

Universitat Autònoma de Barcelona
Servei de Biblioteques



1500449932

F. Xavier VILLALBA NICOLÁS

THE SYNTAX OF SENTENCE PERIPHERY

Tesi doctoral dirigida per la Dra. Gemma Rigau i Oliver

Departament de Filologia Catalana

Facultat de Filosofia i Lletres

UNIVERSITAT AUTÒNOMA DE BARCELONA

Any 2000

Table of contents

Acknowledgments	1
Foreword	2
Abstract	4
Resum	4
CHAPTER 1. THEORETICAL FRAMEWORK	5
1.1. The antisymmetry of syntax	5
<i>1.1.1. Linear order</i>	9
<i>1.1.2. Hierarchical structure</i>	10
<i>1.1.3. Phrase structure axioms</i>	15
<i>1.1.4. Linking linear order and hierarchical structure: The Linear Correspondence Axiom</i>	18
1.2. Some consequences	20
<i>1.2.1. Deriving phrase structure assumptions</i>	20
<i>1.2.2. Deriving X' theory</i>	24
<i>1.2.3. Specifiers as adjuncts</i>	27
<i>1.2.4. Against multiple adjunction</i>	29
<i>1.2.5. More on adjunction</i>	31
<i>1.2.6. Restricting the transformational component</i>	36
<i>1.2.7. The emergence of functional categories</i>	40
1.3. Summary	41
CHAPTER 2. A TYPOLOGY OF DETACHMENT PROCESSES IN CATALAN (I): LEFT-DETACHMENTS	43
2.1. Clitic left dislocation	43
<i>2.1.1. Syntactic properties</i>	43
<i>2.1.2. Interpretive properties</i>	60
<i>2.1.3. Summary</i>	80
2.2. Hanging topic left dislocation	80
<i>2.2.1. Syntactic properties</i>	80
<i>2.2.2. Interpretive properties</i>	97
<i>2.2.3. Summary</i>	103
2.3. The <i>as for</i> construction	103
<i>2.3.1. Syntactic properties</i>	106
<i>2.3.2. Interpretive properties</i>	112
<i>2.3.3. Summary</i>	115
2.4. Metalinguistic topics	115
2.5. Conclusions	118
Appendix. Verbal left-detachments	119
A.1. Syntactic properties	120
<i>A.1.1. The detached element</i>	120
<i>A.1.2. The resumptive element</i>	123
<i>A.1.3. Combinatory aspects</i>	125
A.2. Interpretive aspects	126
<i>A.2.1. The detached element is CLLD</i>	126
<i>A.2.2. Explaining some differences</i>	129
A.3. Summary	131

CHAPTER 3. A TYPOLOGY OF DETACHMENT PROCESSES IN CATALAN (II):	133
RIGHT-DETACHMENTS	
3.1. Clitic right dislocation	133
3.1.1. <i>Syntactic properties</i>	134
3.1.2. <i>Interpretive properties</i>	144
3.1.3. <i>Summary</i>	154
3.2. Afterthoughts	155
3.2.1. <i>Syntactic properties</i>	155
3.2.2. <i>Interpretive properties</i>	159
3.2.3. <i>Summary</i>	160
3.3. Conclusions	160
Appendix A: Deaccenting and prosodic movement	161
A.1. English	162
A.2. Spanish	165
A.3. Italian	172
A.4. Conclusions	174
Appendix B. Verbal right-detachments	176
B.1. Syntactic properties	176
B.1.1. <i>The detached element</i>	176
B.1.2. <i>The resumptive element</i>	177
B.1.3. <i>Combinatory aspects</i>	177
B.2. Interpretive aspects	178
CHAPTER 4. CLITIC LEFT DISLOCATION VS. CLITIC RIGHT DISLOCATION:	181
THREE ANALYSES	
4.1. The symmetric analysis	182
4.1.1. <i>Evidence favoring the analysis</i>	183
4.1.1.1. The resumptive strategy and free word order	183
4.1.1.2. Dislocates are opaque domains for extraction	185
4.1.2. <i>Evidence against the analysis</i>	186
4.1.2.1. Boundedness	186
4.1.2.2. Creation of island effects	188
4.1.2.3. Licensing of negative polarity items	189
4.1.2.4. Principle C	190
4.1.2.5. Pronouns bound by a quantifier	191
4.1.2.6. Interactions between CLRD and CLLD	192
4.1.2.7. Non-finite clauses	193
4.1.2.8. A theoretical remark	194
4.1.3. <i>Conclusions</i>	195
4.2. The covert movement analysis	195
4.2.1. <i>Evidence favoring the analysis</i>	196
4.2.1.1. Upward boundedness	196
4.2.1.2. Creation of island effects	197
4.2.1.3. Licensing of negative polarity items	197
4.2.1.4. Theoretical advantages	198
4.2.2. <i>Evidence against the analysis</i>	199
4.2.2.1. Free word order	199
4.2.2.2. Dislocates are opaque domains for extraction	201
4.2.2.3. Principle C	201
4.2.2.4. Pronouns bound by a quantifier	202
4.2.2.5. Interactions between CLRD and CLLD	203
4.2.2.6. Non-finite clauses	204
4.2.2.7. Informational status	205
4.2.3. <i>Conclusions</i>	205

4.3. The double topicalization analysis	205
<i>4.3.1. Evidence favoring the analysis</i>	206
4.3.1.1. Free word order of the dislocates	206
4.3.1.2. Creation of island effects	207
4.3.1.3. Dislocates are opaque domains for extraction	208
<i>4.3.2. Evidence against the analysis</i>	208
4.3.2.1. Licensing of negative polarity items	208
4.3.2.2. Dislocates are opaque domains for extraction	209
4.3.2.3. Principle C	209
4.3.2.4. Pronouns bound by a quantifier	210
4.3.2.5. Interactions between CLRD and CLLD	211
4.3.2.6. Non-finite clauses	212
4.3.2.7. Theoretical motivation	213
<i>4.3.3. Conclusions</i>	214
4.4. Conclusions	215
CHAPTER 5. A NEW ANALYSIS FOR CLRD: THE SPLIT-TOPIC ANALYSIS	217
5.1. The structural issue	219
<i>5.1.1. Topic-splitting</i>	220
5.1.1.1. Empirical arguments	220
5.1.1.2. Theoretical arguments	224
<i>5.1.2. Clausal architecture</i>	227
5.1.2.1. The placement of ExtTopP	227
5.1.2.2. The placement of IntTopP	230
5.2. The movement issue	234
<i>5.2.1. Movement vs. base-generation</i>	234
5.2.1.1. The presence of a resumptive clitic	235
5.2.1.2. The iteration of CLLD	247
5.2.1.3. The lack of WCO effects	249
5.2.1.4. The failure of CLLD in licensing parasitic gaps	253
5.2.1.5. The sensitivity to (strong) islands	255
5.2.1.6. The existence of reconstruction/connectivity effects	261
<i>5.2.2. The trigger for dislocation</i>	264
<i>5.2.3. The landing-site of dislocation</i>	267
5.2.3.1. Adjunction vs. substitution	268
5.2.3.2. A- vs. A'-movement	270
5.3. The derivational issue	275
<i>5.3.1. CLLD derives from CLRD</i>	275
<i>5.3.2. Multiple dislocation</i>	281
5.4. The interface issue	284
<i>5.4.1. Semantic interpretation</i>	285
<i>5.4.2. Prosody</i>	288
5.5. Conclusions	291
REFERENCES	293

Acknowledgments

Finally, the last word is written. Now it is the time to have a quick glance to the walked path, and there I find, as always, Gemma Rigau. This thesis has been very fortunate to grow under her expertise and me too for having received from her an incessant support, advice and friendship, since a freezing and snowy winter in Massachusetts. My deepest gratitude, Gemma.

And, since we are talking about friendship, let me express my debt with three good friends and excellent linguists. Luis López has always been, although the distance, an intelligent and supportive interlocutor, and a source of encouragement and self-embedding e-mail replies, about linguistics and life. Albert Fontich has always been there as a friend and as a linguist, helping me with the most intricate data, which ice-cream could hardly make just a little more sweet, and revealing me that, after all, there is more fun in syntax than meets the eye. Gorka Elordieta has been virtually there for years, before he happened to be here factually, chatting through the wire, mixing friendship and linguistics. My gratitude to all three.

I also want to make explicit my thankfulness to all the people that over the years have contributed to this thesis in one way or another: Laia Amadas, Anna Bartra, Zulema Borràs, Teresa Cabré i Monné, Linda Escobar, Teresa Espinal, Olga Fullana, Anna Gavarró, Richard Kayne, Mireia Llinàs, Louise McNally, Paco Ordóñez, Montse Pascual, Carme Picallo, Paul M. Postal, Josep Quer, Francesc Roca, Joana Rosselló, Jaume Solà, Enric Vallduví, M. Teresa Ynglès, and María Luisa Zubizarreta, and all the members of the Grup de Gramàtica Teòrica of the UAB.

I am also grateful to the audiences of the congresses where several of the ideas in this work were presented: the "Going Romance 1995" (Amsterdam, 12/7-9/1995), the 6th and 9th Colloquia on Generative Grammar (València, 3/27-29/1996 and Bellaterra, 4/7-9/1999), and the "Third Workshop on the Syntax of Central Romance Languages" (Girona, 11/15-16/1996). Thanks are due also to the colleagues of the Euskal Herriko Unibertsitatea for kindly inviting me to present my confuse thoughts to so intelligent an audience.

Finally, let me thank my father, who is no longer here, my mother, who has always been here, and Sílvia for adding so much happiness and love to the worries and bad temper during the writing of this thesis. This thesis is dedicated to you, Sílvia.

[This research was partially supported by DGES: PB96-1199-C04-02; Generalitat de Catalunya: 1999SGR00113; and CREL-9]

FOREWORD

Since the seventies Clitic Left Dislocation has received a lot of attention and has been deeply investigated, so that many of its properties are pretty well known. In this respect, the work by Guglielmo Cinque (Cinque 1977, 1983, 1990) has been particularly influent. However, nothing similar has happened for Clitic Right Dislocation. The proportion of studies handling with Clitic Left Dislocation overwhelmingly outscores that of studies dedicated to Clitic Right Dislocation, so that from the available literature it is even difficult to say that Clitic Right Dislocation exists as a separate phenomenon. Note, for example, that Cinque (1990), a basic reference for any researcher handling with dislocation constructions, doesn't even mention Clitic Right Dislocation; nor does a quite recent Catalan descriptive grammar of almost 900 pages. Among this quite desolate bibliographical landscape, reference to Clitic Right Dislocation is in general a brief mention within a broader analysis of Clitic Left Dislocation.

Nevertheless, with the advent of the revolutionary work by Richard Kayne on the relationship between phrase structure and word order, which flourishes in his highly remarkable *The Antisymmetry of Syntax*, Clitic Right Dislocation has gained a 'superior status', as the increasing –but still small– number of works appearing in the middle nineties demonstrates. The reason for this recent interest in Clitic Right Dislocation is that this construction, just as other constructions involving apparent adjunction/movement to the right, has become a privileged empirical field where the predictions of Kayne's strong hypothesis on the antisymmetry of syntax can receive a proper verification. Were we able to integrate Clitic Right Dislocation into the highly restrictive system devised by Kayne, a very important step would be made toward a more restrictive theory of phrase structure, and consequently of Universal Grammar.

This thesis is devoted to explore the unexplored territory of Clitic Right Dislocation with the instruments that Kayne's hypothesis of the antisymmetry of syntax furnishes us. The great importance of Kayne's work has a direct reflex in the content of Chapter 1, where I present the bare bones of his proposal.

In Chapter 2, I will make an excursus into the typology of left-detachment constructions, in order to clearly establish the properties of Clitic Left Dislocation against other topic constructions, such as Hanging Topic Left Dislocation, the *as for*

construction, or metalinguistic topics. In this quite arduous task the work of Guglielmo Cinque has proven an invaluable help.

Chapter 3 accomplishes the same function with respect to right-detachments. Hence I will offer the basis for a clear distinction between Clitic Right Dislocation and afterthoughts.

With the background of the preceding chapters, in Chapter 4 I will enter into the properties of Clitic Right Dislocation with respect to Clitic Left Dislocation and will review three basic analyses proposed to accounting for them: the Symmetric Analysis, the Covert Movement Analysis, and the Double Topicalization Analysis.

Finally, in Chapter 5 I offer a distinct proposal, the split-topic analysis, which is an attempt to give an answer to the following questions within the strict limits of the framework presented in Chapter 1 and of minimalist guidelines:

- A) *The structural issue*: Which is the position of left and right dislocates in the sentence?
- B) *The movement issue*: How do left and right dislocates reach their respective positions?
- C) *The derivational issue*: How do we get the correct derivation of Clitic Right Dislocation and Clitic Left Dislocation?
- D) *The interface issue*: How do we get the semantic and prosodic interpretation of Clitic Right Dislocation and Clitic Left Dislocation?

To cope with this set of problems, I will propose that right- and left-dislocates involve two different landing sites: the specifier of the Internal Topic Phrase in the left periphery of the vP area for the former, and that of the External Topic Phrase in the left periphery of sentence for the latter. This proposal will be shown to derive very intriguing properties of both constructions together with quite suggestive possibilities of a more theoretical sort.

I will also emphasize that the split-topic analysis should be considered not *an* optimal analysis of Clitic Right Dislocation and Clitic Left Dislocation, rather than the last word on the issue. Indeed, more justice will be done and more profit will be obtained from taking it as a point of departure for a better understanding of the issue.

ABSTRACT

This dissertation deals with dislocation structures, particularly with Clitic Right Dislocation and with Clitic Left Dislocation. The main guideline of the analysis is that the form of syntactic objects is constrained by Kayne's (1994) Linear Correspondence Analysis. This strict framework will lead us to an analysis (Split-Topic Analysis) involving two different landing sites for dislocated elements: the specifier of the Internal Topic Phrase in the left periphery of the νP area for Clitic Right Dislocation, and that of the External Topic Phrase in the left periphery of sentence for Clitic Left dislocation. It will be argued that this approach is more adequate on both conceptual and empirical grounds, since it complies with the highly restrictive view of syntax imposed by Kayne's Linear Correspondence Axiom, and offers a principled explanation of the consistent set of similarities and differences between Clitic Right Dislocation and Clitic Left Dislocation.

Resum

Aquesta tesi tracta de les estructures amb dislocació, en concret de la dislocació a la dreta i de la dislocació a l'esquerra. Com a guia de l'anàlisi, s'assumeix que el Axioma de Correspondència Lineal de Kayne (1994) constreny la forma dels objectes sintàctics. Això ens porta a una anàlisi (l'anàlisi del tòpic dividit) que fa ús de dues posicions diferents per als elements dislocats: l'especificador del Sintagma Tòpic Intern, que és en la perifèria esquerra del S_{ν} , per a la dislocació a la dreta, i l'especificador del Sintagma Tòpic Extern, que és en la perifèria esquerra del SC, per a la dislocació a l'esquerra. S'argumenta que aquesta proposta resulta més adequada, tant des d'un punt de vista conceptual com empíric, ja que s'ajusta als requisits altament restrictius que l'Axioma de Correspondència Lineal de Kayne imposa a la sintaxi i ofereix una explicació fonamentada del conjunt sistemàtic de semblances i diferències que trobem entre la dislocació a la dreta i la dislocació a l'esquerra.

Chapter 1. Theoretical framework

This chapter is intended as a brief introduction to the framework where this thesis will be developed, namely to the hypothesis of the antisymmetry of syntax, as originally conceived in Kayne (1994). However, since this work is often too concise or elusive in different basic questions, my exposition will respect more its gist than its letter. So then, even though I will take as a point of departure Kayne's monograph, avoiding for the ease of fluency making reference to this work continually, I will explicitly point out when the exposition strictly follows Kayne's proposal or rather my own interpretation or development of his ideas.

1.1. The antisymmetry of syntax

Languages vary with respect to word order. For instance, it is a fact that the order of verb and complement is just the opposite in Basque and Catalan:

- (1) a. Jon-ek eskutitz-a bidali zuen.
Jon-ERG letter-ABS send AUX
'Jon send the letter.'
- b. Jon va enviar la carta.
Jon AUX send the letter
'Jon send the letter.'

Basque is a consistent head-final language: noun/adjective/verbal complements precede the noun/adjective/verb, and it has postpositions. On the contrary, Catalan is consistently head initial across categories.

However, word order variation is not restricted to the crosslinguistic field. We also find differences across categories in a given language. Take for example, Dutch. Dutch is verb final (leaving aside V-2 effects), but it has prepositions, and the complementizer always appears at the beginning of the sentence.

Furthermore, not even a given category of a given language is always consistent with respect to linear order. Take for instance PPs in Dutch (the same is true of German):

- (2) a. onder het viaduct
under the viaduct
'under the viaduct'
- b. er onder
it under
'under it'

At least apparently, Dutch prepositions turn into postpositions when the complement is the inanimate pronoun *er* 'it'.

