement in the LFRJ database). The one preverbal DO is an instance of CILD, as the fronted object is coreferent with a resumptive clitic pronoun that occurs enclitically in the core of the clause. In this case, the preverbal constituent is directly merged in the left periphery, and therefore, could not satisfy a hypothetical V2 requirement

⁶⁶ Latin phrases are inserted within clauses with clear SVO grammar, as expected in OCat, not SOV, as expected in Latin. This, together with the pervasiveness of OCat functional morphemes allows us to establish that the matrix language (that which provides the syntax within Myers-Scotton 4-Morpheme model) of this text is OCat (Myers-Scotton & Jake, 2009; Myers-Scotton & Jake, 2017).

(as discussed in Chapter 5, section 5.2, the status of topicalised subjects is unclear given the lack of resumptive clitic pronouns.

Before concluding this section, let us consider a string of absolute V1 clauses:

(5) [**Rancur**-me·n de la cavalleria de Mir Guilabert
upset.1SG=REFL.1SG=of;it.CL of the knights of Mir Guilabert
qui no m'es servida.]₁
who not to;me.CL=is served

[**Rancur**-me quar desmentist ma mulier ante me (...)]₂
upset.1SG=REFL.1SG=of;it.CL since refuted.3SG my wife in front of me

et [rancur-me de Mir Arnall de la casa de
and upset.1SG=REFL.1SG=of;it.CL of Mir Arnall of the house of

Boxtera que s'a presa e m'a
Boxtera that for;himself.CL=has taken and to;me.CL=has

tolta (...)]₃

taken

[**Rancur**-me de I bou d'Oliba d'Ares, que (...)]₄
upset.1SG=REFL.1SG=of;it.CL of 1 bull of;Oliba d'Ares that

[**Rancur**-me de I vaca de Guilelm Oler, que (...)]₅
upset.1SG=REFL.1SG=of;it.CL of 1 cow of Guillem Oller, that

[**Rancur**-me de la casa de Pere Baro, que (...)]₆
upset.1SG=REFL.1SG=of;it.CL of the house of Pere Baró that

I am upset about the knights of Mir Guilabert, who are not available to me. I am upset since he refuted my wife in front of me (...), and I am upset about Mir Arnall having taken for himself and from me the house of Boxtera (...). I am upset about one bull of Oliba of Ares (...). I am upset about one cow of Guillem Oller, that (...). I am upset about the house of Pere Baro, that (...).

Greuges, 141-48

In (4), we are presented with a string of V1 sentences that are not linked by *e*. This fragment is located towards the end of the *Greuges*, and therefore, the discourse topic (the offence of Guilelm Arnau's son against Guitard Isarn) is already established, as well as the

only possible referent for a 1st person singular subject: Guitard Isarn. Like in the case of the LFRJ, absolute V1 clauses need not present all-focus: they can add content on the discourse topic, as is the case here. The string of *e*-less clauses adds specific offences to the main one that Guitard Isarn is suffering. The choice of repeating the predicate *rancur-me de* ‘I am upset/angered about’ might be purely a rhetoric device emulating oral language, the text being written to be heard and to sound as if it were spoken in a court, probably. This further piece of evidence suggests that 11th century vernacular Catalan might be structurally quite similar to that of the 13th century, which has been described in previous chapters.

Hence, it can be established that formally, the *Greuges* strongly resemble the data of the LFRJ database described above. The difference in the frequency of V1 and V2 orders is thus not linked to syntactic divergences, but to register: the *Greuges* are subject not only to formal constraints dictated by the nature of the text, which was to be read aloud in a court, but also material. It is written on a 102 x 527 mm piece of parchment, taking advantage of every millimetre of surface to write. The lower frequency of V2 and V3 orders can be thus connected to the lack of Scene Setters, partly not needed given that the *Greuges* only have one discourse topic, which is established from the first line, and everything else builds on it linearly, partly due to space constraints.⁶⁷

6.4 SUMMARY

In this Chapter, we have explored the impact of register on word order. In section 6.1, we have considered whether *El Llibre dels Fets* can be taken to be representative of 13th century OCat given its register. In section 6.2, we have embarked on a comparison between the data from the LFRJ database and two MCat texts. It has been shown that even across centuries, register is a determining factor in establishing word order patterns, specifically in regard to where texts are located in the written/oral continuum. Finally, in section 6.3, data from the LFRJ database has been compared with that of a 12th century legal text that superficially, presents rather different proportions of linear verb order. Nevertheless, it is shown that structurally, the *Greuges* are parallel to *El Llibre dels Fets* and that variation in verb position frequency is determined by register and material constraints.

⁶⁷ See Pujol i Campeny & Sitaridou (2017) for a study of word order variation linked to register in OOc.

7 CONCLUSION AND FUTURE PROSPECTS

7.1 SUMMARY OF THE THESIS AND ITS MAIN FINDINGS

In Chapter 3 we have argued that in OCat, the verb does not raise beyond the inflectional domain in unmarked main declarative clauses. This conclusion has been reached after examining verb position in unmarked main clauses from different perspectives.

Firstly, in section 3.1, main clauses are considered. In 3.1.2, V1 clauses have been shown to be actual V1 clauses, whether preceded by the *e* or not. This has been done by examining the relation of *e* and V1 clauses, the link between V1 clauses and certain predicate types, and absolute V1 clauses and Informational structure configurations. The possibility of *e* rendering V1 clauses grammatical within a V2 grammar has been dispelled by showing that (i) *e* is located within a clause external position, and that (ii) while it can be linked to discourse topic continuity, its function is not only restricted to V1 clauses (*e* preceding 88.33% of the total of clauses occurring in the LFRJ database). The link between V1 clauses and certain predicate types, mainly *verba dicendi* and unaccusative predicates. In the case of *verba dicendi*, it has been shown that while there is a tendency for this predicate type to occur in V1 configurations, it is not the only possible word order pattern found with these predicates. The informational value of the subject, the object, as well as the informational value of the clause as a whole interact in determining whether a clause with a *verbum dicendi* will be V1 or not. In the case of unaccusative predicates, it has been shown that V1 orders are linked to certain clausal informational readings and involve the presence of a null locative in SpecTP. Finally, absolute V1 clauses are considered. It is shown that they do not always correlate with all-focus: all-focus can be given to non-V1 clauses, and absolute V1 clauses, not preceded by *e*, can have bipartite information structure (old/new divide). It is concluded that both *eV* and absolute verb-initial verbal clauses are truly V1 and do not conform to a V2 grammar. Instead, they exhibit features parallel to SVO grammars.

Then, the position of the verb in V2 and V3 clauses is considered. The description of V3< clauses shows that several constituents can occur preverbally. Their relative order, linked to Informational Structure, is explored at length in Chapter 5. While the existence of V3 orders has been used in the literature as evidence against V2 analyses of the Old Romance languages, Wolfe (2015a) defends that they should not hinder a V2 analysis. In section 3.1.4, it is shown that in unmarked declarative clauses, the verb did not raise to the left periphery by means of Cinque (1999) adverbial hierarchy. In section 3.1.5, the interaction of clitic

placement and verb position is explored. It is shown that the oscillation between enclisis and proclisis is not linked to polarity-related verb movement, as proposed by Fischer (2002). Instead, data from the LFRJ database confirms Batllori et al. (2005) hypothesis that links the oscillation between proclisis and enclisis to the saturation of FocP, like in Modern Western Iberian varieties. Finally, in section 3.1.6 the interaction between emphatic positive polarity and verb position is explored. It is shown that OCat did not display any of the strategies for the expression of emphatic polarity in positive answers to direct questions available in languages that independently license V-to-Pol, mainly V-echo and do-support. Instead, the expression of emphatic positive polarity is already always linked to the presence of the Emphatic Positive Polarity Particle (EPPA) *sí*, even in cases where V-echo and do-support are the chosen strategies. The syntactic contexts in which *sí* occurs in *El Llibre dels Fets* have allowed us to establish that already by the 13th century, *sí* had undergone grammaticalisation, from being a manner adverb in Latin, SIC ‘thus’, and had become an EPPA. The distribution of *sí* differs significantly from the distribution of *sí* in neighbouring varieties and allows us to label Old Catalan as innovative in this respect. In the presence of *sí*, verb movement to the left periphery is triggered by the same mechanism that triggers V-to-C in Focus Fronting cases. Therefore, we have established that OCat does not display V-to-C movement in unmarked main clauses, but that it does take place when Focus Fronting, a movement operation linked to emphatic positive polarity, takes place.