What is the source of such a huge amount of variation? The standard answer in the Principles and Parameters framework founded on Chomsky (1981) has been the resort to directionality parameters. The differences in word order just seen have standardly been explained as the outcome of the different values for the parameter that sets the directionality of branching of complements, be it one directly affecting phrase structure (see Chomsky 1986, Giorgi and Longobardi 1991, among others) or one establishing the direction of the assignment of Case and Theta Role (see Travis 1984). In essence, it is claimed that only the hierarchical structure of phrase structure is preserved across languages and categories, the linear order being just a particular aspect of each language or category. In other words, Universal Grammar (more specifically X-bar Theory) forces head and complement to be both immediately dominated by X', but it leaves their ordering open. The final word order will be a consequence of the interaction of independent principles (namely, the direction of Case and Theta Role assignment).

Kayne (1994) strongly argues against this conception of phrase structure and takes the opposite way. Instead of assuming a parametrized X-bar Theory, he advocates for a universal basic word order: Specifier-Head-Complement. Any departure from this order will be the result of a movement forced by independent reasons (checking of a formal feature, focus-presupposition articulation, etc.). According to this scholar, the strict subjacent Specifier-Head-Complement order is the side effect of a direct and universal relationship between hierarchical and linear relations: precedence in a phrase marker necessarily entails a higher structural position (in fact, it implies asymmetric c-command; we will see it soon). Let us call this proposal the Hypothesis of the Antisymmetry of Syntax (HAS). The details of Kayne's HAS will be worked out in the

Chapter 1. Theoretical framework

This chapter is intended as a brief introduction to the framework where this thesis will be developed, namely to the hypothesis of the antisymmetry of syntax, as originally conceived in Kayne (1994). However, since this work is often too concise or elusive in different basic questions, my exposition will respect more its gist than its letter. So then, even though I will take as a point of departure Kayne's monograph, avoiding for the ease of fluency making reference to this work continually, I will explicitly point out when the exposition strictly follows Kayne's proposal or rather my own interpretation or development of his ideas.

1.1. The antisymmetry of syntax

Languages vary with respect to word order. For instance, it is a fact that the order of verb and complement is just the opposite in Basque and Catalan:

(1) a. Jon-ek eskutitz-a bidali zuen.

Jon-ERG letter-ABS send AUX

'Jon send the letter.'

b. Jon va enviar la carta.

Jon AUX send the letter

'Jon send the letter.'

Basque is a consistent head-final language: noun/adjective/verbal complements precede the noun/adjective/verb, and it has postpositions. On the contrary, Catalan is consistently head initial across categories.

However, word order variation is not restricted to the crosslinguistic field. We also find differences across categories in a given language. Take for example, Dutch. Dutch is verb final (leaving aside V-2 effects), but it has prepositions, and the complementizer always appears at the beginning of the sentence.

Furthermore, not even a given category of a given language is always consistent with respect to linear order. Take for instance PPs in Dutch (the same is true of German):

- (2) a. onder het viaduct
 under the viaduct
 ‘under the viaduct’
- b. er onder
 it under
 ‘under it’

At least apparently, Dutch prepositions turn into postpositions when the complement is the inanimate pronoun *er* ‘it’.

What is the source of such a huge amount of variation? The standard answer in the Principles and Parameters framework founded on Chomsky (1981) has been the resort to directionality parameters. The differences in word order just seen have standardly been explained as the outcome of the different values for the parameter that sets the directionality of branching of complements, be it one directly affecting phrase structure (see Chomsky 1986, Giorgi and Longobardi 1991, among others) or one establishing the direction of the assignment of Case and Theta Role (see Travis 1984). In essence, it is claimed that only the hierarchical structure of phrase structure is preserved across languages and categories, the linear order being just a particular aspect of each language or category. In other words, Universal Grammar (more specifically X-bar Theory) forces head and complement to be both immediately dominated by X’, but it leaves their ordering open. The final word order will be a consequence of the interaction of independent principles (namely, the direction of Case and Theta Role assignment).

Kayne (1994) strongly argues against this conception of phrase structure and takes the opposite way. Instead of assuming a parametrized X-bar Theory, he advocates for a universal basic word order: Specifier-Head-Complement. Any departure from this order will be the result of a movement forced by independent reasons (checking of a formal feature, focus-presupposition articulation, etc.). According to this scholar, the strict subjacent Specifier-Head-Complement order is the side effect of a direct and universal relationship between hierarchical and linear relations: precedence in a phrase marker necessarily entails a higher structural position (in fact, it implies asymmetric c-command; we will see it soon). Let us call this proposal the Hypothesis of the Antisymmetry of Syntax (HAS). The details of Kayne’s HAS will be worked out in the

next sections, but before that let us briefly consider a general picture of the resultant Universal Grammar.

On the one hand, assuming a basic universal Specifier-Head-Complement order is the strongest hypothesis on the properties of phrase structure. We eliminate all parameters dealing with phrase structure ordering, a desirable upshot, and place variation in the way different languages satisfy universal principles.¹ Obviously, since the hypothesis makes phrase structure so restrictive —only the Specifier-Head-Complement order is allowed—, it is to be preferred on conceptual grounds. It might be argued that this restrictiveness in the phrase structure component necessarily increases the irrestrictiveness of another module of grammar, namely the transformational component. It is true that in order to deal with the huge amount of word order variation, movement takes a more important role, since it becomes responsible of any departure from the basic Specifier-Head-Complement order.² However, this argument cannot withstand a close scrutiny. Note that a model using directionality parameters necessarily makes crucial use of movement to account for some surface orders. For instance, the standard model cannot base generate VSO order, which is commonly analyzed by means of leftward V-movement. Hence, we are not comparing a model just relying on phrase structure conditions and another making use of both phrase structure conditions and movement. Both models make use of movement, so then the one resulting from the HAS, which imposes the most severe conditions on phrase structure is to be preferred unless strong empirical evidence points the opposite way. Furthermore, as will become apparent below in 1.2.6, the model with the HAS has a more restrictive transformational component as well.

On the other hand, the model based upon directionality parameters has serious flaws. Firstly, the OVS order is predicted to be as usual as its mirror image SVO and the same is true of VOS with respect to SOV. However, languages having basic OVS or VOS order are quite strange —especially the former. For instance, in a survey of 63

¹ Note that this is more in the spirit of many of the well-established parameters of the Principles and Parameters framework. Take for example, movement of *wh*-elements in Catalan and Japanese. It has been argued that the contrast is a consequence of applying the same movement but in two different levels (at overt syntax in Catalan; at LF in Japanese). In the present case, the different movements giving rise to the final word order would be forced by the necessity to satisfy, for example, case requirements or by pragmatic reasons. The different way of satisfying these principles would be at the basis of (superficial) word order variation.

² The increasing importance of movement strictly correlates with the emergence of functional categories, a basic element for hosting the moved categories. See paragraph 1.2.7.

languages of almost all family groups, Steele (1978) found 30 having basic SOV, 20 having basic SVO, 10 having basic VSO and 3 having basic VOS —neither OVS nor OSV languages were attested.³ In fact, as Cinque (1996: 447) acutely remarks, '[a] well-known anomaly of all theories of syntax in the Sixties, the Seventies and the Eighties was the existence of various (unexpected) left-right asymmetries in the syntax of natural languages, both within single languages, and cross-linguistically.' The account based on directionality parameters makes an implicit claim for symmetry. If the basic difference between OV and VO languages is to be handled by means of a directionality parameter setting the correct order, we do not expect a strong asymmetry beyond the limits of relative word order. Nonetheless, such a claim is known to be false: mirror-image phenomena are extremely more rare than expected. Consider for instance one of Ross' (1967) perceptive insights, namely the upward boundedness of rightward movement in English (and right-branching languages in general). Interestingly enough, nothing in the standard account of word order typology makes us expect that in the mirror-image left branching languages —i.e., in OV languages— right movement be also upward bounded. However, this prediction is far from being true. Something similar might be said of V2 phenomenon. Even though V2 movement is a clearly well-defined process in many languages, we don't have its mirror-image counterpart: there is no attested language requiring the verb to move to the second-from-last position under the same conditions. Obviously, for a theory making use of directionality parameters, the lack of the rightward counterpart of the two leftward movements considered (see Kayne 1994 and Cinque 1996 for other gaps unexpected in a theory of word order based on directionality parameters) must be an accidental gap. Note, however, that the HAS makes the correct prediction that the big number of left-right asymmetries highlighted by syntacticians and typologists might, and indeed must, exist.

Secondly, the model leaves some room for basic word order ambiguity. A superficial head-final language may be the result of assigning theta-role to the right but case to the left across all categories: all complements will base generated to the right, but will surface to the left, since they have to move for case. How can such a language be distinguished from one assigning both case and theta role to the left? Furthermore,

³ In *The Cambridge Encyclopedia of Language* (1997²) the sum of SVO and SOV is attributed to over the 75% of spoken languages and VSO to the 10-15%. Interestingly, this figures nicely parallel the ones in Steele's survey: SOV plus SVO orders represent over a 79% of the sample, VSO over a 15%, and VOS less than a 5%.

take a VSO language. The verb has moved, but which is the basic word order, SVO or SOV? How does the directionality parameter handle these cases?

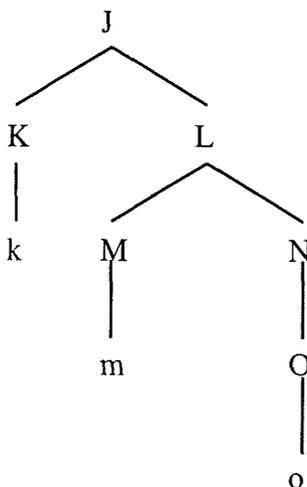
Finally, it is unclear what the reformulation of directionality parameters would be in a framework where case assignment is eliminated in terms of checking in the specifier of functional categories? The resort to establishing the directionality in phrase structure seems the best solution, even if it is a step back.

To sum up, the HAS seems better motivated on conceptual grounds than a model based on directionality parameters, because a more restricted grammar results. Whether it is capable of accounting for the bulk of empirical data is still to be seen, and the brunt of this thesis will be mainly devoted to showing that it is. Nevertheless, before going on in this goal let us work out in more detail Kayne's (1994) proposal.

1.1.1. Linear order

The arrangement of words in discourse necessarily obeys precedence. Hence, when taking into account a phrase marker, it must be the case that its terminal nodes be arranged according to this relation. Consider the following phrase marker, where terminal nodes appear in lowercase for the ease of reference:

(3)



We can define precedence as follows (see Wall 1972):⁴

⁴ Wall's (1972: 140) definition is more general: 'If R is an order, either weak or strict, and $(x, y) \in R$, we say that x precedes y (...).' However, since we are interested in phrase markers, where the precedence relation coexists with other (partial) orders, namely dominance and c-command, the requirement that R be a linear (or total) order is necessary. Note that we restrict the scope of precedence to terminal nodes. This allows us to dispense with special conditions stating that precedence is only significant when the nodes

Precedence =_{def} terminal α precedes terminal β if there is a relation R , such that R is a linear (or total) order and $(\alpha, \beta) \in R$.

Precedence has the following properties:

- 1) it is *irreflexive*: α doesn't precede itself;
- 2) it is *transitive*: if α precedes β and β precedes γ , then α precedes γ ;
- 3) it is *asymmetric*: if α precedes β , then $\alpha \neq \beta$, and β doesn't precede α ;⁵
- 4) it is *total*: if α precedes β and γ precedes β , then either α precedes γ or β precedes γ (or $\alpha = \gamma$).

These four properties define a *strict linear* (or *total*) *order*. This is apparent enough in the previous tree, so I skip the proof. Note that, given the previous properties, the unordered set of terminals $\{k, m, o\}$ can always be converted into an ordered set $\langle k, m, o \rangle$.

1.1.2. Hierarchical structure

Dominance is another basic relation in a phrase marker. We can define a dominance structure as follows (see Grefe & Kracht 1996: 3):

domination structure =_{def} $\langle N, \downarrow \rangle$, where N is a non empty set (of nodes) and \downarrow a binary irreflexive and transitive relation on N .

involved do not stand in the dominance relation (see Wall's 1972 Exclusivity Condition or McCawley's 1982 (3f)).

⁵ Surprisingly enough, even though Kayne (1994: 4) talks of *antisymmetry* instead of *asymmetry*, his definition —not(xLy & yLx)— corresponds to the latter. Obviously, once we assume, as he does (Kayne 1994: 134, note 8), that precedence is irreflexive, it can only be asymmetric, and never antisymmetric. Compare both properties:

Asymmetry: if α precedes β , then $\alpha \neq \beta$, and β doesn't precede α .

Antisymmetry: if α precedes β and β doesn't precede α , then $\alpha = \beta$.

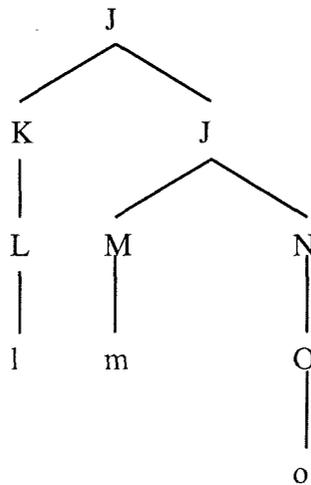
It can be seen at first glance that asymmetry necessarily implies irreflexivity: it cannot be that x precede itself, otherwise it would be possible for a terminal to precede itself, violating asymmetry. However, this is

To put it in tree-based notions (Wall 1972:145):⁶

Dominance =_{def} α dominates β if there is a connected sequence of descending branches in the tree extending from α to β .

However, some refinement is necessary when adjunction structures are involved. Consider the case of J and K in the following phrase marker:

(4)



The crucial point is: does J dominate K? The issue has raised much debate, but there is some consensus in assuming that the answer is no. In the preceding phrase marker, the category J splits into two *segments* and only the higher one dominates K. Once the distinction between categories and segments (proposed in May 1985) is introduced, a refinement of the definition of dominance is necessary to make clear that dominance is a relation between categories, giving segments no role at all.

Dominance =_{def} a category α dominates a category β if there is a connected sequence of descending branches in the tree extending from α to β .

not the case with antisymmetry: for it to be satisfied it must crucially be the case that a terminal precede itself (reflexivity).

⁶ Wall's original definition doesn't include the gerund 'descending', as I do. His definition makes no explicit mention to the direction of the sequence of branches. In fact, Wall presupposes for the reader a theoretical background where sentence structure is constructed from the initial axiom S by means of an up-bottom process. So then, even though my modification technically adds nothing to Wall's definition, I think it makes it clearer for any reader.

Since categories are sets of segments (which in general equal one), it is clear that there is no connected sequence of descending branches extending from category J to K. The only possible relation would be between the higher segment of J and category K, but this is not a relation between two categories. We'll turn back to this issue when discussing adjunction structures (1.2.3.-1.2.5.).

Dominance has the following set of properties:

- 1) it is *irreflexive*: α doesn't dominate itself;⁷
- 2) it is *transitive*: if α dominates β and β dominates γ , then α dominates γ ;
- 3) it is *asymmetric*: if α dominates β , then β doesn't dominate α .⁸

Note that the relation is not total: we can always have at least a pair of nodes that do not dominate nor are dominated by each other. In the preceding phrase marker, it is self-evident that K doesn't dominate J, M, N, nor O; and conversely. As a consequence, the dominance relation cannot be considered a linear ordering: there is no way for this relation to convert the set {J, M, N, O} in an ordered set. Nevertheless, if we reduce our focus to the subset of nodes dominating a given node, the relation turns total. For example, if we focus on the subset formed by node O and the nodes dominating it (J and N), it becomes apparent that dominance can form the ordered set $\langle J, N, O \rangle$. Since the relation is total only in a local domain, we can say that dominance has the following properties:

- 1) it is *irreflexive*: α doesn't dominate itself;
- 2) it is *transitive*: if α dominates β and β dominates γ , then α dominates γ ;
- 3) it is *asymmetric*: if α dominates β , then β doesn't dominate γ ;
- 4) it is *locally total*: for a given node γ dominated by α and β , then either α dominates β or β dominates α (or $\alpha = \beta$).

These properties define a (strict) local linear ordering.

⁷ The irreflexivity of the dominance relation is controversial. As Wall (1972: 145) points out, '[a]s a technical convenience, it is usually assumed that every node dominates itself, i.e. that the dominance relation is reflexive.' Chomsky (1986: fn. 11) and Kayne (1994: 134, fn. 8) explicitly assume that dominance is irreflexive. We will see in 1.1.3, when introducing phrase structure axioms, why this must be so.

Once we put together precedence and dominance, we have the tools to define what a tree is:

Tree =_{def} $\langle T, r, \rightarrow, > \rangle$, where T is a finite set of nodes, r is a node that dominates all nodes, \rightarrow is a strict linear order (i.e. precedence), and $>$ is a strict local linear order (dominance).

On this mathematical configuration, we can define a rather different hierarchical relation: c-command. C-command is based on the notions of *dominance* and *exclusion* (the following definitions are borrowed from Kayne 1994):⁹

C-command =_{def} α c-commands β iff (i) α and β are categories, (ii) α excludes β , and (iii) every category dominating α dominates β .

Exclusion =_{def} α excludes β if no segment of α dominates β .

The definitions become very apparent when we consider again phrase marker (4), repeated here for the ease of reference:

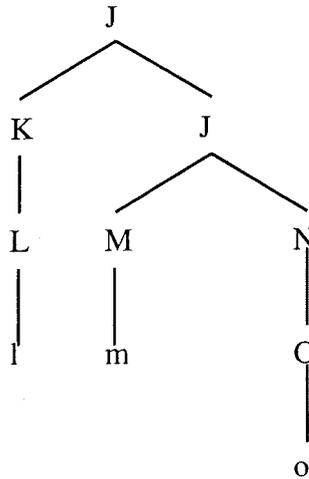
⁸ Again, Kayne (1994: 4) talks of *antisymmetry*. See fn. 5 for discussion.

⁹ It is noteworthy pointing out that the preceding relations hold under the definition of c-command assumed, but a different formulation would establish different relations. Consider, for instance, the definition proposed by Reinhart (1976) which is crucially based on branching nodes:

C-command =_{def} α c-commands β iff the first branching node γ that dominates α either dominates β , or is dominated by a node γ' which dominates β , where γ and γ' are of the same category.

Consider now the c-command relations for phrase marker (4). Crucially, the things radically change for category L and O: since neither K nor N are branching nodes, the first branching node must be J for both L and O, giving as a result that L c-commands J and everything it dominates and O c-commands M. Note as well that, even though the category-segment distinction were assumed, which would prevent O from c-commanding L, L would still c-command J and everything it dominates. As will become apparent later, this is an undesirable consequence for Kayne's hypothesis, so he explicitly rejects Reinhart's formulation of c-command (Kayne 1994: 7).

(5)



K is not dominated by category J, but only by one segment of it, consequently, K c-commands J, M, N, and O. M c-commands N and O. N c-commands M. Note that J doesn't c-command any category because it doesn't exclude them: for any category at least the higher segment of J dominates it.

C-command has the following property:

1) it is *transitive*: if α c-commands β and β c-commands γ , then α c-commands γ .

It is not asymmetric, since two nodes may c-command each other (e.g., M and N in (5)). Furthermore, it is not total either: there are nodes that do not have a c-command relation (e.g. nodes having a dominance relation or nodes like L and O with respect to each other in (5); see the discussion in footnote 9). However, the asymmetric property directly follows if we impose an asymmetric condition on c-command:

Asymmetric c-command =_{def} α asymmetrically c-commands β iff α c-commands β and β doesn't c-command α .

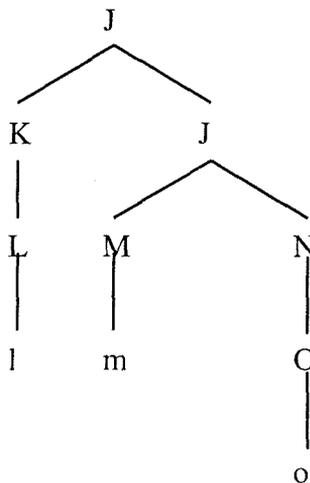
Moreover, asymmetric c-command can also be locally total when we reduce permissible phrase markers to the ones allowing only binary branching. Hence, the properties of asymmetric c-command are:

1) it is *irreflexive*: α doesn't asymmetrically c-command itself;

- 2) it is *transitive*: if α asymmetrically c-commands β and β asymmetrically c-commands γ , then α asymmetrically c-commands γ ;
- 3) it is *asymmetric*: if α asymmetrically c-commands β , then β doesn't asymmetrically c-command α ;
- 4) it is *locally total*: for a given node γ asymmetrically c-commanded by α and β , then either α asymmetrically c-commands β or β asymmetrically c-commands α (or $\alpha = \beta$).

Consider now the asymmetric c-command relations in the previous phrase marker:

(6)



Transitivity is quite apparent: K asymmetrically c-commands M, and M asymmetrically c-commands O, hence K asymmetrically c-commands O. Asymmetry not even needs a comment. Finally, local totality is easy to test: O is asymmetrically c-commanded by both K and M, hence it must be the case that either K asymmetrically c-commands M or vice versa. As we have just seen, K does asymmetrically c-command M. Consequently, the set {O, K, M} becomes the ordered set $\langle K, M, O \rangle$. We can thus say that asymmetric c-command is a (strict) local linear order.

1.1.3. Phrase structure axioms

Until now, we have introduced three basic relations between the nodes of a phrase marker, namely *precedence*, *dominance*, and *asymmetric c-command*. Nevertheless, nothing has been said about the axioms regulating the way of building up syntactic

structures. The first requirement a phrase marker must satisfy by definition is to have a root node (cf. Wall's (1972) Single Root Condition):¹⁰

(7) Axiom 1

Given the set of nodes N forming a phrase marker, $\exists S \in N$ such that $\forall Z \in N, Z \neq S, S$ dominates Z .

Note that since, by definition, the root node is the only node not dominated by any other node, it must be the case that dominance be irreflexive, like the two other relations considered, namely precedence and asymmetric c-command.

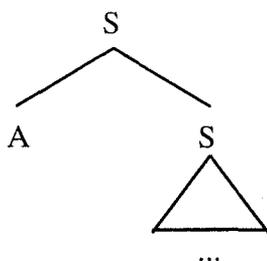
Moreover, just as Axiom 1 states the existence of a root node from which to establish the set of dominance relations, Kayne (1994) argues for another root node which will be at the basis of asymmetric c-command relations. Hence:

(8) Axiom 2

Given the set of nodes N forming a phrase marker, $\exists A \in N$ such that $\forall Z \in N, Z \neq A, A$ asymmetrically c-commands Z .

According to Kayne, A will be adjoined to the root node S , asymmetrically c-commanding everything that S dominates. Schematically:

(9)



Note that this introduces a contradiction in the system since S doesn't dominate A (only the higher segment of S does). In order to avoid this problem we have to refine Axiom 1:

¹⁰ The formulation of this axiom and of the two other axioms tries to formalize Kayne's quite concise comments on this issue. They are to be understood as conditions regulating the wellformedness of phrase

(10) Axiom 1 (revised)

Given the set of nodes N forming a phrase marker, $\exists S \in N$ such that $\forall Z \in N, Z \neq S, S$ doesn't exclude Z .

Since the higher segment of root node S dominates node A , S doesn't exclude A , as required.

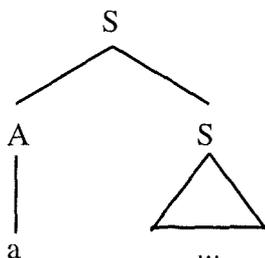
Finally, it is plausible to assume that the root node A dominates a terminal node a from which to establish the precedence relation between terminals:

(11) Axiom 3

Given the set of terminals T forming a phrase marker, $\exists a \in N$ such that $\forall z \in N, z \neq a, a$ precedes z .

Therefore, the initial structure for a phrase marker will be:

(12)



Note that we are dealing with abstract nodes that function as the initial symbols from which to build up phrase structure.

To sum up, the resultant elements of phrase structure are:

- Three axioms:
 - ◆ AXIOM 1 (introduction of root S)
 - ◆ AXIOM 2 (introduction of root A)
 - ◆ AXIOM 3 (introduction of terminal a)
- Three relations:
 - ◆ PRECEDENCE
 - ◆ DOMINANCE

1.1.4. Linking linear order and hierarchical structure: the Linear Correspondence

Axiom

As we have just seen, beyond the three initial axioms, phrase structure has three relations:

- PRECEDENCE is a linear order
- DOMINANCE is a local linear order
- ASYMMETRIC C-COMMAND is a local linear order

Precedence is a relation mainly affecting terminal nodes, whereas asymmetric c-command is rather a relation between nonterminal nodes. Just in the middle, dominance can be a relation between terminal and nonterminal nodes—which we will label *d*, following Kayne (1994)—or a relation between nonterminal nodes only—which we will label *D*, following Kayne (1994). The gist of the issue is the following: is there any mechanism relating the linear order of terminals—precedence—with the local linear order of nonterminals—asymmetric c-command?

The answer has always been negative. There has been consensus in the framework(s) of generative grammar that precedence and hierarchical relations must be kept aside. In fact, Generative Grammar has strongly argued for the privileged role of hierarchical relations against precedence, which has received a null or rather vicarious role.¹¹

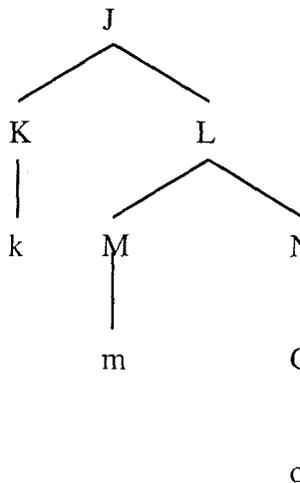
Kayne (1994) is the strongest hypothesis for a close link between precedence and c-command. He argues that these two relations are mutually dependent—*asymmetric c-command implies precedence*—but indirectly, that is, by means of dominance—in fact

¹¹ Langacker (1969) originally introduced both precedence and command ('a node A "commands" another node B if (1) neither A nor B dominates the other; and (2) the S-node that most immediately dominates A also dominates B') as conditions regulating anaphor relations. Lasnik (1976) treatment of disjoint reference conditions was based on both kommand ('a kommands b if the minimal cyclic node dominating a also dominates b') and precedence. Nevertheless, the standard version of Binding Theory has been grounded on purely hierarchical terms (i.e., c-command), following Reinhart (1976). Barss and Lasnik (1986) or Bresnan (1994) are two quite differently oriented attempts to give precedence a role in Universal Grammar.

by means of relation d , which stands for dominance between nonterminals and terminals. Let us see how the whole system works.

Relation d is a function between a nonterminal node and a set of terminal nodes equal to or higher than one. Hence, $d(X)$ defines the set of terminals dominated by the nonterminal X . Consider a concrete case:

(13)



In this phrase marker we have the following functions: $d(J) = \{k, m, o\}$, $d(K) = \{k\}$, $d(L) = \{m, o\}$, $d(M) = \{m\}$, $d(N) = \{o\}$, $d(O) = \{o\}$.

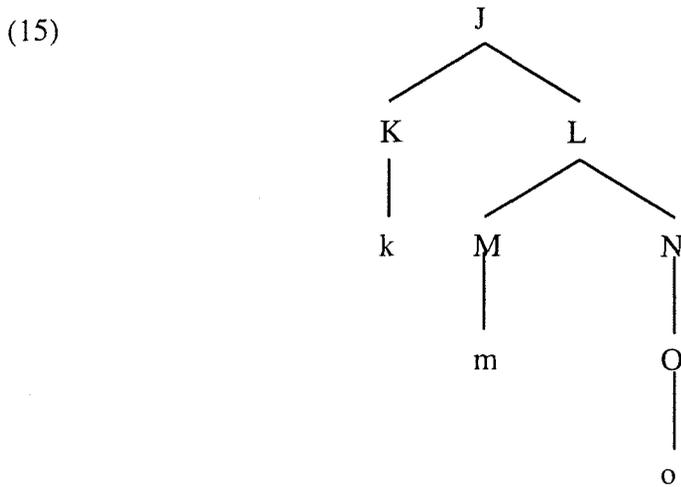
Function d can also affect ordered pairs of nodes. For example, we can define $d\langle X, Y \rangle$ as the Cartesian product $d(X) \times d(Y)$. Namely, $d\langle X, Y \rangle$ defines a set of ordered pairs $\{\langle a, b \rangle, \dots\}$ such that a belongs to $d(X)$ and b belongs to $d(Y)$. However, we haven't make precise the exact relation between X and Y . Let us say that it is asymmetric c-command. Hence, for a given phrase marker we can always form a set Σ containing all pairs of nonterminal nodes such that the first asymmetrically c-commands the second: in phrase marker (13), $\Sigma = \{\langle K, M \rangle, \langle K, N \rangle, \langle K, O \rangle, \langle M, O \rangle\}$. As stated, $d(\Sigma)$ will define the set of pairs of terminals dominated by Σ : $d\langle K, M \rangle = \langle k, m \rangle$, $d\langle K, N \rangle = \langle k, o \rangle$, $d\langle K, O \rangle = \langle k, o \rangle$, $d\langle M, O \rangle = \langle m, o \rangle$.

Having all these concepts in mind, consider now Kayne's proposal (where T is the set of terminal nodes):

(14) Linear Correspondence Axiom

$d(\Sigma)$ is a linear ordering of T .

The most intuitive consequence of the Linear Correspondence Axiom (henceforth, LCA) is the requirement that asymmetric c-command entail precedence. Consider how the LCA works in the previous phrase marker (repeated here for the ease of reference):



We have just seen that $d(\Sigma) = \langle k, m \rangle, \langle k, o \rangle, \langle m, o \rangle$. The LCA requires that $d(\Sigma)$ be a linear ordering of the set of terminals $\{k, m, o\}$. As we can easily see, the requirement is fulfilled, since it respects transitivity — k asymmetrically c-commands m , m asymmetrically c-commands o , so then k asymmetrically c-commands o —, asymmetry — k precedes m and o but not conversely, and m precedes o but not conversely—, and totality — all terminal nodes enter into the order relation.

The consequences of the LCA are widespread on both theoretical and empirical grounds. They will be the core of the next section.

1.2. Some consequences

1.2.1. Deriving phrase structure assumptions

The wellformedness conditions of phrase structure have commonly been grounded on the X' schema and on general assumptions on the form of phrase markers. In general, phrase structure relies on:¹²

¹² The X' schema in (16)a is the standard two-level application of the more general one

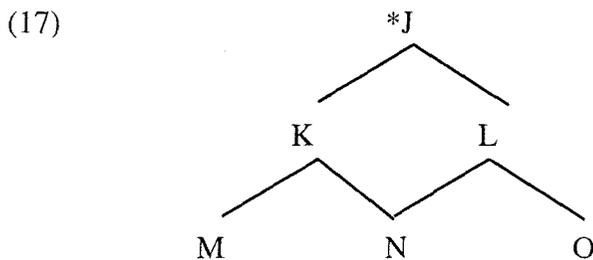
(i) $X^n \rightarrow X^{n-1}$

Other applications have been proposed in the literature, the most outstanding that of Jackendoff (1977).

- (16) a. i) $X'' \rightarrow (YP) X'$
 ii) $X' \rightarrow X (ZP)$
- b. i) Multidominance is not allowed.
 ii) Loops are not allowed.
 iii) Discontinuous constituents are not allowed.¹³
 iv) Floating constituents are not allowed.

As Speas (1990) pointed out, the conditions in (16)b are quite different from the X' schema. Whereas the former are intended to define what counts as a possible linguistic structure, the latter constrain the way to give grammatical content to the built structures. In this paragraph we will only be concerned with the former, handling with X-bar Theory in the next paragraph.

First of all, consider *multidominance*:

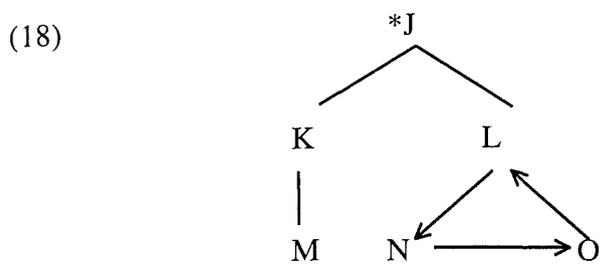


We have seen in 1.2.2. that dominance is a locally total relation, that is, for a given node X dominated by both Y and Z it must be the case that either Y dominates Z or Z dominates Y (or $Y = Z$). Note that such a property doesn't hold in (17): since N is dominated by both K and L , it must be the case that either K dominates L or L dominates K , which is against the facts. Hence, multidominance is excluded by one of the properties of the dominance relation, namely local totality.¹⁴

¹³ This is a fairly standard assumption, even though it has been argued against by several scholars. The most outstanding proposal is that of McCawley (1982), which proposes a set of phrase structure conditions allowing discontinuous constituents. See also Baltin (1987), McCawley (1987), and the collective monographic volumes Huck & Ojeda (eds.) Bunt (ed.).