In section 3.2, verb position in embedded clauses is examined. It is established that the position of the verb in embedded clauses is the same as the position of the verb in main clauses: it does not raise to the left periphery. Instead, it remains in the inflectional domain.

From the evidence presented in Chapter 3, it is established that in the OCat of the LFRJ database, there was no V-to-C movement in unmarked declarative clauses.

Chapter 4 has focused on the description of subject position in the data from the LFRJ database. Firstly, in section 4.2, the Old Catalan preverbal field is explored, and preverbal positions available for subjects are identified. It is shown that subjects can be topicalised in the left periphery, even though it is not clear whether they can be Clitic Left Dislocated or Hanging Topics, given that Old Catalan did not have subject clitic pronouns. They can also be focalised, to receive an emphatic reading (object clitic pronouns have been used as a test to identify whether subjects are focalised or topicalised). Finally, the third preverbal position identified for subjects is SpecTP, specialised for hosting subjects. Apart from nominal expressions, other elements can also occupy this position, given its featural make-up. It is thus

concluded that unlike the Old Romance languages that have been described as having a V2 grammar, Old Catalan has a specialised position for hosting preverbal subjects, which are, incidentally, the most common type of preverbal constituent. Three postverbal subject positions are identified: postverbal subjects can occur in their base-generated position, right dislocated (a phenomenon not instantiated often in the Old Romance languages), and focalised in the low left periphery, which in Old Catalan was still active. Another finding of this section is that Old Catalan did not have a definiteness effect. In section 4.3, the distribution of subjects in embedded clauses is described. It is shown that, unlike in those languages that have been described as V2, there is no asymmetry in the distribution of pro-drop subjects between main and embedded clauses, as it is the case in Modern Catalan. Therefore, this Chapter provides further evidence against a V2 analysis of Old Catalan and in favour of an SVO one.

In Chapter 5 we have explored the Information Structure of Old Catalan. It has been shown that, while the left periphery of Old Catalan shares the structure proposed by Benincà (2006), data from the LFRJ database differs from descriptions of the left periphery made for other Old Romance varieties in one key and important aspect: Old Catalan did not have an Information Focus position in the left periphery. Instead, like its Modern counterpart, Information Focus was encoded postverbally. Apart from describing the left periphery of Old Catalan, in this Chapter we have also explored two other phenomena: anteposition of non-finite forms, in section 5.4.1 and recomplementation, in section 5.4.2. In the former, we agree with Batllori (2015) analysis of participial fronting and link it to Focus Fronting and emphatic polarity. In the case of recomplementation, we analyse it along the lines of Villa-García (2015) for recomplementation of Modern Spanish, in that 2nd *que* is an optional lexicalisation of TopP.

In Chapter 6, we have explored the connection between textual mode (written and oral) and register with word order. To do that, we have used two Modern Catalan databases of very different registers: one built on the recording of an informal conversation and the other on a historiographic novel with very similar narrative features to *El Llibre dels Fets*. Surprisingly, in spite of sharing many genre features with the latter, the data from the LFRJ database strongly patterns with the oral text, and not with the one with which it shares genre features. This is so because *El Llibre dels Fets*, in spite of being a written text, it was conceived to be *heard*, to be received orally, and thus, it displays features that tend to correlate with oral texts, such as polysyndeton and abundance of V1 clauses. This conclusion has been enhanced by

the comparison between the LFRJ database and the data from an OCat legal text. In this case, superficial word order differences are also due to the interaction between genre and register, while syntactically, the text abides by the same rules as *El Llibre dels Fets*.

7.2 FUTURE PROSPECTS

The synchronic and descriptive nature of this thesis has pointed at several issues that deserve further investigation.

Firstly, more needs to be known about Old Catalan to allow us to draw a picture of how the language evolved from Latin, a SOV language, to the SVO language that has been described above. The main questions to be answered are: did Old Catalan go through a V2 phase, as it has been proposed for other Old Romance languages? If so, was it innovative, in evolving into developing SVO grammar before other varieties? If not, what was the evolutionary path from Latin to Old Catalan? To answer these questions, the study of older texts and Late Latin/Early Romance data will be required. The pool of Old Catalan texts dating from before the 13th is relatively small. Nevertheless, the *Usatges de Barcelona*, a 12th century text, present themselves as a good candidate from which data can be drawn. A quick glance through it suggests that the syntax of this text differs significantly from that of *El Llibre dels Fets* and may add valuable information to the history of the evolution of Catalan word order.

Connected to this matter is the relation of Old Catalan with Southern Occitan dialects, which have been found in the literature to show fewer V2 features than northern Occitan varieties. If, by the 13th century, Old Catalan was already an SVO language, as the data from the LFRJ database has shown, it could be proposed that it was an innovative variety also in this respect, and that Southern Occitan Dialects constituted the boundary where the isogloss between innovative SVO varieties and archaising V2 ones is found. Of course, more research about the interaction between Old Catalan and Southern Occitan dialects is needed to better understand the interaction between these two varieties that, during the early middle ages, probably constituted a strong and active dialectal continuum.

Similar questions arise in regard to the connection between Old Catalan and Old Spanish. Certain phenomena, such as the increase of proclisis, have already been described as following an east-to-west expansion in the peninsula, which ends up yielding the current picture of the distribution of clitics. Could it be possible that Old Catalan was also innovative

in terms of word order and information structure, particularly in regard to the distribution of Information Focus?

Apart from these matters that concern the role of Old Catalan within the broader Old Romance picture, more specific questions regarding 13th century Old Catalan. Among them we find questions connected to the expression of positive polarity. As Batllori & Hernanz (2008) point out, Modern Catalan has a wealth of emphatic positive polarity particles, including *bé*, *pla* and *ja*. It would be interesting to explore whether they follow exactly the same grammaticalisation path as *sí*, in (non)veridical contexts, or whether their grammaticalisation occurs in different circumstances. This could contribute to the better understanding of the role of Polarity in the Catalan left periphery.

Connected to polarity might be the expression *sí Déus m'ajut*, which does not lend itself to a straightforward analysis, given the multiplicity of meanings that can be attributed to *sí* in the Old Romance varieties. Understanding the origin of this expression, as well as its syntax and spread through the Old Romance languages may contribute to our understanding of contact dynamics among them.

A further matter related to *sí* that requires further investigation is the interaction of the emergence of *sí* as a positive polarity particle and the decrease and subsequent obsolescence of the subordinating adverbial conjunction *sí que*. The grammaticalisation of *sí que* as the means of expression in Modern Catalan is also to be explained.

Furthermore, as pointed out on various occasions throughout this thesis, Old Catalan had an active low left periphery that is still accessible to certain lexical items in Modern Catalan. The study of the evolution of the low left periphery and its connection to word order changes also remains to be explored.

Further study of the relationship between word order and register and genre, not only within Old Catalan, but also across the Old Romance languages, would also contribute to providing us with more tools to interpret the data used in historical syntax. A potential start point could be the study of Old Catalan legalese, especially from documents such as the *Greuges*, where the language reflected is closer to vernacular languages than in other document types. The exploration of code-switching into Latin in this kind of document can also shed light on the weakening of Latin-Romance diglossia.

Last, but not least, the creation of a syntactically parsed corpus of Old Catalan would not only benefit work along the lines of this thesis, but also make the language more accessible to those working in Old Romance linguistics.