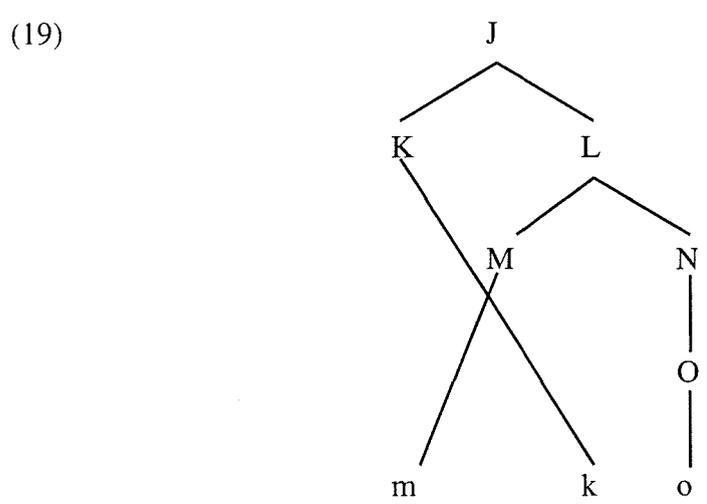
¹⁴ As the perspicacious reader have already noticed, multidominance cannot be accounted for by means of phrase structure rules (but see Ojeda 1987, who develops a GPSG framework allowing multidominance). Wall (1972: 148) introduces the Nontangling Condition to exclude multidominance (and trees with 'crossing branches'):

Consider now the possibility of having loops in our phrase markers (for the sake of clarity I use arrows to indicate dominance relations within the loop):



If we assume that L dominates N, N dominates O, and O dominates L, then it might be said that N precedes and dominates O. However, note that in this phrase marker the dominance relation doesn't meet the asymmetry requirement: O dominates L, but since L dominates N and N dominates O, it must also be the case —by transitivity— that L dominates O. Since we have assumed that dominance is *both* asymmetric and transitive, there is no way to build up a phrase marker including loops.¹⁵

Take now the case of discontinuous constituents. Consider the following phrase marker:



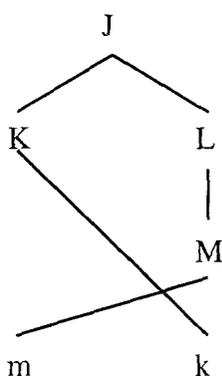
The Nontangling Condition

In any well-formed constituent-structure tree, for any nodes *x* and *y*, if *x* precedes *y*, then all nodes dominated by *x* precede all nodes dominated by *y*.

¹⁵ Note that assuming that L dominates N, N dominates O, and L dominates O brings us to a case of multidominance (the node O is dominated by both L and N), which is against the (local) totality requirement of dominance, as we have just seen.

Here *m* and *o*, which belong to constituent *L*, are separated by a terminal *k*, which doesn't. The LCA prevents the possibility of such a structure. Since *K* asymmetrically *c*-commands *M*, *N*, and *O*, and *M* asymmetrically *c*-commands *O*, we obtain the following ordered pairs of terminals $\langle k, m \rangle$, $\langle k, o \rangle$, $\langle m, o \rangle$. Then, by the LCA it must be the case that these three pairs form a linear ordering $\langle k, m, o \rangle$, which is against the facts. There is thus no way for a phrase marker containing a discontinuous constituent to satisfy the LCA. In fact, the ban against discontinuous constituents is a subcase of a more general ban against 'crossing branches'. Since asymmetric *c*-command implies precedence, any phrase marker of the following sort

(20)



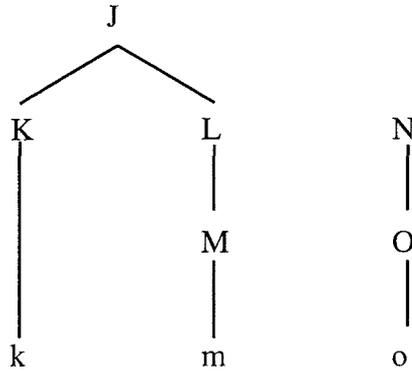
will violate the LCA: given that *K* asymmetrically *c*-commands *M*, *k* must precede *m*.

This ban against crossing branches, and consequently against discontinuous constituents seems desirable on theoretical grounds, since it makes Universal Grammar more restrictive. Assuming a more permissive phrase structure, like the one advocated for in McCawley (1968, 1982) should be grounded on clear empirical evidence, which is not.¹⁶

Finally, consider the case of floating constituents:

¹⁶ McCawley (1982) claims that the only reason for rejecting discontinuous constituency is the primacy of the vision of strings of terminal nodes as the outcome of phrase structure rules. He argues that if tree structures are taken as primitives there is no reason to reject multidominance. It is true that phrase structure rules cannot generate phrase markers like (19) or (20). However, this criticism doesn't extend to the present framework, which dispenses with phrase structure rules.

(21)



In this case the constituent N is completely disconnected of the rest of the phrase marker, and there is no obvious way to establish c-command relations. That amounts to saying that no ordering can be established between the terminals dominated by N and those dominated by J. Hence, the totality requirement of the precedence relation is not met, and {k, m, o} do not form a linear ordering, violating the LCA.¹⁷

To sum up, we have just shown that crucial intuitive assumptions on the wellformedness of phrase markers straightforwardly follow from the interaction between our three phrase structure relations —precedence, dominance, and asymmetric c-command— and the LCA.

1.2.2. Deriving X' Theory

The X' schema

- (22) a. $X'' \rightarrow (YP) X'$
 b. $X' \rightarrow X (ZP)$

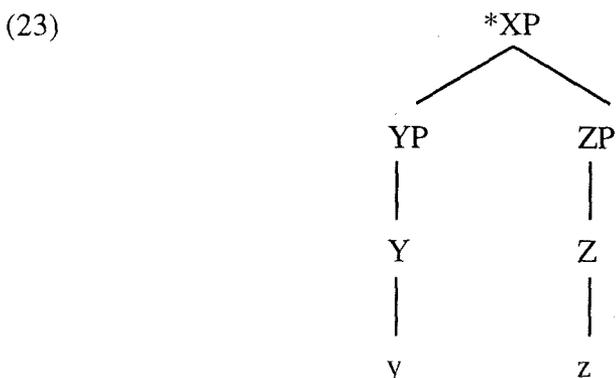
implicitly embodies two assumptions on the form of phrases:

- every projection has a head
- no projection has more than a head

¹⁷ Moreover, the existence of floating constituents violates Axiom 1, since the floating constituent is excluded by the root node S. See also Wall's (1972: 146) Single Root Condition.

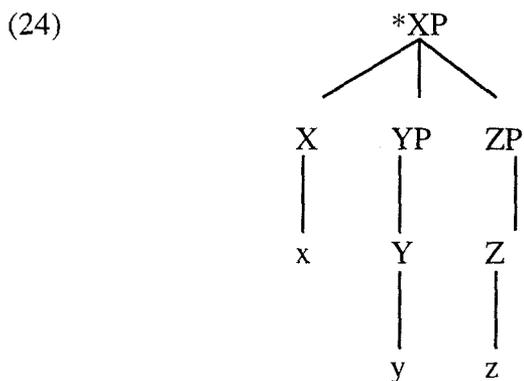
Even though, both statements seem highly desirable, on empirical grounds, they are by no means conceptually necessary. In other words, these two conditions are a good generalization on the form of phrase markers, but they cannot be considered principles governing their wellformedness.

However, when they are considered in the light of the system presented so far, an interesting turn follows. Consider a structure without a head:



We don't want such a structure to be possible at all, and the LCA provides us with a principled reason for rejecting it. Since YP asymmetrically c-commands Z and ZP asymmetrically c-commands Y, $d(\Sigma) = \langle y, z \rangle, \langle z, y \rangle$. Obviously, asymmetry is not respected and hence no linear ordering can be given to $\{z, y\}$. Hence, nonheaded structures will be rejected on the basis of the asymmetry property of asymmetric c-command.

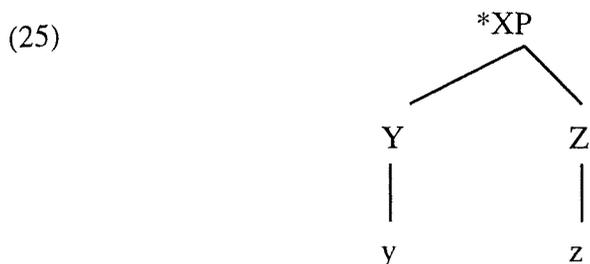
The following structure suffers from the same problem:



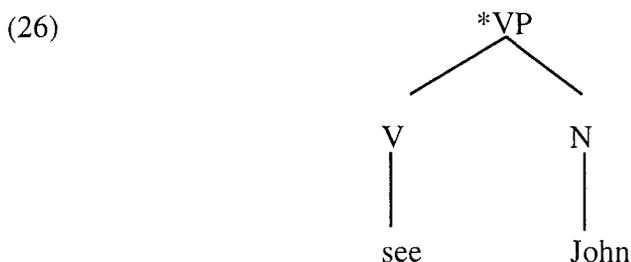
The relation between YP and ZP is symmetric, making the ordering between y and z impossible. Thus, the LCA offers a principled explanation for the obligatoriness of

binary branching, and hence the impossibility of having two complements at the same level.

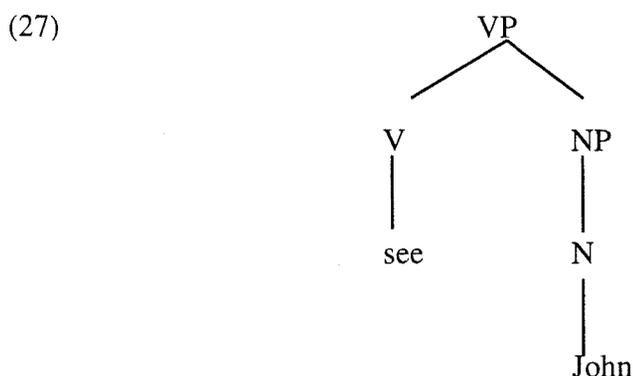
Let us analyze now the ban against structures with two heads:



Here Y and Z c-command each other, so the asymmetry property doesn't hold. As a consequence, we cannot assign a linear ordering to the set of terminals {y, z}. (25) is formally identical to the impossible phrase marker (irrelevant details are omitted)



The only alternative allowed by the LCA is the following:¹⁸



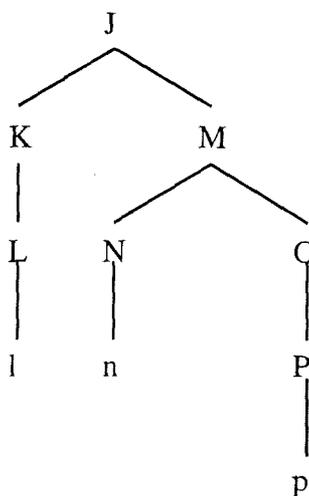
¹⁸ At this point becomes apparent the choice in the formulation of c-command. The one adopted, based in 'the first node up', gives the correct results. Note, however, that a formulation based in 'the first branching node', like the one proposed by Reinhart (1976) —see footnote 9—, would imply that the relevant node for establishing the c-commanding domain of N should be VP. In other words, under Reinhart's formulation, N would c-command V, which would introduce an undesirable symmetry and would render the linear ordering between *see* and *John* impossible.

Here the symmetry between V and N is broken. Hence, $\Sigma = \langle V, N \rangle$, and $d(\Sigma) = \langle \text{see}, \text{John} \rangle$, which is a linear ordering for the set $\{\text{see}, \text{John}\}$.

1.2.3. Specifiers as adjuncts

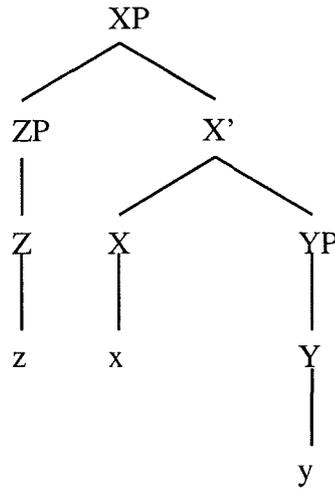
One of the consequences of the X' schema is the introduction of a nonterminal immediately dominated by the maximal projection, namely a specifier. However, it has to be seen which is the status of specifiers in our system. Consider the following phrase marker:

(28)



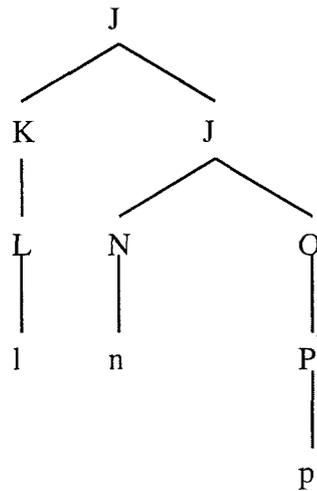
Here we have the following asymmetric c-commanding relations: $\langle K, N \rangle$, $\langle K, O \rangle$, $\langle K, P \rangle$, $\langle M, L \rangle$, and $\langle N, P \rangle$. Hence, $d(\Sigma) = \langle l, n \rangle$, $\langle l, p \rangle$, $\langle n, l \rangle$, $\langle p, l \rangle$, and $\langle n, p \rangle$. Obviously, there is no linear ordering for the set of terminals, and the phrase marker is to be rejected because of the LCA. Nevertheless, the preceding structure is identical to the standard phrase marker

(29)



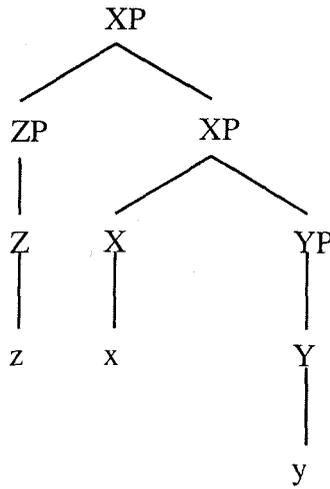
In other words, the system just developed doesn't allow specifiers. This is the truth, but not the whole truth. Note that the crucial point is the nonterminal M/X', which breaks asymmetry. The phrase marker in (28) becomes admissible if $M = J$:

(30)



Here the asymmetric c-commanding relations are $\langle K, J \rangle$, $\langle K, N \rangle$, $\langle K, O \rangle$, $\langle K, P \rangle$, and $\langle N, P \rangle$. Hence, $d(\Sigma) = \langle l, n \rangle$, $\langle l, p \rangle$, and $\langle n, p \rangle$, which crucially differs from the preceding case in the exclusion of $\langle n, l \rangle$ and $\langle p, l \rangle$: since the lower segment of J is not a category, it cannot count for asymmetric c-command. As a direct consequence of the change, $d(\Sigma)$ can be converted in the linear ordering $\langle l, n, p \rangle$. The translation in X' terms is quite apparent:

(31)

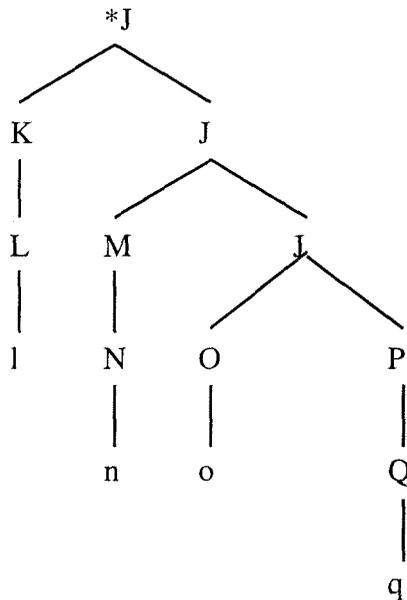


We come up to the result that specifiers must be adjuncts. This is the only configuration allowed by the LCA.

1.2.4. Against multiple adjunction

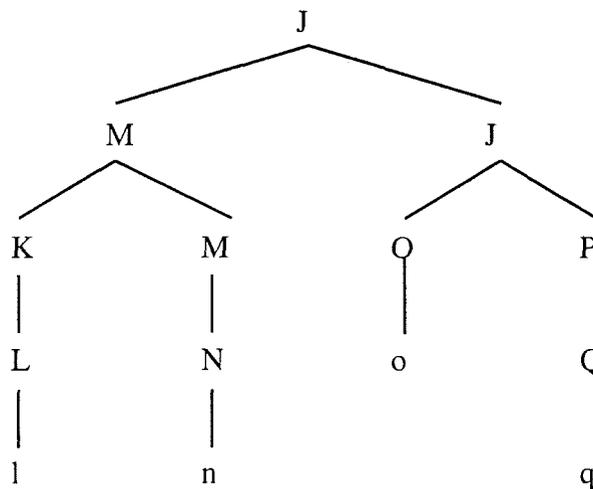
In the standard conception of phrase structure XPs were expected to have just one specifier, whereas no limit was imposed on adjuncts. Now, in a system that doesn't make any distinction between specifiers and adjuncts, we must come up against the issue of multiple adjunction structures. Take for example the following phrase marker:

(32)



K and M are two maximal projections adjoined to maximal projection J. Given the definition of asymmetric c-command, which makes crucial use of the category-segment distinction, K asymmetrically c-commands N and M asymmetrically c-commands L. As a consequence, $d(\Sigma)$ includes both $\langle l, n \rangle$ and $\langle n, l \rangle$, which is against the asymmetry requirement. So then, the system bars the possibility of having two XPs adjoined to the same maximal projection. Obviously, this seems to be too strong a restriction. Nonetheless, the system allows us for a different configuration capable of explaining the cases of multiple adjunction:

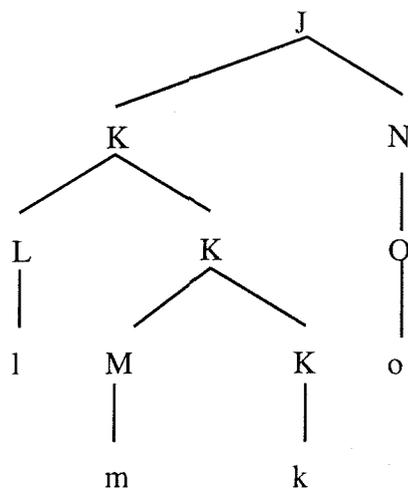
(33)



Here, the problem of symmetry disappears. K asymmetrically c-commands both M and N (and J, O, P, and Q), and M asymmetrically c-commands J, O, P, and Q. Hence, $d(\Sigma) = \langle l, n \rangle, \langle l, o \rangle, \langle l, q \rangle, \langle n, o \rangle, \langle n, q \rangle$, and $\langle o, q \rangle$, which is a linear ordering (it can form the ordered set $\langle l, n, o, q \rangle$). So, in principle, the only possibility of multiple adjunction is adjunction to a already adjoined category.

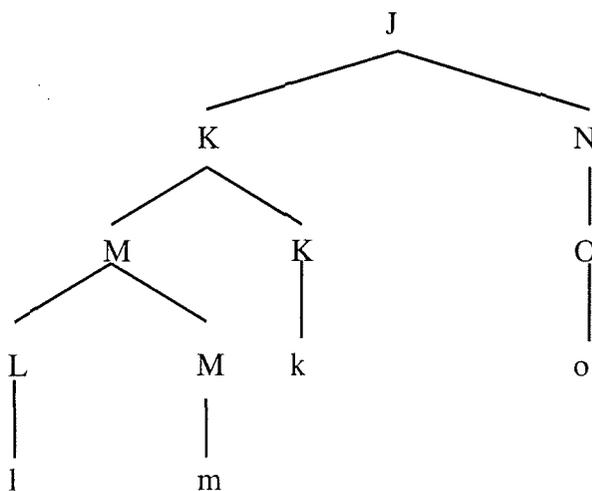
Note that the same reasoning extends to multiple adjunction to a head. The standard configuration violates the LCA :

(34)



In this structure L and M c-command each other: they exclude one the other, both are categories and they are dominated by the same categories, namely J. Hence, asymmetry doesn't hold and no ordering can be given to terminals {l, m}. The solution has been suggested before:

(35)



In this configuration, L asymmetrically c-commands both M and K (plus N and O). Consequently, $d(\Sigma) = \langle l, m \rangle, \langle l, k \rangle, \langle l, o \rangle, \langle m, k \rangle, \langle m, o \rangle, \langle k, o \rangle$, which is a linear ordering.

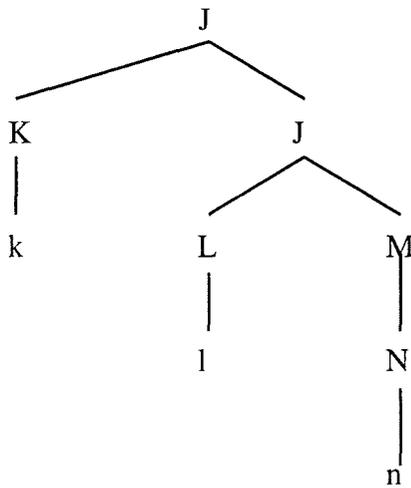
1.2.5. More on adjunction

Note that when arguing against the possibility of multiple adjunction in standard terms, we have implicitly reduced the conditions of application of adjunction to two cases:

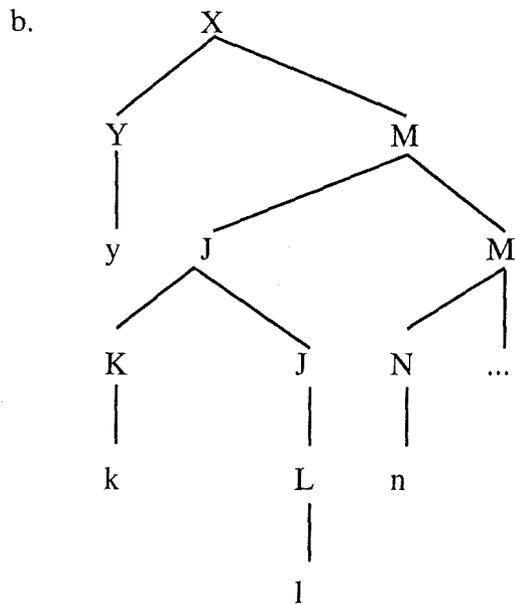
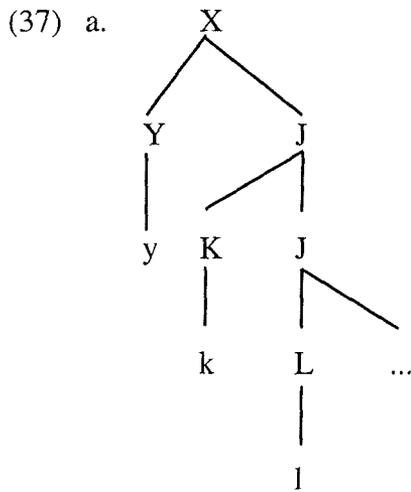
- adjunction of a maximal projection to a maximal projection
- adjunction of a head to a head

These are well established assumptions in the Principles and Parameters framework. However, there is no obvious reason of why they should hold. Consider the case of adjunction of a head to a maximal projection:

(36)



In principle the structure fulfills the requirements of the LCA: K asymmetrically c-commands J, and everything dominated by J, namely, L, M, and N. As a consequence, $d(\Sigma) = \langle k, l \rangle, \langle k, n \rangle, \langle l, n \rangle$, which is a linear ordering (see Kayne 1994: 31). However, even though the isolated structure is unproblematic, the things radically change when it is embedded in a higher one. Since, by Axiom 2, there must be a node A asymmetrically c-commanding all the structure, the preceding structure must necessarily be embedded in either of the following configurations:

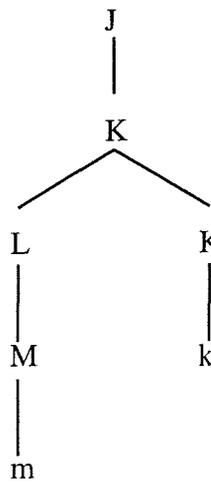


Firstly consider (37)a. Here neither Y nor K asymmetrically c-command each other, so no ordering can be provided for terminals y and k (note that it must be the case that $Y \neq A$, otherwise a contradiction with Axiom 2 will arise), which entails a violation of the totality requirement. In phrase marker (37)b, we still have the symmetric relation between Y and K. However, it seems *prima facie* that since Y asymmetrically c-commands J, we have the ordered pair $\langle y, k \rangle$ under $d(Y)$. Nevertheless, when the category/segment distinction is taken into account, this illusion dispels, for J doesn't dominate K. In other words, in this configuration no ordering is provided for y and k either (again, Y must be distinct from A, for the reason just given before). To sum up, the impossibility of adjoining a head to a maximal projection directly follows from the interaction of the LCA with basic axioms of phrase structure.¹⁹

Let us examine now the case of XP adjunction to a head:

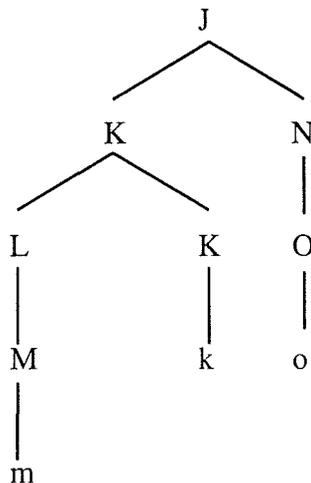
¹⁹ Kayne himself attempts to forbid heads in specifier position by means of a condition requiring 'that the highest element in a chain of heads must have a specifier' (p. 31). However, as it has been shown, the development of the system renders such an *ad hoc* condition superfluous.

(38)



As Johnson (1997) correctly points out, in this configuration there is no problem at all with the LCA: L asymmetrically c-commands K, and hence $d(\Sigma) = \langle m, k \rangle$, which is a linear ordering. Nevertheless, things take a different turn when more structure is added. First of all, the head K cannot have a complement:

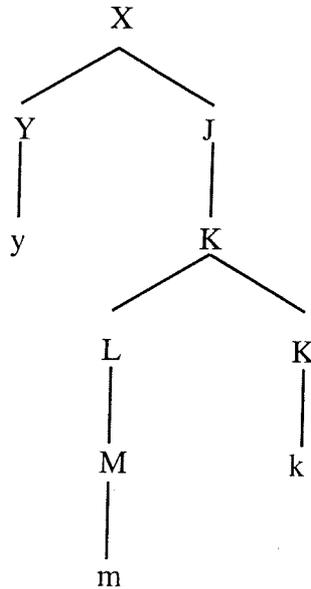
(39)



Here, L asymmetrically c-commands O and N M, yielding the contradictory pairs $\langle m, o \rangle$ and $\langle o, m \rangle$. Hence, the asymmetry requirement is not met.

Moreover, the intransitive head cannot be projected within a complement position:

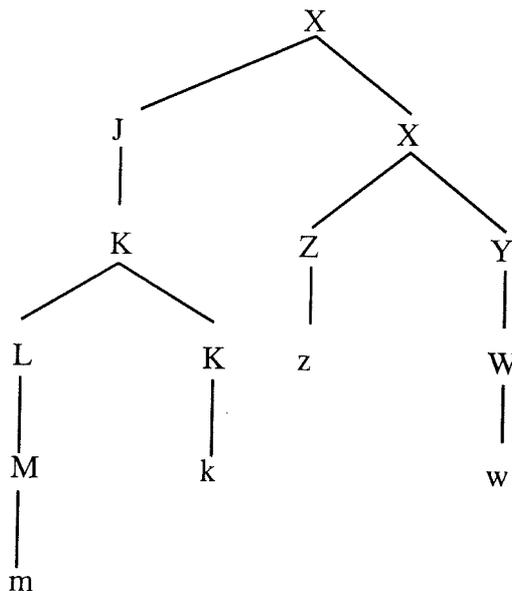
(40)



Even though the structure is perfect with respect to the LCA, Kayne (1994: Chapter 5, note 10) correctly notes that the crucial point here is to establish what the source would be for the adjoined phrase if there is no material to the right. Obviously, there is no way for this structure to be derived.

Finally, let us study the possibility that the head project in a specifier position:

(41)



Again the structure is unproblematic for the LCA, but as in the previous case there is something amiss with the source of L. It must be the case that its source be Y (or a phrase dominated by Y). However, it easy to test that moving L from this position goes

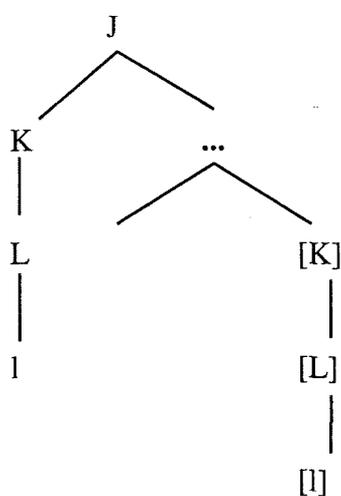
against all standard assumptions on movement, namely, that the landing site c-command the source. In this structure L doesn't c-command Y, so then it is an instance of improper movement.

To sum up, adjunction of an XP to a head, an undesirable structure on standard hypothesis about phrase structure, is also rejected by the LCA, even though indirectly: the only structures compatible with the LCA either have no source for the adjunct —i.e. (40)— or involve an instance of improper movement —i.e. (41).

1.2.6. Restricting the transformational component

The problem stated by phrase marker (41) brings us to another issue having crucial consequences for the design of Universal Grammar: movement. It is a standard assumption since the seventies (e.g. the Proper Binding Condition of Fiengo 1977), that movement must be upward, namely that the landing site must c-command the source. However, in the system developed here c-command is too weak a relation. Kayne (1994: chapter 5, note 2) suggests that the copy theory of movement along the lines pursued by Chomsky (1993, 1995) would directly force movement to respect asymmetric c-command. Take the following hypothetical configuration where it is assumed that the moved XP and its copy c-command each other (I conventionally mark the copy by means of square brackets):

(42)

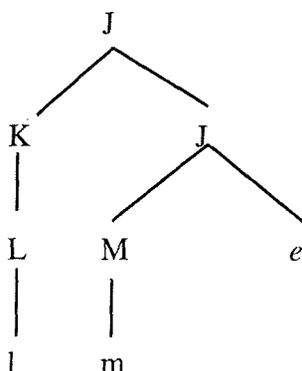


Kayne argues that this configuration violates the LCA since K asymmetrically c-commands [L], and [K] L. Hence, if a copy theory of movement is assumed, the

requirement that the trace be asymmetrically c-commanded by the moved element straightforwardly follows.

However, does a configuration like the one in (42) ever arise? Consider the movement possibilities allowed in our system. Firstly, substitution/adjunction to a specifier must move a maximal projection having its source in the complement (or in a projection dominated by it):

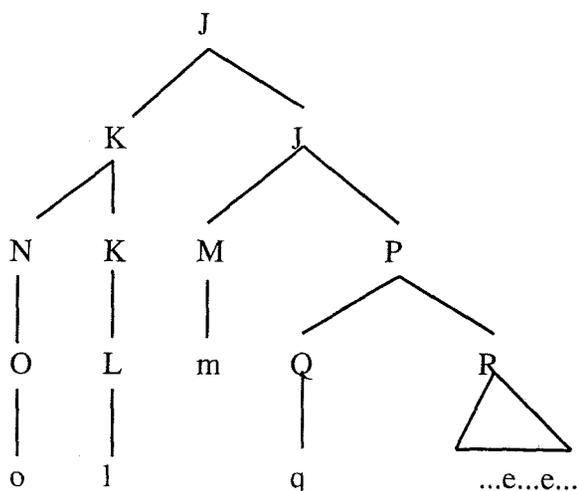
(43)



We have already seen that the specifier asymmetrically c-commands all the material under the maximal projection it is adjoined to. Hence, this option is unproblematic for the LCA, because it necessarily forces an asymmetric c-command relation between the moved element and its trace.

The second option is adjunction of a maximal projection to the specifier of another specifier:

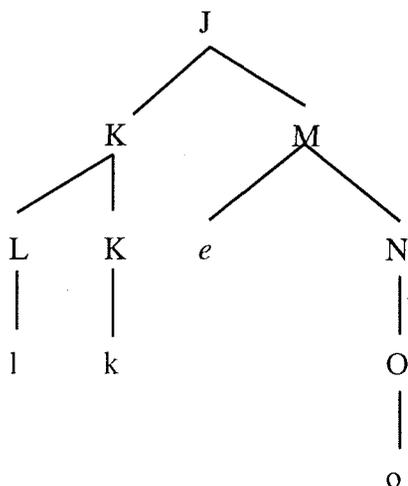
(44)



Here both N and K asymmetrically c-command P, hence they asymmetrically c-command their traces as well.

Consider the third possibility: adjunction of a head to another head. Three possible configurations follow. Firstly, a head moving from within the complement of the landing site:

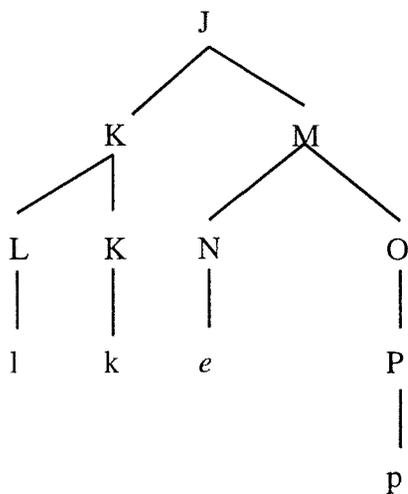
(45)



L asymmetrically c-commands M, and consequently its trace. So the movement is unproblematic.

Secondly, consider a head moving from within a specifier to the higher head:

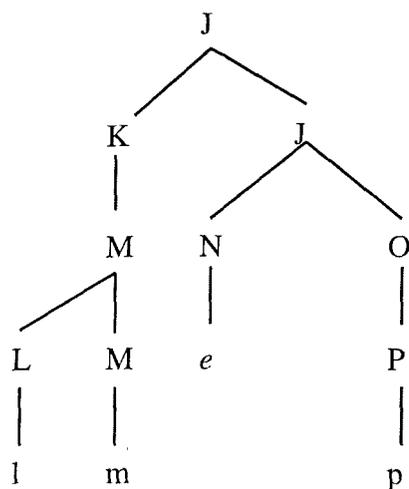
(46)



Again, the configuration is perfect regarding the LCA: L asymmetrically c-commands M, and hence its trace.

Finally, consider the case in which a head adjoins to the head of its own specifier:

(47)



Here the things completely change: L doesn't asymmetrically c-command N, because the first category dominating L —i.e. K— doesn't dominate N. So then, in this configuration the moved head would not even c-command its trace, clearly a case of improper movement on all respects.

In sum, the requirement that the moved element *asymmetrically* c-command its trace doesn't have to be stipulated, nor linked to a copy theory of movement, but it independently follows from the design of Universal Grammar. Obviously, since asymmetric c-command maps into precedence, the immediate consequence of this requirement is that the landing site must precede the source. In other words, movement must necessarily be to the left. This highly restricts the possibilities of the transformational component —already restricted by phrase structure conditions—, certainly a desirable result on conceptual grounds.