8 REFERENCES

- Adams, J. N. (1994). Wackernagel's Law and the Position of Unstressed Personal Pronouns in Classical Latin. *Transactions of the Philological Society*, 92, 103–178.
- Adams, M. (1987). From Old French to the Theory of Pro-Drop. *Natural Language & Linguistic Theory*, 5(1), 1–32. Retrieved from <http://www.jstor.org/stable/4047573>
- Alboiu, G., Hill, V., & Sitaridou, I. (2015). Discourse-driven V-to-C in Early Modern Romanian. *Natural Language and Linguistic Theory*, 33(4), 1057–1088. <https://doi.org/10.1007/s11049-014-9270-8>
- Alegre, M. (1991). *Dialectologia catalana*. Barcelona: Teide.
- Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing Agr: Word Order, V-Movement and EPP-Checking. *Natural Language & Linguistic Theory*, 16(3), 491–539.
- Alexiadou, A., Anagnostopoulou, E., & Everaert, M. (Eds.). (2003). *The unaccusativity puzzle*. Oxford: Oxford University Press.
- Argenter, J. A. (2006). Responsibility in Discourse : Evidence, Report and Entitlement to Speak in the " Book of Deeds of King James ". *Language in Society*, 35(1), 1–25. Retrieved from <http://www.jstor.org/stable/4169476>
- Aurell, J. (2005). From Genealogies to Chronicles. The Power of the Form in Medieval Catalan Historiography. *Viator*, 36, 235–264.
- Aurell, J. (2008). La chronique de Jacques Ier, une fiction autobiographique. Auteur, auctorialité et autorité au Moyen Âge. *Annales. Histoire, Sciences Sociales*, 63, 301–318.
- Axel, K. (2007). Verb first declaratives. In *Studies on Old High German Syntax Left Sentence Periphery, Verb Placement and Verb-Second*. (pp. 113–172). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Badan, L., & del Globo, F. (2010). On the Syntax of Topic and Focus in Chinese. In P. Benincà & N. Munaro (Eds.), *Mapping the Left Periphery*. Oxford; New York: Oxford University Press.
- Batllori, M. (1993). Participle preposing and other related phenomena in Old Spanish and Old Catalan. Girona.

- Batllori, M. (2009). *Ordre de mots i estructura informativa a la llengua de Francesc Eiximenis*. Girona. Retrieved from <http://hdl.handle.net/10256/3961>
- Batllori, M. (2015). Análisis comparativo de las construcciones de dislocación a la derecha y su incidencia en el orden de palabras del español, catalán y portugués medievales. In M. Castillo Lluch & M. López Izquierdo (Eds.), *El orden de palabras en la historia del español*. Madrid: Visor.
- Batllori, M. (2016). El valor modal de haber en los futuros y condicionales analíticos. In C. de Benito Moreno & O. S. Octavio de Toledo y Huerta (Eds.), *En torno a haber: construcciones, usos y variación desde el latín hasta la actualidad. Studia Romanica et Linguistica: 46* (pp. 33–78). Frankfurt am Main.
- Batllori, M., & Hernanz, M. L. (2008). *Emphatic Polarity from Latin to Romance. Poster. The 10th Diachronic Generative Syntax Conference*. New York. Retrieved from http://habilis.udg.edu/~info/Activitat_Docent_Recerca_Gestio/Montserrat_Batllori/BatlloriHernanzPosterDIGSXCornell.pdf
- Batllori, M., & Hernanz, M. L. (2011). Variación y periférea izquierda: la anteposición de foco débil en español y en catalán. *18 Congreso de La Asociación de Hispanistas, Sección 13*.
- Batllori, M., & Hernanz, M. L. (2013). Emphatic polarity particles in Spanish and Catalan. *Lingua*, 128, 9–30. <https://doi.org/10.1016/j.lingua.2012.11.010>
- Batllori, M., Iglesias, N., & Martins, A. M. (2005). Sintaxi dels clítics pronominals en català medieval. In Pérez Saldanya, M. & Martines, J. *Caplletra 38. Revista Internacional de Filologia (Primavera 2005). Monogràfic sobre "Sintaxi Històrica"*. Valencia: Institut Interuniversitari de Filologia Valenciana & Publicacions de l'Abadia de Montserrat.
- Belletti, A. (1988). The case of unaccusatives. *Linguistic Inquiry*, 19(1), 1–34.
- Belletti, A. (1999). Italian/Romance clitics: Structure and derivation. In H. Van Riemsdijk (Ed.), *Clitics in the Languages of Europe* (pp. 543–580). Berlin: Mouton de Gruyter.
- Belletti, A. (2001). Inversion as focalization. In *Subject Inversion in Romance and the Theory of Universal Grammar* (pp. 60–90). Oxford: Oxford University Press.
- Belletti, A. (2004). Aspects of the Low IP Area. In L. Rizzi (Ed.), *The Structure of CP and IP: the cartography of syntactic structures. Volume 2*. (pp. 16–51). New York:

Oxford University Press.

- Benincà, P. (1984). Un'ipotesi sulla sintassi delle lingue romanze medievali. *Quaderni Patavini Di Linguistica*, (4), 3–19.
- Benincà, P. (2004). The left periphery of Medieval Romance. *Studi Linguistici e Filologici Online*, 2, 243—297. Retrieved from <http://www.doaj.org/abstract?id=106082&toc=y>
- Benincà, P. (2006). A Detailed Map of the Left Periphery of Medieval Romance. *Crosslinguistic Research in Syntax and Semantics Negation Tense and Clausal Architecture*, 2, 53–86. Retrieved from <http://www.doaj.org/abstract?id=106082&toc=y>
- Benincà, P., & Poletto, C. (2004a). A Case of Do-Support in Romance. *Natural Language & Linguistic Theory*, 22(1), 51–94.
- Benincà, P., & Poletto, C. (2004b). Topic, focus, and V2. In L. Rizzi (Ed.), *The Structure of CP and IP* (pp. 52–75). Oxford: Oxford University Press.
- Berman, R. A. (2015). Language development and use beyond the sentence. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge Handbook of Child Language* (2nd editio). Cambridge: Cambridge University Press.
<https://doi.org/https://doi.org/10.1017/CBO9781316095829.021>
- Biberauer, T., Holmberg, A., Sheehan, M., & Roberts, I. G. (2010). *Parametric Variation Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.
- Biberauer, T., & Roberts, I. (2008). Subjects, Tense and Verb Movement in Germanic and Romance. *Cambridge Occasional Papers in Linguistics*, 3, 24–43.
- Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language*, 83–96.
- Bonet, E. (1990). Subjects in Catalan. *MIT Working Papers in Linguistics*, 13, 1–26.
- Borer, H. (1980). Empty subjects in Modern Hebrew and constraints on thematic relations. *Proceedings of the North Eastern Linguistic Society*, 10.
- Borer, H. (2010). Locales. In M. Rappaport-Hovav & I. Sichel (Eds.), *Papers in honor of Anita Mittwoch* (pp. 309–337). Oxford: Oxford University Press.
- Borik, O. & Gehrke, B. (2016). The Syntax and Semantics of Pseudo-Incorporation. *Syntax and Semantics*, Volume: 40. Leiden; Boston: Brill.