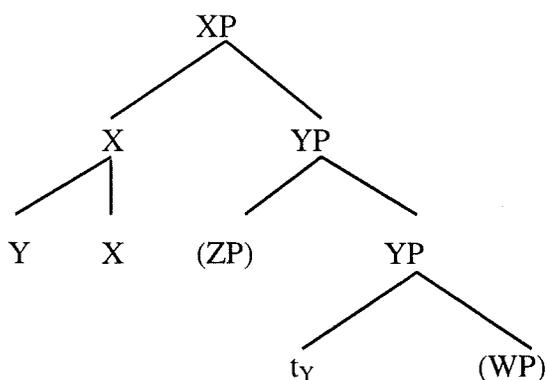
Note, however, that the restrictions the LCA imposes on the transformational component, namely, that movement must be to the left, go further off the limits of conceptual desirability: it has many empirical consequences. All the operations crucially involving right movement —Heavy NP Shift, Right Node Raising, Extraposition, or Clitic Right Dislocation, to name a few— must now receive a completely different analysis (see Kayne 1994 for some detailed proposals; see also Roca 1997: Ch. 1 for a summary). If Kayne's framework were capable of giving a principled explanation for all these constructions without resorting to right movement, it would receive a strong empirical support. This dissertation is devoted to showing that this highly restrictive

hypothesis on the form of Universal Grammar is strongly supported by data on Clitic Right Dislocation.

1.2.7. The emergence of functional categories

For those who work in the Principles and Parameters framework, the explosion of functional categories is a well-known phenomenon. It is extremely difficult to name all of them, but every syntactician is able to make a list of at least ten different functional categories forming a simple sentence. This proliferation has been criticized from different scholars working within and outside the Principles and Parameters framework, but its critics have mainly focused on the somewhat baroque aspect of the resultant clause structure. Nevertheless, very few attention has been paid to the theoretical advantages derived from the articulation of a functional skeleton. This is not the place to enter into the details of the advantages of assuming that grammatical concepts such as Tense, Aspect, or Focus should project into syntactic categories (a more restrictive hypothesis on the internal structure of phrases, a more transparent syntax-semantics relationship, a clear approximation to crosslinguistic variation, a different approach to the acquisition problem, etc.; see Hoekstra 1995 for a brief review); moreover, we'll argue for two functional topic projections in Chapter 3. In any case, note that, independently of the exact motivation for each functional category, their proliferation may be understood as a side effect of the constraints the LCA imposes on phrase structure. Since only one adjunct/specifier is allowed by phrase, the room for moved phrases is extremely shortened. For example, once it is assumed that the Specifier-Head-Complement is universal, the existence of verb-initial languages entails the movement of the verb to a higher functional projection (without a lexically filled specifier):

(48)



In other terms, the decreasing of accessible landing sites forces the increasing of functional phrases in order to host moved phrases. Whether this is an accurate explanation of the human faculty is an empirical issue yet to be elucidated, not a theoretical shortcoming *per se*.

1.3. Summary

In this chapter it has been shown that the Hypothesis of the Antisymmetry of Syntax, as fleshed out in Kayne (1994), argues for a direct link between linear precedence and hierarchical structure. This link, formulated as the Linear Correspondence Axiom, ensures that if a category *A* asymmetrically *c*-commands *B*, the set of terminals dominated by *A* will precede the set of terminals dominated by *B*. As a consequence, specifiers must precede heads, and heads complements, yielding a universal specifier-head-complement order. Moreover, the LCA allows us to dispense with the conditions imposed by X-bar Theory, and constrains the transformational component, since movement must necessarily be to the left. The resultant picture is thus conceptually more desirable, even though its own conceptual strength forces many standard analysis to be reconsidered. In Chapters 4 and 5 we will test its empirical coverage in one concrete area, that of clitic right dislocation.



Chapter 2. A typology of detachment processes in Catalan (I): Left-detachments

This chapter is devoted to characterize the range of left-detachment processes in Catalan and other Romance languages —a companion chapter will deal with right-detachments. A Cartesian spirit guides this commitment: to convert the somewhat messy realm of topic elements marked by left-detachment into a typology of clear and distinct constructions. As always happens in science, many of the insights found here are inspired by previous work by many scholars. This will become apparent by the cross-examination of sections 2.1 and 2.2, which have a big debt with the work of Guglielmo Cinque. The existence of such achievements will make our Cartesian endeavor easier, since I will be able to refer the reader to the relevant literature whenever a thorough description would prove tedious. In turn, those aspects that have received less attention in the literature will be discussed with more detail.

The chapter begins with a description of one of the best-known topic constructions we found in the left periphery, Clitic Left Dislocation (section 2.1). In section 2.2, I characterize Hanging Topic Left Dislocation, following and broadening Guglielmo Cinque's work. In section 2.3, I discuss the independent status of the *as for* construction with respect to other topic constructions, particularly Hanging Topic Left Dislocation. Finally, in section 2.4, I briefly consider a putative fourth topic construction, metalinguistic topics, which I show that has an independent status on both syntactic and interpretative grounds.

2.1. Clitic left dislocation

2.1.1. Syntactic properties

Since Cinque (1977) it is well established that Romance languages have a typical construction apparently similar to Germanic Left Dislocation, but which is indeed a quite distinct one: he labels it *Clitic Left Dislocation* (hereafter: CLLD) in Cinque (1983). We have examples of CLLD in (the use of the comma doesn't entail a real pause nor an intonational break; see Dolci 1986 and Zubizarreta 1998):

- (1) a. De la Maria, en vam parlar ahir. [Catalan]
of the Maria of.it PAST-2PL talk yesterday
'About Maria, we talked yesterday.'
- b. A Lucille, je lui écris constamment. [French: Postal 1991]
to Lucille I his/her write constantly
'Lucille, I write to constantly.'
- c. Esos zapatos queroos eu. [Galician: Álvarez et al. 1986]
these shoes want-them I
'These shoes, I want.'
- d. Al mare, ci siamo già stati. [Italian: Cinque 1990]
to.the seaside there were-2PL already been
'In the seaside, we have already been.'
- e. Del libre de Joan, n'ai parlat. [Occitan: Sauzet 1989]
of-the book of Joan of.it-have-1 talked
'John's book, I have talked about.'
- f. Ao corvo, a raposa roubou-lhe o queijo. [Portuguese: Mateus et al. 1983]
to.the raven the fox stole-to.him the cheese
'From the raven, the fox stole the cheese.'
- g. Florile, le cumparase Maria. [Romanian: Motapanyane 1994]
flowers-the them bought-3 Maria
'As for the flowers, Maria had bought them.'
- h. La littera, Maria l'at dza mandata. [Sardinian: Jones 1990]
the letter Maria her has already sent
'The letter, Maria has already sent.'
- i. Todo esto, lo dijo un profesional. [Spanish]
all this it said a professional
'All this, a professional said.'

Beyond the Romance family, many other languages show (a more or less equivalent to) CLLD:¹

¹ Aoun & Benmamoun (1998) consider Lebanese Arabic, who they claim to show CLLD. They offer examples like the following:

- (i) a. *Naadya jeef-a Kariim mbeerih.*
Nadia saw-3-her Karim yesterday
'Nadia, Karim saw yesterday.'

- (2) a. Ton Janni den ton ksero. (Modern Greek: Anagnostopoulou 1997)
 the John not him know-1
 'John, I don't know.'
- b. Ja tzyaq ch'ooyaa' x-ee-tij-ow-i. (Tzutujil: Aissen 1992).
 the clothes rats ASP-A3PL-eat-FOC-IV
 'Rats were the ones who ate the clothes.'
- c. Alenje njũchi zin-ná-wá-lum-a. (Chichewa: Bresnan & Mchombo 1987)
 hunters bees SM-PAST-OM-bite-INDIC
 'The hunters, the bees bit.'

Furthermore, Baker (1996) has strongly argued that polysynthetic languages have CLLD (and CLRD) as the only option for a lexical NP argument to surface. Hence, if his analysis is correct, Mohawk, Nahuatl, Gunwinjguan, Wichita, Chukchee, Ainu, and Tanoan languages should be included in the list of languages having CLLD. Probably, CLLD is present in every language licensing null pronominals either by means of agreement or by means of clitics.

Descriptively speaking, CLLD involves a phrase in the left periphery of the sentence —the dislocate— which is linked to a resumptive clitic (or agreement marker) within the sentence (be it either visible or null). Schematically:

- (3) CLLD₁ [_S ... clitic/AGR₁...]

The clitic/AGR fulfills all the functions the dislocate would have done if it had not been displaced.

Syntactically, CLLD can be characterized as follows (some of these properties will be studied in more detail in Chapter 4):

D) *It affects maximal projections of any category:*

-
- b. Sməʃt ʔənnə Naadya rəht mən duun ma təhke maʔ-a.
 heard-1 that Nadia left-2 without COMP talking-2 with-her
 'I heard that Diana, you left without talking to her.'

As we will see during this chapter, even though Aoun and Benmamoun label this construction CLLD, its properties seem closer -even though not identical- to those of Romance HTLD than to those of Romance CLLD.

- (4) a. El llibre, el vam comprar a Barcelona. [DP]
 the book him PAST-2PL buy in Barcelona
 ‘The book, we bought in Barcelona.’
- b. De llibre, ja me’n vaig comprar un. [NP]
 of book already to.me-of.it PAST-1 buy one
 ‘I already bought one book.’
- c. De la Maria, en vam parlar ahir. [PP]
 of the Maria of.it PAST-2PL talk yesterday
 ‘Maria, we talked about yesterday.’
- d. D’intel·ligent, no ho és pas. [AP]
 of intelligent not it is NEG
 ‘Intelligent, (s)he is not.’
- e. Obertament, la Maria no hi ha parlat mai. [AdvP]
 overtly the Maria not LOC has talked never
 ‘Overtly, Maria has never talked.’
- f. Que té por de la Maria, ho sap tothom. [CP_[+tense]]
 that has fear of the Maria it knows everybody
 ‘That (s)he is afraid of Maria, everybody knows.’
- g. Comprar un cotxe, la Maria sempre ho ha intentat. [CP_[-tense]]
 buy a car the Maria always it has tried
 ‘To buy a car, Maria has always tried.’

The restrictions affecting NPs and VPs merit a more extensive comment: the former will be discussed in 2.1.2., when reviewing the interpretive properties of CLLD; the latter will be described in the Appendix to this chapter.

II) *It is iterative:*

- (5) a. Al Pere₁, del llibre₂, li₁’n₂ va parlar ahir.
 with the Pere of-the book of.it+LOC PAST-3 talk yesterday
 ‘(S)he talked with Pere about the book.’
- b. Les cerveses₁, a Barcelona₂, les₁ hi₂ venen molt cares.
 the beers at Barcelona them LOC sell-3PL very expensive-FEM.PL
 ‘They sell beers very expensive at Barcelona.’

Again, this property of CLLD is well documented in many other languages:

- (6) a. O queijo, ao corvo, a raposa roubou-lho. [Portuguese: Mateus et al. 1983]
 the cheese to.the raven the fox stole-to.him-it
 ‘The fox steal the cheese from the raven.’
- b. *De vinu, a sos pitzinnos, non nde lis dao.* [Sardinian: Jones 1993]
 of wine to the children not of.it to.them give-1
 ‘Wine, to the children, I do not give them any.’
- c. *Oggi, a Giorgio, io, un lavoro, non posso offrirlglielo.* [Italian: Benincà et al. 1988]
 today to Giorgio I a job not can-1 offer-to.him-it
 ‘Today I cannot offer Giorgio a job.’
- d. *Tis Marias, ta vivlia, tis-ta-estile o Yanis.* [Greek: Tsimpli 1995]
 the-GEN Maria the-ACC books her-them-sent the Yanis
 ‘Yanis sent the books to Maria.’
- e. *Ja nata’ ooxiij nb’e K’oqol Keej.* [Tzutujil: Aissen 1992]
 the my.father in.3.days goes Masatenango
 ‘In three days, my father is going to Masatenango.’²

III) *It occurs in both root and embedded contexts:*

- (7) a. Del llibre, en va parlar ahir la Maria.
 of-the book of.it PAST-3 talk yesterday the Maria
 ‘The book, Maria talked about yesterday.’
- b. Sembla que, *del llibre, en va parlar ahir la Maria.*
 seems that of-the book of.it PAST-3 talk yesterday the Maria
 ‘It seems that the book, Maria talked about yesterday.’

² The topic character of the adverbial *ooxiij* becomes clear when it is taken into account that the preverbal position is occupied by focalized elements. Consider (i), where the first element (in italics) is a topic and the preverbal one (in boldface) a focus:

- (i) *Ja tzyaq* **ch’ooyaa’** x-ee-tij-ow-i.
 the clothes rats ASP-A3PL-eat-FOC-IV
 ‘Rats were the ones who ate the clothes.’

If we compare (i) with sentence (6)e, it becomes quite apparent that *ooxiij* is not associated with focus even though it appears preverbally (as the translation shows). Hence, (6)e is a genuine case of multiple CLLD.

Again, this behavior is attested crosslinguistically:

- (8) a. E depois a avó contou que, ao corvo, a raposa lhe tinha roubado o queijo.
 ‘And afterward, the grandfather told that the fox had stolen the cheese from the raven.’ [Portuguese: Mateus et al. 1983]
- b. Juanne at natu ki, *sa littera*, l’áiát dza mandata.
 ‘John said that, the letter, he had already sent it.’ [Sardinian: Jones 1993]
- c. Mi sembra che, *di Giorgio*, nessuno (*ne*) abbia parlato bene.
 ‘It seems to me that nobody talked about Giorgio favorably.’ [Italian: Benincà et al. 1988]
- d. Anarotieme an tin Maria tin agapas akoma
 wonder-1 if the-Maria_{ACC} CL_{ACC} love-2 still
 ‘I wonder if you still love Maria.’ [Greek: Anagnostopoulou, 1997]
- e. Aa Xwaan n-0-b’ij chi *Ta Mari’y* ma t-r-aajo’.
 youth Juan ASP-E3-say that Miss Maria NEG ASP-E3-want
 ‘Juan says that María doesn’t wants it.’ [Tzutujil: Aissen 1992]³

IV) *In cases of multiple CLLD, the ordering of the dislocates is free:*

- (9) a. La Maria, al Pere, del llibre, li’n va parlar ahir.
 the Maria with the Pere of-the book of.it+LOC PAST-3 talk yesterday
 ‘Maria talked with Pere about the book.’
- b. La Maria, del llibre, al Pere, li’n va parlar ahir.
- c. Al Pere, la Maria, del llibre, li’n va parlar ahir.
- d. Al Pere, del llibre, la Maria, li’n va parlar ahir.
- e. Del llibre, al Pere, la Maria, li’n va parlar ahir.
- f. Del llibre, la Maria, al Pere, li’n va parlar ahir.

The free ordering of the dislocates is found cross-linguistically as well. With respect to Italian, Benincà et al. (1988: 144) observe that “[I]’ordine reciproco dei vari elementi dislocati è indifferente.” This is illustrated by the following pair from Sanfillipo (1998:ex. 15):

³ As Aissen (1992: 73) notes, the topic character of *Ta Mari’y* follows from ‘the absence of focus semantics, the absence of focus verb forms, and the position of the negation in 82 [= (8)e].’

- (10) a. A Gianni, Carlo, gliel'ha presentato Maria.
to Gianni, Carlo, to.him-him has introduced Maria
b. *Carlo, a Gianni, gliel'ha presentato Maria*

The same description is valid for Sardinian, according to Jones (1993: 317): “[a]s far as we can ascertain, the order of left-dislocated [=CLLD] items is not fixed, though it may be affected by factors such as relative length, empathy or topicworthiness.” The same behavior is found in Spanish:

- (11) a. A Juan, este libro, es mejor no enseñárselo.
to Juan, this book, is better not show-to.him-it
b. *Este libro, a Juan, es mejor no enseñárselo*

If we move to more distant languages, like Modern Greek or Chichewa, the same pattern holds (examples from Iatridou 1990 and Bresnan and Mchombo 1987, respectively):

- (12) a. O Yanis, tin Maria, tin agapa.
the-NOM Yanis the-ACC Maria her loves
b. Tin Maria, o Yanis, tin agapa.
the-ACC Maria the-NOM Yanis her loves
'Yanis loves Maria.'
- (13) a. Njûchi alenjje zi-ná-wá-lum-a.
bees hunters SM-PAST-OM-bite-INDIC
b. *Alenjje njûchi zi-ná-wá-lum-a.*
hunters bees SM-PAST-OM-bite-INDIC
'The bees bit the hunters.'

Note that the neutral word order seems to be VSO in Modern Greek (according to Tsimpli 1995) and SOV in Chichewa (according to Bresnan & Mchombo 1987).

V) *There must be a resumptive element and it must be a clitic pronoun:*

- (14) a. **De la Maria, vam parlar ahir.*
of the Maria PAST-2PL talk yesterday
'Maria, we talked about yesterday.'
- b. **De la Maria, vam parlar ahir d'ella.*
of the Maria PAST-2PL talk yesterday of her
'Maria, we talked about yesterday.'