- Bossong, G. (2006). La sintaxis de las Glosas Emilianenses en una perspectiva tipológica. In J. L. Girón Alconchel & J. J. de Bustos Tovar (Eds.), *Actas del VI Congreso Internacional de Historia de la Lengua española (Madrid, 29 de septiembre-3 octubre 2003)* (pp. 529–544). Madrid: Arco/Libros.
- Bouzouita, M. (2008). At the syntax-pragmatics interface: clitics in the history of Spanish. *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, 221–263.
- Bouzouita, M. (2014). Left Dislocation phenomena in Old Spanish: an examination of their structural properties. In *Left sentence peripheries in Spanish: diachronic, variationist and comparative perspectives*, (pp. 23–51). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Bruguera, J. (1991). *El llibre dels fets del Rei en Jaume*. Barcelona: Barcino.
- Bruguera, J. (1998). *Llibre dels Fets*. Barcelona: Proa.
- Bruguera, J. (2012). Prioritats textuais i filiació dialectal del ‘Llibre dels fets’ de Jaume I. In G. Colón I Domènech (Ed.), *El Llibre dels feits. Aproximació crítica. Col·lecció Actes*. (pp. 37–46). València: Publicacions de l’Acadèmia Valenciana de la Llengua.
- Büring, D. (1997). *The Meaning of Topic and Focus : The 59th Street Bridge Accent*. London: Routledge.
- Büring, D. (2003). On D-trees, beans, and B-accent. *Linguistics and Philosophy*, 26(5), 511–545.
- Bybee, J. L. (2001). Main clauses are innovative, subordinate clauses are conservative. Consequences for the nature of constructions. *Complex Sentences in Grammar and Discourse. Essays in Honor of Sandra A. Thompson.*, 1–19.
<https://doi.org/10.1075/z.110>
- Calabrese, A. (1992). Calabrese, Andrea. In S. Kuno & H. Thräinsson (Eds.), *Harvard Working Papers in Linguistics* (Vol. 1, pp. 91–127).
- Cardinaletti, A. (2004). Towards a Cartography of Subject Positions. In L. Rizzi (Ed.), *The Structure of CP and IP: the cartography of syntactic structures. Volume 2*. (pp. 115–165). Oxford: Oxford University Press.
- Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case-study

- of the three classes of pronouns. In H. van Riemsdijk (Ed.), *Clitics in the languages of Europe* (pp. 145–233). Berlin: Mouton de Gruyter.
- Casanova, E. (2012). La variació lèxica del ‘Llibre dels feits’ en el context de l’època. *El Llibre Dels Feits. Aproximació Crítica. Col·lecció Actes.*, 10, 85–100.
- Castellà Lidon, J. M. (2002). *La complexitat lingüística en el discurs oral i escrit: densitat lèxica, composició oracional i connexió textual*. Universitat Pompeu Fabra.
- Castellà Lidon, J. M. (2004). Llengua oral i sintaxi. *Articles de Didàctica de La Llengua i de La Literatura*, 33(April), 27–39.
- Castillo Lluch, M., & López Izquierdo, M. (Eds.). (2010). *Modelos latinos en la Castilla medieval*. Madrid: Iberoamericana Vervuert.
- Cecchetto, C. (1999). A comparative analysis of left and right dislocation in Romance. *Studia Linguistica*, 53, 40–67.
- Chafe, W. L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In *Subject and Topic* (pp. 25–55). New York: Academic Press.
- Cinque, G. (1977). The movement nature of left dislocation. *Linguistic Inquiry*, 8(2), 397–412.
- Cinque, G. (1997). “Topic” Constructions in Some European Languages and “Connectedness.” In E. Anagnostopoulou, H. van Riemsdijk, & F. Zwarts (Eds.), *Materians on Left Dislocation* (pp. 93–118). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Cinque, G. (1999). *Adverbs and Functional Heads: a Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Cinque, G., & Rizzi, L. (2010). *Mapping Spatial PPs. The Cartography of Syntactic Structures, Volume 6*. Oxford: Oxford University Press.
- Clackson, J. (2007). *Indo-European linguistics : an introduction*. Cambridge: Cambridge University Press.
- Costa, J. (2001). Spec, IP ou deslocado?: Prós e contras das duas análises dos sujeitos pré-verbais. *DELTA*, 17(2), 283–303.
- Costa, J. (2008). *Subject Positions and Interfaces. The Case of European Portuguese*. Berlin; New York: Mouton de Gruyter.

- Cuenca, M. J. (2008). Els connectors textuais i les interjeccions. In J. Solà, M.-R. Lloret, J. Mascaró, & M. Pérez Saldanya (Eds.), *Gramàtica del Català Contemporani, III* (pp. 3173–3238). Barcelona: Editorial Empúries.
- D'Alessandro, R. (2014). The null subject parameter: Where are we and where are we headed. *Leiden University Centre for Linguistics*. [https://doi.org/10.1016/0024-3841\(91\)90031-Y](https://doi.org/10.1016/0024-3841(91)90031-Y)
- De Cat, C. (2007). French dislocation without movement. *Natural Language & Linguistic Theory*, 25(3), 485–534.
- de Jong, I. (2015). *Narratology and Classics. A Practical Guide*. Oxford: Oxford University Press.
- Devine, A. M., & Stephens, L. D. (2006). *Latin Word Order* (1st ed.). Oxford: Oxford University Press.
- Diesing, M. (1990). Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory*, 8, 41–79.
- Donaldson, B. (2016). Preverbal subjects, information structure, and object clitic position in Old Occitan. *Journal of Linguistics*, 52(01), 37–69. <https://doi.org/10.1017/S0022226714000619>
- Dorian, N. C. (2002). Diglossia and the simplification of linguistic space. *International Journal of the Sociology of Language*, 2002(157), 63–69. <https://doi.org/10.1515/ijsl.2002.042>
- Dretske, F. (1972). Contrastive statements. *Philosophical Review*, 81(4), 411.437.
- Eide, K., & Sitaridou, I. (2014). Contrastivity and information structure in the old Ibero-Romance languages. In K. Bech & K. Gunn Eide (Eds.), *Information Structure and Syntactic Change in Germanic and Romance languages* (pp. 377–412). Amsterdam: John Benjamins Publishing Company.
- Elspaß, S. (2007). A twofold view ‘from below’: New perspectives on language histories and language historiographies. In S. Elspaß, N. Langer, J. Scharloth, & W. Vandenbussche (Eds.), *Germanic language histories ‘from below’ (1700-2000)*. *Studia Linguistica Germanica* 86. (pp. 3–9). Berlin; New York: Mouton de Gruyter.
- Elvira, J. (2015). Stylistic Fronting in Old Spanish Texts. In M. Bouzouita, I. Sitaridou, &

- E. Pato (Eds.), *Studies in Historical Ibero-Romance Morphosyntax, Issues in Hispanic and Lusophone Linguistics series (IHLL)*. Amsterdam: John Benjamins Publishing Company.
- Ernst, T. (2009). Speaker-Oriented Adverbs. *Natural Language & Linguistic Theory*, 27(3), 497–544. <https://doi.org/10.1007/sl>
- Erteschik-Shir, N. (1997). *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- Farkas, D. F., & Bruce, K. B. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, 27(1), 81–118. <https://doi.org/10.1093/jos/ffp010>
- Fernández-Ordóñez, I. (2008). Orden de palabra, tópicos y focos en la prosa alfonsí. *Revista de Estudios Alfonsíes*, (6), 139–172.
- Fernández-Sánchez, J. (2017). *Right dislocation as a biclausal phenomenon: evidence from the Romance languages*. Universitat Autònoma de Barcelona.
- Ferrando Francès, A. (2012). Interès dels manuscrits C i D del “Llibre dels Fets” per a la fixació textual de la crònica jaumina. In *El Llibre dels Feits. Aproximació crítica. Col·lecció Actes*. (pp. 47–84). València: Publicacions de l’Acadèmia Valenciana de la Llengua.
- Fischer, S. (2002). *Catalan clitic system : a diachronic perspective on its syntax and phonology*. Berlin; New York: Mouton de Gruyter.
- Fischer, S. (2005). Construcciones con avance estilístico en el catalán antiguo. *Caplletra*, (38), 119–135.
- Fleischman, S. (1991). Discourse pragmatics and the grammar of Old French: a functional reinterpretation of *si* and the personal pronouns. *Romance Philology*, 44, 251–283.
- Fontana, J. M. (1993). *Phrase Structure and the Syntax of Clitics in the History of Spanish*. University of Pennsylvania.
- Fontana, J. M. (1997). On the integration of second position phenomena. In A. van Kemende & N. Vincent (Eds.), *Parameters of morphosyntactic change* (pp. 207–249). Cambridge: Cambridge University Press. Retrieved from <papers2://publication/uuid/4C1227C4-CE4D-43D8-9903-169ED66ABBBC>
- Forcadell, M. (2013). Subject informational status and word order: Catalan as an SVO

- language. *Journal of Pragmatics*, 53, 39–63.
<https://doi.org/10.1016/j.pragma.2013.03.017>
- Forcadell, M., & Llopis, J. (2014). Structural change in Catalan discourse. *Studies in Language*, 38(2), 237–274. <https://doi.org/10.1075/sl.38.2.01for>
- Frascarelli, M. (2000). *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian. Studies in Natural Language and Linguistic Theory 50. Natural Language and Linguistic Theory* (Vol. 50). Dordrecht: Kluwer Academic Publishers.
- Frascarelli, M. (2004). Dislocation, Clitic Resumption and Minimality: A Comparative Analysis of Left and Right Topic Constructions in Italian. In R. Bok-Bennema, B. Bart Hollebrandse, M. Kampers, & P. Sleeman (Eds.), *Romance Languages and Linguistic Theory 2002* (pp. 99–118). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Frascarelli, M., & Hinterhölzl, R. (2007). Types of topics in German and Italian. In S. Winkler & W. Schwabe (Eds.), *On Information Structure, Meaning, and Form* (p. 87–116). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Fraser, B. (1998). Contrastive discourse markers in English. In A. Jucker & Y. Ziv (Eds.), *Discourse Markers: Description and Theory* (pp. 301–26). Amsterdam; Philadelphia: Benjamins.
- Fraser, B. (2009). An Account of Discourse Markers. *International Review of Pragmatics*, 1(2), 293–320.
- Gallego, Á. J. (2013). Object shift in Romance. *Natural Language and Linguistic Theory*, 31(2), 409–451. <https://doi.org/10.1007/s11049-013-9188-6>
- García Cornejo, R. (2006a). A vueltas con las construcciones que...que: el llamado que pleonástico. In J. L. Girón Alconchel & J. J. de Bustos Tovar (Eds.), *Actas del VI Congreso Internacional de Historia de la Lengua española : Madrid, 29 de septiembre-3 octubre 2003, Vol. 1* (pp. 715–728).
- García Cornejo, R. (2006b). Otra vez sobre el que español. *Actas Del V Congreso Andaluz de Lingüística General: Homenaje Al Profesor José Andrés de Molina Redondo*, 2, 593–608.
- Gärtner, H. M. (2001). Are there V2 relative clauses in German? *Journal of Comparative Germanic Linguistics*, 3, 97–141.

- Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins Publishing Company.
- Giannakidou, A. (1999). Affective dependencies. *Linguistics and Philosophy*, 22, 367–421.
- Giannakidou, A. (2001). Varieties of polarity items and the (non)veridicality hypothesis. In J. Hoeksema, H. Rullman, V. Sánchez-Vanencia, & T. van der Wouden (Eds.), *Perspectives on negation and polarity items*. Amsterdam: John Benjamins Publishing Company.
- Giannakidou, A. (2008). Negative and positive polarity items: Variation, licensing, and compositionality. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.
- González i Planas, F. (2010). *La duplicació del complementador en les llengües iberoromàniques*. Universitat de Girona.
- González i Planas, F. (2014). On quotative recomplementation: Between pragmatics and morphosyntax. *Lingua*, 146, 39–74.
- Gundel, J. (2012). Information Structure. In *The Encyclopedia of Applied Linguistics*. [https://doi.org/https://doi.org/10.1002/9781405198431.wbeal0537](https://doi.org/10.1002/9781405198431.wbeal0537)
- Gundel, J. K., & Fretheim, T. (2004). Topic and Focus. In L. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 175–196). Oxford: Blackwell.
- Gupton, T. (2010). *The Syntax Information Structure Interface: Subjects and Clausal Word Order in Galician*. University of Iowa.
- Haegeman, L. (2003). Conditional clauses: External and internal syntax. *Mind & Language*, 18(4), 317–339. <https://doi.org/10.1111/1468-0017.00230>
- Haegeman, L. (2009). The movement analysis of temporal adverbial clauses. *English Language and Linguistics*, 13(October 2009), 385. <https://doi.org/10.1017/S1360674309990165>
- Haegeman, L. (2010). The internal syntax of adverbial clauses. *Lingua*, 120, 628–648. <https://doi.org/10.1016/j.lingua.2008.07.007>
- Haegeman, L. (2014). West Flemish verb-based discourse markers and the articulation of the speech act layer. *Studia Linguistica*, 68(1), 116–139.

<https://doi.org/10.1111/stul.12023>

Haumann, D. (2007). *Adverb Licensing and Clause Structure in English*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Hernanz, M. L., & Rigau, G. (2006). Variación dialectal y periferia izquierda. In B. Fernández & I. Laka (Eds.), *Andolin gogoan. Essays in Honour of Professor Eguzkitza* (pp. 435–452). Bilbao: Euskal Herriko Unibertsitateko Argitalpen Zerbitzua.

Higgins, R. (1988). *Where the Old English sentence begins*. University of Massachusetts, Amherst.

Hinzelin, M.-O. (2006). Marc-Olivier Hinzelin (Konstanz) Die neutralen Pronomina « ell », « això », « allò », « açò » und « ço » in unpersönlichen Konstruktionen in diachroner und dialektaler Variation. In C. Pusch (Ed.), *La gramàtica pronominal del català: variació - evolució - funció. The Grammar of Catalan Pronouns: Variation - Evolution - Function*. (pp. 47–84). Aachen: Shaker.

Hinzelin, M.-O. (2010). Catalan and Spanish Neuter Pronouns in Impersonal Constructions: Diachronic and Dialectal Variation. *Dialectologia*, 4, 23–45.

Holmberg, A. (2001). The syntax of yes and no in Finnish, 55(2), 141–175.
<https://doi.org/10.1093/acprof:oso/9780198701859.001.0001>

Holmberg, A. (2005). Is there a Little Pro? Evidence from Finnish? *Technology*, 36(4), 533–564.

Holmberg, A. (2006). Stylistic Fronting. In M. Everaert & H. van Riemsdijk (Eds.) (Vol. 1, pp. 532–565). London: Blackwell Publishing.

Holmberg, A. (2015). Verb Second. In T. Kiss & A. Alexiadou (Eds.), *Syntax - Theory and Analysis : An International Handbook*. (pp. 342–383). Berlin; Boston: Mouton de Gruyter.

Holmberg, A. (2016). *The Syntax of Yes and No*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198701859.001.0001>

Hooper, P. J., & Thompson, S. A. (1973). On the applicability of root transformations. *Linguistic Inquiry*, 4, 465–497.