Note that similar sentences may be grammatical under different interpretations and intonation patterns. For instance, (14)a minimally contrasts with licit instances of Focus Preposing:

- (15) *DE LA MARIA, vam parlar ahir.*
of the Maria PAST-2PL talk yesterday
'We talked ABOUT MARIA yesterday.'

Moreover, (14)b has the following counterpart involving Hanging Topic Left Dislocation, a construction we will take into account in 2.1.2:

- (16) *La Maria, ara sí que parlaran d'ella.*
the Maria now yes that talk.FUT.3PL of her
'Maria, they WILL now talk about her.'

It is worth noting that the use of a clitic pronoun with CLLD (the same holds true for CLRD, see 3.1.1) is constrained by several factors both within and across languages. First of all, not all constituents must be replaced by a clitic. Consider the quite complex case of Catalan non-selected PPs. Some instances do not require the presence of the clitic:

- (17) a. *Amb un pot de greix (hi) hem greixat el motor.*
with a can of grease (LOC) have-2PL lubricated the engine
- b. *En aquesta sala, la Judit (hi) toca l'arpa. [Solà 1990: ex. 41b]*
'In this room, Judit plays the harp.'

However, in other cases, what seems optionality implies indeed two different interpretations, as the following contrast from Solà (1990: ex. 39), which I provide with his (translated) glosses in the prime examples:

- (18) a. Al despatx, atenc els clients.
'At the office, I attend to clients.'
a'. 'When at the office, I attend to clients.'
b. Al despatx, hi atenc els clients.
'At the office, I attend to clients.'
'I attend to clients precisely in this office (we were talking about).'

Several equivalent examples easily come to mind, as the classical ones due to Gabriel Ferrater:

- (19) a. En aquest pis, no toca el sol.
'In this apartment, (s)he doesn't touches the sun.'
b. En aquest pis, no hi toca el sol.
'In this apartment the sun doesn't get in.'

Intuitively, the preposed PP without a resumptive clitic receives an interpretation close to that of free adjuncts. Consider a clear case from Hernanz (1993: ex. 32):

- (20) a. Amb aquest professor, tothom treballa.
'With this teacher, everybody works.'
b. Amb aquest professor, tothom hi ha treballat.
'This teacher, everybody has worked with.'

(20)a is a free adjunct with a conditional flavor, which doesn't have a resumptive element within the sentence. This is confirmed by the possibility of having

- (21) Amb aquest professor, tothom treballa amb tothom.
'With this teacher everybody works with everybody.'

Here no place is left for the free adjunct within the sentence, which suggests we are not dealing with CLLD —in (20)b, instead, the presence of the clitic ensures the connectivity of the fronted PP, which is interpreted as a comitative adjunct.⁴ The issue is complex enough to deserve a more detailed study, which we cannot attempt here. See Bartra 1985, 2.3.3.1-2, Fernández Soriano 1989, 1993, Hernanz 1993, Solà 1977: 25-26, 1990: 112ff, 1994: 14.2.

If we make a step further we find instances of CLLD where no clitic is even possible. Take the case of VPs, which are replaced by the pro-verb *fer-ho* ‘do it’:

- (22) a. Menjar espinacs, cap nen no vol fer-ho.
eat spinachs no child not wants do-it
- b. *Jugar amb l'ordinador*, ho fan tots els nens.
play with the computer it do-they all the children

⁴ In principle, a way to test this analysis would be to add a clear CLLDed element to the structure. Since the ordering of the dislocates is free, as we have seen above, if the PP were a CLLDed element it would allow a permutation with the other clitic left dislocate. *Prima facie*, this prediction is borne out:

- (i) a. Amb aquest professor, els exercicis, els vaig fer a temps.
‘With this teacher, I did the exercises on time.’
- b. Els exercicis, amb aquest professor, els vaig fer a temps.
‘With this teacher I did the exercises on time.’

Nevertheless, this might be a premature conclusion. We positively know about the extreme flexibility of placement of free adjuncts (I indicate with ___ the positions where the PP may surface):

- (ii) (Amb aquestes condicions) la Maria ___ no pot ___ treballar ___ en la tesi ___.
‘With such conditions, Maria cannot work on her thesis.’

None of these placements produce any change in the truth-conditions of the sentence. CLLD, in contrast, do not have such a freedom.

Another difference is that free adjuncts license an arbitrary second person subject (see Hernanz 1993), whereas CLLD doesn't:

- (iii) a. Amb aquest professor, treballes.
‘With this teacher, you work/one works.’
- b. Amb aquest professor, hi treballes.
‘With this teacher, you work/*one works.’

These differences point to a separate analysis of both constructions, which should extend to cases with adverbs like

- (iv) D'aquest tema, *històricament*, a les aules en parlaven molt.
of this subject historically in the classrooms of.it talked-they much
‘Historically people talked much about this subject in the classrooms.’

Secondly, there is a great deal of cross-linguistic variation with respect to the pronominal system available for CLLD. Catalan, like French (see Larsson 1979) or Sardinian (see Jones 1993), is a rather extreme example of a rich clitic system with a pervasive clitic resumption strategy. However, Italian, another language with a rich clitic system, seems to allow CLLD without clitic more easily. Compare (all Italian examples are from Benincà 1988: ex. 128):

- (23) a. A Giorgio, (gli) darò un libro.
a'. Al Joan, *(li) donaré un llibre.
'To John, I'll give a book.'
- b. Con Giorgio, non (ci) esco mai.
b'. Amb el Joan, no (hi) surto mai.
'With John, I never go out.'
- c. Dei suoi lavori, non (ne) parla volentieri.
c'. De les seves feines, no (en) parla de grat.
'About his jobs, (s)he doesn't talk willingly.'
- d. A casa, (ci) andrò domani.
d'. A casa, *(hi) aniré demà.
'I'll go home tomorrow.'

Another extreme case is found in languages lacking a rich clitic system, like Spanish. Compare the preceding Italian and Catalan sentences with their Spanish counterparts:

- (24) a. A Juan, *(le) daré un libro.
b. Con Juan, no salgo nunca.
c. De sus trabajos, no habla con gusto.
d. A casa, iré mañana.

Romance languages lacking prepositional clitics may even resort to definite clitics to reassume indefinite NPs. Consider the following paradigm based in Longa et al. (1998: ex. 32):

- (25) A: Were there witches?
- B: a. No, no *n'hi havia*. (Catalan)
not not PART-LOC had
- b. No, nun *les* había. (Asturian)
not not them had
- c. Non, non *as* había. (Galician)
not not them had
- d. No, no (*las*) había. (Spanish)
not not them had
'No, there weren't any.'

See Longa et al. (1998) for details. See also Chapter 3.1.1.

VI) *The dislocate appears to the left of wh-elements but to the right of the complementizer:*

- (26) a. De la Maria quan en parlarem?
of the Maria when of.it talk-FUT-1PL
'When will we talk about Maria?'
- b. No sé pas *de la Maria* quan *en* parlarem.
not know-1 NEG of the Maria when of.it talk-FUT-1PL
'I do not know when will we talk about Maria.'
- c. Tinc un amic amb el qual, *d'aquests temes*, no *en* parlem mai.
have-1 a friend with the which of-these subjects not of.it talk-1 never
'I have a friend with whom we never talk about these subjects.'
- d. (És increíble,) *d'aquestes cadires*, el preu que *en* podem treure.
is incredible of-these chairs the business that of.it can-2PL take
'(It is incredible) which price we can obtain by these chairs.'
- (27) a. Em va dir que, de la Maria, no en parlarem mai.
to.me PAST-3 that of the Maria not of.it talk-FUT-1PL never
'(S)he told me that we will never talk about Maria.'
- b. No vindré perquè, *d'aquests temes*, no *en* parlem mai.
not come-FUT-1 because of-these subjects not of.it talk-1 never
'I am not coming because we never talk about these subjects.'

This behavior is common to all languages having CLLD (e.g., see Jones 1993 for Sardinian or Aissen 1992 for Mayan language Tz'utujil). However, beyond this common observation, there is no consensus on the proper analysis. Two main proposals have been suggested.

Some authors have argued that dislocates occur within the CP. Beyond this point, many differences arise: movement of the topic to [Spec, CP], as defended by Aissen (1992) for Tz'utujil or by Svolacchia et al. (1995) for Somali; base-adjunction of the topic to the CP, and movement of a coindexed null operator to [Spec, CP], as in Chomsky (1977) for English Left Dislocation or in Cinque (1983, 1990) for Italian CLLD. Regardless of the exact details of each proposal, all of them are committed with CP-recursion in order to offer an explanation to the fact that topics occupy an intermediate position between the complementizer and interrogative *wh*-elements.

The other alternative proposes that topics are adjoined to IP. Again, differences arise as to whether the topic is base generated in this position or moved (Lasnik & Saito 1992 for English topicalization, Rochemont 1989, Vallduví 1990).⁵ This proposal has to assume that interrogative *wh*-elements are not in [Spec, CP], but rather in [Spec, IP].

More recently, Rizzi (1997) develops a split CP area, where a Topic Phrase dominates the projection hosting focused elements (his *Focus Phrase*) but it is dominated by the projection hosting complementizers (his *Force Phrase*). A similar proposal is found in Villalba (1997, to appear). We'll remain neutral with respect to this point until Chapter 4, where a more articulated vision of the sentence architecture is proposed.

To wind up this paragraph, consider a fact reported by Fontich (1998:§7.1). This author notes that clitic dislocates do not surface to the left of the interrogative marker *que* lit. 'that', which may head yes/no questions with a confirmatory character (even though the prime versions, where the clitic left dislocate follows the interrogative marker, are marked equally as bad as the other sentences, it must be noted that speakers have the strong feeling that the later are clearly less ill-formed):

⁵ Specific proposals are found in other frameworks, e.g. Vallduví & Engdahl (1996), Engdahl (1999), Sanfillipo (1998) in a HPSG framework; Butt & King (1997) in LFG; Choi (1996, 1997) in OT. A thorough comparison would exceed the limits of this work.

- (28) a. *De la poesia bucòlica etrusca, que en parlarem avui?
of the poetry bucolic Etrurian Q of-it talk-FUT-1PL today
‘Etrurian bucolic poetry, will we talk about today?’
- a’. *Que, de la poesia bucòlica etrusca, en parlarem avui?
- b. *Amb la Maria, que no hi penses parlar?
with the Maria Q not LOC think-2 talk
‘Mary, aren’t you going to talk with her?’
- b’. *Que, amb la Maria, no hi penses parlar?

He further notes that the slightly modified version in (29) is fully acceptable:

- (29) a. I de la poesia bucòlica etrusca, que no en parlarem avui?
‘And Etrurian bucolic poetry, will we talk about today?’
- b. I amb la Maria, que no hi penses parlar?
‘And Mary, aren’t you going to talk with her?’

Here the topic is associated with an interrogative intonational pattern and appears separated off the sentence by a pause. These sentences pose us two problems. First of all, it has to be explained why a contrast exists between (28) and (29). Secondly, we have to explain why an apparently clear case of CLLD is associated with such special intonational pattern. My proposal is that we are not dealing with a topic construction, but rather with two different sentences, the former being elliptical. Let me put forward my argument. First of all, this construction conveys a confirmatory character: the speaker essentially demands for the validity of applying a previous assertion to a different entity. This becomes clearer when integrating the sentences in a dialogue:

- (30) a. A: Avui parlarem de la literatura satírica anglosaxona.
‘Today we will talk about Anglo-Saxon satiric literature.’
- B: I de la poesia bucòlica etrusca, que no en parlarem avui?
‘And Etrurian bucolic poetry, aren’t we talking about today?’
- b. A: Serà millor que parli amb en Carles.
‘It will be better that I talk with Carles.’
- B: I amb la Maria, que no hi penses parlar?
‘And Mary, aren’t you going to talk with her?’

Take (30)a. Here we have a property $\lambda x[\text{talk.about}(x)]$, for $x = [\text{Anglo-Saxon satiric literature}]$, and we are asking for the validity of applying this property to $x = [\text{Etrurian bucolic poetry}]$. This special relation would be the determining factor for explaining the obligatoriness of the conjunction *i* ‘and’, namely the contrast between (28) and (29), and the unexpected intonational contour. Furthermore, many other unexpected facts follow straightforwardly. First of all, it becomes apparent why this kind of sentences are impossible in an out-of-the-blue context. Secondly, the analysis predicts that the content of the open sentence is presupposed, a prediction that receives strong confirmation from the incoherence of a dialogue like the following:

- (31) A: Serà millor que parli amb en Carles.
 ‘It will be better that I talk with Carles.’
 B: #I amb la Maria, que no hi penses anar al cine?
 ‘And Mary, aren’t you going to go to the cinema with her?’

The necessity that the open sentence be presupposed would explain the fact that the open sentence preceded by the topic can be omitted without any semantic loss:

- (32) a. A: Avui parlarem de la literatura satírica anglosaxona.
 ‘Today we will talk about Anglo-Saxon satiric literature.’
 B: I de la poesia bucòlica etrusca?
 ‘And about Etrurian bucolic poetry?’
 b. A: Serà millor que parli amb en Carles.
 ‘It will be better that I talk with Carles.’
 B: I amb la Maria?
 ‘And with Mary?’

To sum up, the construction under question is not an instance of CLLD, but rather the juxtaposition of two sentences. Graphically:

- (33) a. I de la poesia bucòlica etrusca? Que no en parlarem avui?
 ‘And what about Etrurian bucolic poetry? Will we talk about it today?’
 b. I amb la Maria? Que no hi penses parlar?
 ‘And what about Mary? Aren’t you going to talk with her?’

VII) *There is obligatory connectedness between the dislocate and the resumptive element:*

The term *connectedness* comprises a number of different syntactic phenomena, which reflect the fact that, to put it in Cinque's (1983) own words, "the lefthand phrase [...] behaves with respect to the various syntactic and Logical Form (LF) principles as if it occupied the position actually filled by the resumptive clitic pronoun within the associated S." This amounts to saying that the dislocate and the clitic must share Case and categorial features, and theta role:

- (34) a. Crec que *jo/a mi, no em van dir res del tema.
think-1 that to me/I not to.me PAST-3PL say anything of-the subject
'I think that they said me nothing about the subject.'
- b. Crec que *(a) la Maria, no li van dir res del tema.
think-1 that to the Maria not to.her PAST-3PL say anything of-the subject
'I think that they said Maria nothing about the subject.'

In connection with this, Cinque examines the behavior of CLLD with respect to binding principles. It seems that dislocates satisfy the binding relations they would have maintained if in situ:

- (35) a. La Maria₁ no confia en ella mateixa₁.
b. En ella mateixa₁, la Maria₁ no hi confia.
'Maria doesn't rely in herself.'
- (36) a. *La Maria₁ la₁ va veure a ella₁ de seguida.
b. *A ella₁, la Maria₁ la₁ va veure de seguida.
'*Maria saw her immediately.'
- (37) a. *pro₁ va veure el fill de la Maria₁ de seguida.
b. *El fill de la Maria₁, pro₁ el va veure de seguida.
'*She saw Maria's son immediately.'

However, it is very unlikely that this is a genuine connectedness effect. Consider for the sake of illustration a minimal variation of (35):

- (38) a. No confia en ella mateixa la Maria.
 ‘It is Maria who doesn’t rely in herself.’
 b. *En *ella mateixa*, no hi confia *la Maria*.
 ‘Herself, Maria doesn’t rely in.’

The difference has to do with the informational status of the binder: it is part of the presupposition in (35), but the focus in (38). The fact that only the presupposed referent may bind the anaphora suggests that we are dealing with an instance of discourse binding: the anaphora within the dislocate is bound by a previous mention of the referent *Maria*. This intuition seems confirmed by data involving pronoun binding:

- (39) a. *A ell_{1/2}*, en Pere₂ creu que no *li* diran res.
 to him the Pere thinks that not to.him/her say-FUT-3PL anything
 ‘To him, Pere thinks that they will say nothing.’
 b. *Amb ella_{1/2}*, la Maria₂ sospitava que no *hi* volien parlar.
 with her the Maria suspected that not LOC wanted-3PL talk
 ‘Her, Maria suspected that they didn’t want to talk with.’

The dislocated pronoun may corefer with the subject without yielding a violation of Principle C. However, this coreference is only possible if the referent of the name has already been introduced in the discourse. Note what happens if the name is focused:

- (40) a. *A ell_{1/*2}}*, és en Pere₂ qui creu que no *li* diran res.
 ‘To him, it is Pere who thinks that they will say nothing.’
 b. *Amb ella_{1/*2}}*, és la Maria₂ qui sospitava que no *hi* volien parlar.
 ‘Her, it is Maria who suspected that they didn’t want to talk with.’

Note that the explanation of such binding facts cannot be based on a presumed reconstruction of the CLLDed element into its original position: the name can be coreferent with the clitic pronoun both in (39) and (40). A more promising approach resorting to discourse binding relations is pursued in Villalba (forthcoming).