Hsu, B. (2017). Verb second and its deviations: An argument for Feature Scattering in the

- left periphery. *Glossa: A Journal of General Linguistics*, 2(1), 35. 1-3.
[https://doi.org/10.1016/0025-326X\(78\)90261-8](https://doi.org/10.1016/0025-326X(78)90261-8)
- Hulk, A., & Pollock, J.-Y. (2001). *Subject Inversion in Romance and the Theory of Universal Grammar. Subject inversion in Romance and the theory of Universal Grammar*. Retrieved from <http://books.google.fr/books?id=RawWTnEg49UC>
- Jensen, F. (1994). *Syntaxe de l'ancien occitan*. Tübingen: Max Niemeyer Verlag.
- Kaiser, G. A. (2002). *Verbstellung Und Verbstellungswandel in Den Romanischen Sprachen*. Tübingen: Max Niemeyer Verlag.
- Kaiser, G. A. & Zimmermann, M. (2011). On the decrease in subject-verb inversion in French declaratives. In E. Rinke & Kupisch, T. *The Development of Grammar. Language Acquisition and Diachronic Change. In honour of Jürgen Meisel* (pp. 354-382). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Kiss, K. É. (1998). Identificational focus versus information focus. *Language*.
<https://doi.org/10.1353/lan.1998.0211>
- Koenenman, O., & Neeleman, A. (2001). Predication, verb movement and the distribution of expletives. *Lingua*, 111(3), 189–233. [https://doi.org/10.1016/S0024-3841\(00\)00030-9](https://doi.org/10.1016/S0024-3841(00)00030-9)
- Kroch, A. (1989). Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change*.
- Kunert, H. P. (2003). L'ordine degli elementi della frase in occitano antico. *Romanische Forschungen*, 115(2), 194–209.
- Labelle, M. (2007). Clausal architecture in early Old French. *Lingua*, 117, 289–316.
- Lafont, R. (1967). *La phrase occitan : essai d'analyse systématique*. Paris: Presses Universitaires de France.
- Lahousse, K. (2011). *Quand passent les cigognes. Le sujet nominal postverbal en français moderne*. Paris: Presses Universitaires de Vincennes.
- Ledgeway, A. (2005). Moving through the Left Periphery: The Dual Complementiser System in the Dialects of Southern Italy. *Transactions of the Philological Society*, 103(3), 339–96. <https://doi.org/10.1111/j.1467-968X.2005.00157.x>
- Ledgeway, A. (2008). Satisfying V2 in early Romance: Merge vs. Move. *Journal of*

- Linguistics*, 44(02), 437–470. <https://doi.org/10.1017/S0022226708005173>
- Leonetti, M. (2010). La expresión de la estructura informativa en la sintaxis: un parámetro de variación en las lenguas románicas. *Romanistisches Jahrbuch*, 61, 338–355.
- Leonetti, M. (2012). Spanish VSX. In K. Lahousse & M. Marzo (Eds.), *Romance Languages and Linguistic Theory* (pp. 37–63). Amsterdam: John Benjamins Publishing Company.
- Leonetti, M. (2015). On Word Order in Spanish Copular Sentences. In I. Pérez Jiménez, M. Leonetti, & S. Gumiel (Eds.), *New Perspectives in the Study of Ser & Estar* (pp. 203–235). Amsterdam: John Benjamins Publishing Company.
- Leonetti, M. (2017). Basic constituent orders. In A. Dufter & E. Stark (Eds.) *Manual of Romance Morphosyntax and Syntax*. *Manuals of Romance Linguistics* 17 (pp. 887–932). Berlin; Boston: Mouton de Gruyter.
- Leonetti, M., & Escandell-Vidal, M. V. (2009a). Fronting and Verum-Focus in Romance. In *Focus and Background in Romance Languages*.
- Leonetti, M., & Escandell-Vidal, M. V. (2009b). La expresión del Verum Focus en español. *Español Actual*, 92, 11–46.
- López, L. (2009). *A Derivational Syntax of Information Structure*. Oxford: Oxford University Press.
- Maling, J. (1980). Inversion in Embedded Clauses in Modern Icelandic. *Íslenskt Mál Og Almenn Málfræði*, 2, 75–93.
- Marchello-Nizia, C. (1985). *Dire le vrai : l'adverbe "si" en français médiéval : essai de linguistique historique*. Genève: Droz.
- Martí i Castell, J. (2002). *Estudi Lingüístic dels Usatges de Barcelona. El codi a mitjan segle XII*. Barcelona: Curial Edicions Catalanes. Publicacions de l'Abadia de Montserrat.
- Martins, A. M. (1994). *Clíticos na história do português*. Universidade de Lisboa.
- Martins, A. M. (2004). Emphatic Affirmation and Polarity. Contrasting European Portuguese with Brazilian Portuguese, Spanish, Catalan, and Galician. In J. Doetjies & P. González (Eds.), *Romance Languages and Linguistic Theory* (pp. 197–223). Amsterdam; Philadelphia: John Benjamins Publishing Company.

- Martins, A. M. (2005). Clitic Placement, VP-Ellipsis, and Scrambling in Romance. In *Grammaticalization and Parametric Variation* (pp. 175–193). Oxford: Oxford University Press.
- Martins, A. M. (2011). Scrambling and information focus in old and contemporary Portuguese. *Catalan Journal of Linguistics*, 10(1), 133–158.
- Martins, A. M. (2013). Emphatic polarity in European Portuguese and beyond. *Lingua*, 128, 95–123. <https://doi.org/10.1016/j.lingua.2012.11.002>
- Mascarenhas, S. (2015). *Complementizer doubling in European Portuguese*. University of Oxford.
- Mathieu, E. (2006). Quirky subjects in old French. *Studia Linguistica*, 60(3), 282–312. <https://doi.org/10.1111/j.1467-9582.2006.00128.x>
- Menéndez Pidal, R. (1908). *Cantar de mio Cid: Texto, gramática y vocabulario*. Madrid: Imprenta de Bailly-Bailliere e hijos.
- Molnár, V. (2002). Contrast - from a contrastive perspective. In *Symposium on Information Structure in a Cross-Linguistic Perspective* (pp. 147–161). Rodopi.
- Munaro, N. (2005). Computational puzzles of conditional clause preposing. In A. M. Di Sciulo (Ed.), *UG and external systems: Language, brain and computation* (pp. 73–94). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Munaro, N. (2010). Toward a hierarchy of clause types. In *Mapping the Left Periphery* (pp. 125–162). Oxford: Oxford University Press.
- Munby, J. (1978). *Communicative syllabus design : a sociolinguistic model for defining the content of purpose-specific language programmes*. London: Cambridge University Press.
- Mussafia, A. (1888). Enclisi o proclisi del pronome personale atono quale oggetto. *Romania*, 27, 145–146.
- Myers-Scotton, C., & Jake, J. (2009). A universal model of code-switching and bilingual language processing and production. *The Cambridge Handbook of Linguistic Code-Switching*, 336–357.
- Myers-Scotton, C. M., & Jake, J. L. (2017). Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism*, 21(3),

340–366. <https://doi.org/10.1177/1367006915626588>

- Octavio de Toledo y Huerta, A. (2015). Futuros que se miran el ombligo: mesoclisys y anteposición de formas no personales en la historia del español. In M. Castillo Lluch & M. López Izquierdo (Eds.), *El orden de palabras en la historia del español y otras lenguas iberorromances* (pp. 141–233). Madrid: Visor Libros.
- Ordóñez, F., & Treviño, E. (1999). Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua*, 107(1–2), 39–68. [https://doi.org/10.1016/S0024-3841\(98\)00020-5](https://doi.org/10.1016/S0024-3841(98)00020-5)
- Ott, D. (2015). Connectivity in left-dislocation and the composition of the left periphery. *Linguistic Variation*, 15(2), 225–290.
- Paoli, S. (2003). *Comp and the Left Periphery: Comparative Evidence from Romance*. University of Manchester.
- Paoli, S. (2006). The fine structure of the left periphery: COMPs and subjects: Evidence from Romance. *Lingua*, 117, 057–1079. <https://doi.org/10.1016/j.lingua.2006.05.007>
- Par, A. (1923). Sintaxi catalana, segons los escrits en prosa de Bernat Metge (1398). Halle (Saale): M. Niemeyer,. Retrieved from <http://search.lib.cam.ac.uk/?itemid=%7Ccambrdgedb%7C3724808> CN - PC3871 .P3
- Partridge, S., & Kwakkel, E. (Eds.). (2012). *Medieval authorship in theory and practice*. Toronto: University of Toronto Press.
- Petrova, S., & Hinterhölzl, R. (2010). From V1 to V2 in West Germanic. *Lingua*, 120(2), 315–328. <https://doi.org/10.1016/j.lingua.2008.10.007>
- Pinto, M. (1994). Subjects in Italian: Distribution and interpretation. *Linguistics in the Netherlands*, 11(1), 175–186. <https://doi.org/10.1075/avt.11.18pin>
- Pinto, M. (1997). *Licensing and interpretation of inverted subjects in Italian*. University of Utrecht.
- Poletto, C. (2002). The left-periphery of V2-Rhaetoromance dialects: A new view on V2 and V3. In S. Barbiers, L. Cornips, & S. Van der Kleij (Eds.), *Syntactic Microvariation*.
- Poletto, C. (2005). Sì and e as CP expletives in Old Italian. In M. Batllori, M. L. Hernanz,

- C. Picallo, & F. Roca (Eds.), *Grammaticalization and Parametric Variation* (pp. 206–235). Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199272129.003.0013>
- Poletto, C. (2014). *Word Order in Old Italian*. Oxford: Oxford University Press.
- Poole, G. (2013). Interpolation, verb-second, and the low left periphery periphery in Old Spanish. *International Journal of Theoretical Linguistics*, 5(1), 69–98.
- Potsdam, E. (1999). A Syntax for Adverbs. In E. van Geldern & V. Samii (Eds.), *Proceedings of WECOL 98* (pp. 397–411). Fresno CA: California State University.
- Prince, E. (1988). Discourse analysis: A part of the study of linguistic competence. In F. Newmeyer (Ed.), *Linguistics: the Cambridge Survey* (pp. 164–182). Cambridge: Cambridge University Press.
- Pujol i Campeny, A. (2014). *Long Head Movement in Old Catalan*. University of Cambridge.
- Pujol i Campeny, A. (2017). Verb position in Old Catalan. *Diachronic Generative Syntax 19*.
- Pujol i Campeny, A. (2018). “E sí la hoïren tots”: sí and positive polarity in Old Catalan. *Manuscript Submitted for Review*.
- Pujol i Campeny, A. (2019). Verb First Clauses in Old Catalan. In A. Breitbarth, M. Bouzouita, M. Farasyn, & L. Danckaert (Eds.), *The Determinants of Diachronic Stability*. Amsterdam: John Benjamins Publishing Company.
- Pujol i Campeny, A., & Sitaridou, I. (2017). *Stylistic variation and word order in Old Occitan*. University of Cambridge.
- Reinhard, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27, 53–94.
- Ribeiro, I. (1995). Evidence for a verb-second phase in Old Portuguese. In A. Battye & I. Roberts (Eds.), *Clause Structure and Language Change* (pp. 110–139). Oxford: Oxford University Press.
- Riquer, M. (1964). *Història de la literatura catalana*. Barcelona: Ariel.
- Rivero, M. L. (1993). Long Head Movement vs. V2, and null subjects in Old Romance. *Lingua*, 89, 217–245.

- Rizzi, L. (1982). *Issues in Italian Syntax*. Berlin; Boston: Mouton de Gruyter.
- Rizzi, L. (1997). The Fine Structure of the Left Periphery. In L. Haegeman (Ed.), *Elements of Grammar* (1997th ed., pp. 281–337). Dordrecht, London: Kluwer Academic Publishers.
- Rizzi, L. (2015). Notes on Labeling and Subject Positions. In C. Hamann & S. Matteini (Eds.) *Structures, Strategies and Beyond. Studies in Honour of Adriana Belletti*. Amsterdam: John Benjamins Publishing Company.
- Rizzi, L. (2018). Subjects, Topics and the interpretation of *pro*. In R. Petrosino, P. Cerrone & H. van der Hulst (Eds.), *From Sounds to Structures: Beyond the Veil of Maya. Studies of Generative Grammar (SGG)* (pp. 510-530). Berlin; Boston: Mouton de Gruyter.
- Rizzi, L. & Shlonsky, U. (2006). Satisfying the Subject Criterion by a Non Subject: English Locative Inversion and Heavy NP Shift. In M. Frascarelli (Ed.) *Phases of Interpretation* (pp. 341-362). Berlin: Mouton de Gruyter.
- Roberts, I. (1993). *Verbs and Diachronic Syntax. A Comparative History of English and French*. Dordrecht: Kluwer Academic Publishers.
- Roberts, I. (2010). A deletion analysis of null subjects. In *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 58–87).
- Roberts, I., & Roussou, A. (2003). *Syntactic change: a minimalist approach to grammaticalisation*. Cambridge: Cambridge University Press.
- Rodríguez Molina, J. (2010). *La gramaticalización de los tiempos compuestos en español antiguo. Cinco cambios diacrónicos*. Universidad Autónoma de Madrid.
- Rodríguez Molina, J. (2014). La gramática oculta de la polaridad positiva en español antiguo. *Revista de Filología Hispánica*, 30(3), 861–915.
- Rögnvaldsson, E., & Thráinsson, H. (1990). On Icelandic Word Order Once More. In *Syntax and Semantics* (Vol. 24, pp. 3–40). New York: Academic Press.
- Romaine, S. (1982). *Socio-Historical Linguistics: its Status and Methodology*. Cambridge: Cambridge University Press.
- Rosselló, J. (2001). A Minimalist Approach to the Null Subject. *Catalan Working Papers in Linguistics*, 8, 97–128.

- Russell-Gebbett, P. (1965). Grievances of Guitart Isarn, Lord of Caboet. In P. Russell Gebbett (Ed.), *Medieval Catalan Linguistic Texts* (p. 74 et seq.).
- Rutten, G. (2016). Diaglossia , individual variation and the limits of standardization : Evidence from Dutch. In C. Russi, K. Grzegorek, & M. Rocchi (Eds.), *Current Trends in Historical Sociolinguistics*. Warsaw: De Gruyter Open.
- Salvesen, C. M. (2013). Topics and the Left Periphery. A comparison of Old French and Modern Germanic. In T. Lohndal (Ed.), *In search of universal grammar : from Old Norse to Zoque* (pp. 131–172). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Salvesen, C., & Walkden, G. (2017). Diagnosing embedded V2 in Old English and Old French. In *Micro-change and macro-change in diachronic syntax (Oxford studies in diachronic and historical linguistics ; 23)* (pp. 168–181). Oxford: Oxford University Press.
- Santorini, B. (1989). *The generalization of the verb-second constraint in the history of Yiddish*. University of Pennsylvania.
- Schiffrin, D., Tannen, D., Hamilton, H. E., Ward, G., & Birner, B. J. (2001). 6 Discourse and Information Structure, (1992).
- Schneider, E. W. (2008). Investigating Variation and Change in Written Documents. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change* (pp. 67–96). Oxford: Blackwell Publishing.
- Serra i Alegre, E., & Prunyoza i Tomàs, M. (2008). La coordinació. In J. Solà, M.-R. Lloret, J. Mascaró, & M. Pérez Saldanya (Eds.), *Gramàtica del Català Contemporani, III* (pp. 2181–2246). Barcelona: Editorial Empúries.
- Sheehan, M. (2005). *The Status of Preverbal Subjects in Romance Languages*. Cambridge.
- Sheehan, M. (2010). ‘Free’ inversion in Romance and the Null Subject Parameter. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric variation: null subjects in minimalist theory* (pp. 231–262). Cambridge: Cambridge University Press.
- Sheehan, M. (2015). *On the lack of consistency in (Romance) consistent null subject languages*. Cambridge: CamCos4.

- Sifaki, E. (2003). *EPP satisfiers; verb-initial orders in Greek*. University of York.
- Sifaki, E., & Tsoulas, G. (2016). *XP-V orders in Greek and the EPP*.
- Sigurdsson, H. Á. (1990). V1 Declaratives and Verb Raising in Icelandic. *Syntax and Semantics*, 24, 41–69.
- Sitaridou, I. (2004). A corpus-based study of null subjects in Old French and Occitan. In K. D. Pusch & J. Kabatek (Eds.), *Romance Corpus Linguistics II. Corpora and Diachronic Linguistics* (pp. 359–374). Gunter Narr Verlag.
- Sitaridou, I. (2009). On the emergence of personal infinitives in the history of Spanish. *Diachronica*, 26(1), 36–64.
- Sitaridou, I. (2011). Word order and information structure in Old Spanish. *Catalan Journal of Linguistics*, (10), 159–184. Retrieved from <http://ddd.uab.cat/record/88146>
- Sitaridou, I. (2012). A Comparative Study of Word Order in Romance. *Folia Linguistica*, 46(2, “The Pace of Grammaticalisation in Romance”), 553–604.
- Sitaridou, I. (2015). La anteposición de participio en español antiguo debida a la estructura de la información del discurso. In M. López Izquierdo & M. Castillo Lluch (Eds.), *El orden de palabras en español medieval* (pp. 111–143). Madrid: Visor.
- Sitaridou, I. (2016). Invited DiGS talk. In A. Breitbarth, M. Bouzouita, M. Farasyn, & L. Danckaert (Eds.), *The Determinants of Diachronic Stability*. Amsterdam: John Benjamins Publishing Company.
- Sitaridou, I. (2019). Word Order in Old Spanish: V2 or not V2? In A. Breitbarth, M. Bouzouita, M. Farasyn, & L. Danckaert (Eds.), *The Determinants of Diachronic Stability*. Amsterdam: John Benjamins Publishing Company.
- Solà, J. (1992). *Agreement and Subjects*. Universitat Autònoma de Barcelona.
- Soldevila, F. (2007). *Les quatre grans cròniques, Vol. 1: Llibre dels Feits del Rei En Jaume*. Barcelona: Institut d’Estudis Catalans.
- Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 315–332). New York: Academic Press.
- Steiner, B. D. J. (2014). *The Evolution of Information Structure and Verb Secpnd in The History of French*. Indiana University.

- Steuk, J. (2016). Complement clauses in conversation. In A. Cuza, L. Czerwionka, & D. Olson (Eds.), *Inquiries in Hispanic Linguistics: From theory to empirical evidence. Issues in Hispanic and Lusophone Linguistics 12* (pp. 73-94). Amsterdam: John Benjamins Publishing Company.
- Störzer, M., & Stolterfoht, B. (2015). Syntactic base positions for adjuncts? Psycholinguistic studies on frame and sentence adverbials. *Questions and Answers in Linguistics*, 1(2), 57–72. <https://doi.org/https://doi.org/10.1515/qal-2015-0004>
- Stowell, T., & Massam, D. (2018). Introducing register variation and syntactic theory. *Linguistic Variation*, 17(2), 149–156. <https://doi.org/10.1075/lv.00001.mas>
- Sturgeon, A. (2008). *The left periphery: the interaction of syntax, pragmatics and prosody in Czech*. Amsterdam: John Benjamins Publishing Company.
- Talmy, L. (1972). *Semantic Structures in English and Atsugewi*. Berkely: University of California.
- Thräinsson, H. (2007). *The syntax of Icelandic*. Cambridge: Cambridge University Press.
- Tobler, A. (1875). De l'ordre des mots dans Chretien de Troyes. *Vermischte Beiträge Zur Französischen Grammatik*, 5, 395–414.
- Torruella, J., Pérez Saldanya, M. & Martines, J. (2009). *Corpus Informatitzat del Català Antic*. <http://lexicon.uab.cat/cica>. Accessed 10/2014-03/2019.
- Tortora, C. (2001). Evidence for a null locative in Italian. In G. Cinque & G. Salvi (Eds.), *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*. (pp. 313–326). Elsevier.
- Uriagereka, J. (1995). An F position in Western Romance. In K. É. Kiss (Ed.), *Discourse Configurational Languages (dies in Comparative Syntax)* (pp. 153–175). Oxford: Oxford University Press.
- Vallduví, E. (1992). Focus Constructions in Catalan. In C. Laeufer & T. A. Morgan (Eds.), *Theoretical Analyses in Romance Linguistics* (pp. 457–479). Amsterdam: John Benjamins Publishing Company.
- Vallduví, E. (1993). Catalan as VOS: evidence from Information Packaging. *Linguistic Perspectives on Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages.*, 103, 335--350.

<https://doi.org/10.1075/cilt.103.31val>

- Vallduví, E. (1994). Detachment in Catalan and information packaging. *Journal of Pragmatics*, 22(6), 573–601. [https://doi.org/10.1016/0378-2166\(94\)90031-0](https://doi.org/10.1016/0378-2166(94)90031-0)
- Vallduví, E. (2008). L'oració com a unitat informativa. In J. Solà, M. R. Lloret, J. Mascaró, & M. Pérez-Saldanya (Eds.), *Gramàtica del Català Contemporani, II* (pp. 1223–1279). Barcelona: Editorial Empúries.
- Vallduví, E., & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics*, 34, 459–519.
- Vallduví, E., & Vilkuna, M. (1998). On rheme and contrast. In *Syntax and Semantics 29: The Limits of Syntax*. (pp. 79–108). San Diego: Academic Press.
- van Reenen, P., & Schlösler, L. (2000). The pragmatic functions of the Old French particles *ainz*, *apres*, *donc*, *lors*, *or*, *puis*, and *si*. In S. C. Herring, L. Schøsler, & P. van Reenen (Eds.), *Textual Parameters in Older Languages* (pp. 59–109). John Benjamins Publishing Company.
- Vance, B. (1993). Verb-first declaratives introduced by *et* and the position of *pro* in old and middle French. *Lingua*, 89(2–3), 281–314. [https://doi.org/10.1016/0024-3841\(93\)90055-2](https://doi.org/10.1016/0024-3841(93)90055-2)
- Vance, B. (1997). *Syntactic change in medieval French : verb-second and null subjects*. Dordrecht, London: Kluwer Academic Publishers.
- Vance, B., Donaldson, B., & Steiner, B. D. (2010). V2 loss in Old French and Old Occitan: The role of fronted clauses. In S. Colina, A. Olarrea, & A. M. Carvalho (Eds.), *Romance Linguistics 2009 - Selected papers from the 39th linguistic symposium on Romance languages (LSRL), Tucson, Arizona, March 2009* (pp. 301–320). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Veny, J. (1983). *Els parlars catalans: síntesi de dialectologia*. Palma de Mallorca: Moll.
- Vikner, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.
- Villa-García, J. (2010). *Recomplementation and locality of movement in Spanish*. University of Connecticut, Storrs.
- Villa-García, J. (2012). *The Spanish Complementizer System: Consequences for the*

- Syntax of Dislocations and Subjects, Locality of Movement, and Clausal Structure.*
University of Connecticut, Storrs.
- Villa-García, J. (2015). *The Syntax of Multiple-que Sentences in Spanish: Along the left Periphery.* Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Wanner, D. (1998). Les subordinées à double complémentateur en roman medieval. In G. Ruffino (Ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza.* Tübingen: Niemeyer.
- Willis, D. (1997). Clausal coordination and the loss of verb-second in Welsh. *Oxford Working Papers in Linguistics, Philology and Phonetics*, 2, 151–172.
- Wolfe, S. (2014). The Old Sardinian Condaghes : A Syntactic Study. *Transactions of the Philological Society*, n/a. <https://doi.org/10.1111/1467-968X.12046>
- Wolfe, S. (2015a). *Microvariation in Medieval Romance Syntax: A Comparative Study.* University of Cambridge.
- Wolfe, S. (2015b). The Nature of Old Spanish Verb Second Reconsidered. *Lingua*, 164, 132–155.
- Wolfe, S. (2015c). Verb-initial orders in Old Romance: a comparative account. *Revue Roumaine de Linguistique*, LX(2–3), 147–172.
- Wolfe, S. (2016). A Comparative Perspective on the Evolution of Romance Clausal Structure. *Diachronica*, 33(4), 461–502.
- Zubizarreta, M. L. (1998). *Prosody, Focus and Word Order.* Cambridge, Massachusetts: MIT Press.
- Zwart, J.-W. (1997). The Germanic SOV Languages and the Universal Base Hypothesis. In L. Haegeman (Ed.), *The New Comparative Syntax* (pp. 246–267). London: Longman.