VIII) *The relation between the dislocate and the clitic is subject to (strong) island constraints* (I conventionally mark islands with braces):

- (41) a. **D'aquest llibre, la Maria va preparar el sopar {i en Pere en va parlar}*.
of-this book, the Maria PAST-3 prepare the dinner and the Pere of.it PAST
talk
- b. **De ric, {ser-ho} no és fàcil*.
of rich be-it not is easy
- c. **De Kant, la Maria va veure {la noia que en parla a la tesi}*.
of Kant the Maria PAST-3 see the girl that of.it talks in the thesis
- d. **De Kant, la Maria confia en {mi, que no en parlo mai}*.
of Kant the Maria relies in me that not of.it talk-1 never
- e. **D'aquest llibre, has de fer el sopar {si en Pere en parla}*.
of this book have-2 of do the dinner if the Pere of.it talks
- f. **D'intel·ligent, la Maria va sortir de la universitat {sense ser-ho més}*.
of-intelligent the Maria PAST-3 come.out of the university without be-it
more

(41)a violates the Coordinate Structure Constraint; (41)b, the Sentential Subject Constraint, (41)c-d, the Complex NP Constraint; (41)e-f, an adjunct island.

All these facts are well-studied (see Cinque 1977, 1983, 1990, Iatridou 1991, Postal 1991), and there is consensus on the independent status of CLLD with respect to the rest of constructions studied in this chapter, especially Hanging Topic Left Dislocation.

2.1.2. Interpretive properties

CLLD is a well-studied syntactic construction. Since Cinque's seminal work (Cinque 1977, 1983, 1990), many works have been devoted to deciding whether it is base generated or it is the result of movement, whether it licenses parasitic gaps or whether it shows WCO effects. However, less attention has been paid to its impact on the interpretation of sentences, mainly due to the fact that it is unclear whether CLLD has any effect on truth-conditions. In this section I'll work out a description of the factors involved in the interpretation of CLLD. The presentation is structured in a

60

twofold way. Firstly, attention will be paid to the factors determining which NP/DP are suitable for CLLD. As will become clear, neither definiteness nor specificity is the property underlying dislocability. As a consequence, the section will continue showing the pragmatics and semantic properties of the construction are better understood when contemplated in the context of discourse relations, and the informational and semantic articulation of sentence.

Mainly based on English topicalization, a huge amount of works have raised examining the restrictions imposed on detached NPs. Three main features have been argued to be at the roots of dislocability, namely definiteness, specificity and ‘poset’ relations—which will be discussed in part B). Consider first definiteness. Even though the literature shows that definite specific NPs, including proper names and personal pronouns, are optimal dislocates, we do have definite NPs which are not allowed in CLLD, like Fauconnier’s (1975) quantificational superlatives (see Postal 1991 for French):

- (42) a. *No feu cap crítica perquè la més petita crítica, no és capaç de suportar-la.
 ‘Don’t make any criticism because she can’t stand any criticism.’
 b. *Abaixeu la música, que en Pere, *la més petita fressa*, no és capaç de suportar-la.
 ‘Turn down the music, that Pere can’t stand the slightest noise.’

Since quantificational superlatives are semantically quantifiers even though formally definite NPs, these examples can be put together with those of dislocated quantifiers that we’ll take into account below. A common property is that these examples ameliorate when interpreted partitively, that is, as ranging from a previously established set:

- (43) a. Serà millor que no digui res, perquè la més petita crítica sortida dels meus llavis, la consideraria un insult.
 ‘It would be better I say nothing, because the slightest criticism coming out from my lips, he’ll find an insult.’
 b. Abaixeu la música, que en Pere, *la més petita fressa que no sigui seva*, no és capaç de suportar-la.

'Turn down the music, that Pere can't stand the slightest noise not produced by him.'

Whichever the reason for the contrast is, it cannot be definiteness, since we are dealing with definite NPs in both cases.

Finally, we do have definites including referentially dependent elements, like *altre, diferent, mateix* (see Culicover & Jackendoff 1995):⁶

(44) A: Tot i tenir molts llibres sobre el tema, en Pere va recomanar els llibres d'un altre.

'Even though he has many books on the subject, Pere recommended someone else's books.'

B: **Doncs els llibres d'un altre, jo no els he recomanat mai.*

well the books of-a other I not them have-1 recommend never

'Well, I have never recommended someone else's books.'

(45) A: M'he comprat una màquina de fer fotos Leica.

'I've bought a Leica camera.'

B: **La mateixa màquina de fer fotos, me la vaig comprar l'any passat.*

the same machine of make photos to.me her PAST-1 buy the-year passed

'Last year I bought the same camera.'

With these examples in mind, it seems clear that having the feature [+definite] cannot be a sufficient condition for being a dislocate.

Another suggested property underlying dislocation is specificity (see Escobar 1995: 123ff —who attributes the idea of linking CLLD and specificity to María Luisa Zubizarreta—, Quer 1993, Roca 1992: 2.3, Solà 1992: 281, or Villalba 1994: 3.3.1.1), whatever it exactly means for each.⁷ It is beyond discussion that there are clear instances of non-specific indefinite and bare NPs that cannot be dislocated:

⁶ Interestingly enough, these sentences improve when in a generic tense (I owe this observation to Gemma Rigau):

(i) La dona d'un altre no te l'has de tirar mai.
'Someone else's wife you should never screw.'

I have no principled explanation for this behavior.

⁷ Escobar (1995) makes a reference to Enç (1991), without further details. Roca (1992) uses 'definite' and 'specific' somewhat ambiguously, without giving any definition of the latter. Quer (1993) uses the

(46) A: Necessito una empresa de neteja.

'I need a cleaning company.'

B: a. *Una empresa (de neteja) qualsevol, en Pere també la busca.

'Any (cleaning) company, Peter is looking for as well.'

b. En Pere també busca una empresa (de neteja) qualsevol i espera trobar-la aviat.

'Peter is also looking for any (cleaning) company and he hopes he'll find it soon.'

(47) A: Tinc molta set, no tens cervesa?

'I am very thirsty, don't you have beer?'

B: a. *Cerveses, tu mateix les pots agafar de la nevera.

'Bottles of beer, you can get them from the refrigerator by your own.'

b. A la nevera hi ha cerveses: tu mateix les pots agafar.

'There are bottles of beer in the refrigerator, you can get them by your own.'

Note that, as the examples in b. show, the ill-formedness of CLLD, doesn't seem to follow from a presumed incompatibility between the indefinite NP and the clitic (against the conclusion in Roca 1992: 2.3).

However, it is also a clear fact that indefinite *specific* NPs are not dislocable either.⁸ Consider Bosque's (1994) observation that indefinite nominals modified by assessment degree adjectives (i.e. superlative adjectives, extreme degree adjectives,

term specific without qualifications. Solà i Pujols (1992) uses 'weak', 'D-linked', and 'referential' interchangeably. Villalba (1994) explicitly assumes Enç's (1991) definition of specificity.

⁸ Larsson (1979: 28) claims that dislocated indefinites can only receive a generic interpretation, not a specific one —she doesn't even mention the possibility of a non-specific interpretation. Postal (1991), follows her description of the facts and offers as evidence the following examples, where the presence of *certain* 'certain' ensures specificity:

(i) a. *A certains hommes, je ne leur parle jamais.

'To certain men, I never speak.'

b. *A certaines idées bizarres, il y a fait allusion.

'To certain strange ideas, he referred.'

It is true that the Catalan equivalents of (i) are equally odd, but the addition of a generic context makes them perfect:

(ii) a. Amb certes persones, és millor no parlar-hi: per exemple, en Pere.

'With certain people, it is better not to speak: for example, Pere.'

b. De certs temes, tothom en pot parlar: per exemple, de futbol.

'About certain subjects, everybody can talk: for example, about football.'

and non-restrictive prenominal adjectives) are interpreted as specific (see also Bosque & Picallo 1996, and Picallo 1994, for a more developed analysis). Bosque tests this observation in several contexts (Bosque's 1994 exs. 56, 58, 59 and 60):

(48) a. A las siete siempre había un programa {interesante/interesantísimo} en la televisión.

'At seven o'clock there was always {an interesting/a very interesting} show on TV.'

b. Todo el que tiene un burro {lento/?*lentísimo} lo golpea.

'Everyone who has {a slow/an extremely slow} donkey beats it.'

c. Recuerdo haber leído alguna que otra {novela interesante/?*interesante novela} de ese autor.

'I remember to have read one interesting novel or another by this author.'

d. Debería escribirse un {drama rural conmovedor/?*conmovedor drama rural}.

'A touching rural drama should be written.'

In (48)a the nominal modified by the excess degree adjective, unlike its counterpart, is not under the scope of the adverbial quantifier *siempre* 'always', which suggests it is specific. The same conclusion follows from the other examples, which introduce non-specific contexts, hence the incompatibility (see Bosque 1994: section 6 for details). Furthermore, Bosque shows the intuition that such indefinites cannot be interpreted as generics is coherent with their oddity in generic contexts (see Bosque 1994: section 6):

(49) a. *Generalment, un magnífic ordinador no acostuma a donar molts problemes.

'Generally, a magnificent computer doesn't tend to provoke a lot of problems.'

b. *Generalment, un bon escriptor sol escriure una interessant novel·la cada dos anys.

'Generally, a good writer usually writes an interesting novel every two years.'

It becomes now apparent that the nominals modified by the excess degree adjective are especially relevant to our discussion on the dislocability of specific indefinites: they are unambiguously specific and they don't admit a generic reading which might lead to unwanted interpretations. Consider:

(50) A: Les novel·les realment interessants es venen sense necessitat de gaire publicitat.

'Really interesting novels sell without resorting to much publicity.'

B: a. *D'una interessant novel·la, no cal parlar-ne gaire: la gent no és tonta.

b. D'una novel·la interessant, no cal parlar-ne gaire: la gent no és tonta.

'An interesting novel, you mustn't talk much about: people are not stupid.'

(51) A: En Pere té un ordinador, però és molt lent.

'Pere has a computer, but it is very slow.'

B: *Doncs, jo, un magnífic ordinador, me'l vaig comprar per dues-centes mil.

'Well, a magnificent computer, I bought for two hundred thousand.'

As the examples make clear, even though the indefinites in question are known to be specific, they cannot be dislocated. It seems clear, then, that specificity is not a sufficient condition for being a dislocate.

Nevertheless, we do have instances of indefinites that can be dislocated. Note the following examples with specific indefinites, due to Solà (1992: 270), which I provide with his glosses:

(52) a. Una sargantana, l'he vista al jardí.

'One (of the)lizard(s), I saw in the garden.'

b. (A) un/algun malalt el visiten tots els metges.

'A/Some patient is visited by all the doctors.'

Consider (52)a. Solà correctly observes that this sentence is an impossible answer to 'What happened?', which amounts to saying that the referent of the clitic left dislocated NP has previously been introduced in the discourse or it is pragmatically available. We can imagine this sentence in a context like:

(53) A: Aquestes sargantanes són molt valuoses i les tinc sempre en aquesta habitació.

'These lizards are very worthy and I always have them in this room.'

B: Doncs vigila, perquè una sargantana l'he vista al jardí.

'Take care then, since one (of the)lizard(s), I saw in the garden.'

It becomes clear now that the indefinite can only be understood partitively, as Solà's gloss reflects, which is confirmed by the fact that the most natural answer in this context would be an overt partitive construction:⁹

(54) a. B': Doncs vigila, perquè una de les sargantanes l'he vista al jardí.

'Take care then, since one of them, I saw in the garden.'

b. B'': Doncs vigila, perquè n'he vist una al jardí.

'Take care then, since I saw one (of them) in the garden.'

⁹ Interestingly enough, the sentence is less natural with a clitic right dislocate:

(i) ^{??}Doncs vigila, perquè l'he vista al jardí, una de les sargantanes.
'Take care then, since I saw it in the garden, one of them.'

This fact seems to be related to the following contrast noticed by Larsson (1979) for French, and reproduced for Catalan by Bartra (1985), from who I borrow the French and Catalan examples, respectively (a context has been added to make the point clearer):

- (ii) [Have you seen all the black horses in the rancho?']
a. [?]De cheval noir/De chevaux noirs, je ai vu le vôtre.
'I have seen YOUR black horse.
b. Je ai vu le vôtre, de cheval noir/*de chevaux noirs.
'I have seen YOUR black horse.
- (iii) [Have you seen all the blue cars in the parking?']
a. De cotxe blau/De cotxes blaus, només he vist el teu.
'I have only seen YOUR blue car.
b. Només he vist el teu, de cotxe blau/*de cotxes blaus.
'I have only seen YOUR blue car.

Larsson (1979: 34) argues that only the (b) sentences are instances of (right) dislocation involving coreference (hence the obligatoriness of agreement). According to her, the (a) sentences are instances of Cinque's hanging topics, i.e. HTLD (hence the optionality of agreement). I agree with Larsson that the non-agreeing NPs must be analyzed as instances of HTLD, however, I think she is wrong in extending such an analysis to *both* kinds of NP. As the following example shows, the agreeing NP allows, and the non-agreeing NP, rejects a leftmost topic:

- (iv) [Have you seen all the blue cars in the parking?']
En aquest aparcament, de cotxe blau/*de cotxes blaus, només he vist el teu.
'In this parking, I have only seen YOUR blue car.

As is well-known (see Benincà et al. 1988, Cinque 1983, 1990), HTLD doesn't allow recursion nor a topic to its left. If Larsson were right, the agreeing NP should be odd, which is not. The correct generalization is that the agreeing NPs are clitic left/right dislocates, whereas the non-agreeing ones are instances of HTLD, which is known to have no right counterpart.

So then, it seems that a partitive reading is necessary somehow to rescue a specific dislocated indefinite. This correlates with the behavior of dislocated QPs:

- (55) a. *A cap alumne no l'he vist avui.
'Any student, I haven't seen today.'
- b. A cap d'aquests alumnes de què em parles no l'he vist avui.
'Any of those students you talk me about, I haven't seen today.'
- c. *A cada alumne l'he vist avui.
'Every student, I have seen today.'
- d. A cada un dels teus alumnes l'he vist avui.
'Each of your students, I haven't seen today.'
- e. *Res (no) ho havia fet ell.
'Nothing, he had made.'
- f. Res del que et va ensenyar (no) ho havia fet ell.
'Nothing of the things he showed to you, he had made.'

Clearly enough, the partitive interpretation makes CLLD possible for QPs otherwise excluded from this position. The customary interpretation of this fact is that partitivity entails specificity—which in turn would allow for dislocability. This is in essence the way Enç (1991: 21) puts it: “specificity involves linking objects to the domain of discourse in some manner or other. One acceptable way of linking is through this assignment function, by relating objects to familiar objects. Another acceptable way of linking is the subset relation, which we have observed in covert and overt partitives.” Given this, the reasoning would be that partitivity entails specificity, and specificity entails dislocability. However, we have already seen that specificity is not a suitable candidate for explaining dislocability. The answer to the puzzle lies, I think, in pursuing the consequences partitivity has for discourse relations. Two aspects of Enç’s approach are worth commenting. First of all, ‘linking objects to the domain of discourse’. Besides the obvious formal resemblance, this notion is quite close to the ‘link’ of Vallduví (1990). Vallduví develops a pragmatic model for explaining the way information is structured in sentences, independently of its propositional content. He calls this model *informatics* and it has a major goal

formalizing *information packaging*.¹⁰ To do so, he resorts to the file-metaphor suggested in Reinhart (1981) and modeled in Heim (1982). According to Vallduví, CLLD is a paradigmatic syntactic incarnation of one of his primitive informative functions: a *link*, which is defined as an instruction for updating the information conveyed by the focus under the file denoted by the CLLDed element.¹¹ Consider:

- (56) a. El llibre, l'he comprat a Can Viader.
 the book him-have-1 bought to Can Viader
 'The book, I have bought at Can Viader.'
- b. A Can Viader, hi he comprat el llibre.
 to Can Viader LOC-have-1 bought the book
 'At Can Viader, I have bought the book.'

In Vallduví's model, (56) would be interpreted (informally) as the instruction 'go to the file with the heading *book* and add the information *I have bought x at Can Viader*'. (56), instead, conveys the instruction 'go to the file with the heading *Can Viader* and add the information *I have bought the book at x*'. This minimal pair shows that the same propositional content ('I have bought the book at Can Viader') is structured in two different ways, which is, according to Vallduví, a proof that information packaging is independent of the evaluation of truth-values.¹²

¹⁰ For an exposition of the goals and methods of information packaging, besides the classical reference —Vallduví (1990)—, it is worth consulting Engdahl (1999), McNally (1998), Vallduví (1994), and Vallduví & Engdahl (1996).

¹¹ Vallduví's notion of link is crucially based on a very particular pragmatic model that takes files as the basic tools for managing information. A function 'link' makes sense in such a system, since it serves as a pointer to connect files cards, ensuring efficient update of information. However, as Hendriks & Dekker (1996) point out, such a conception has to face serious theoretical and empirical problems. Consider, for example, the case of quantified or negative links:

- (i) [L Every/No man] [F WALKS]

It is difficult to imagine which file card these links usher to. In order to solve this problem (and other inadequacies), Hendriks & Dekker (1996) suggest that links are not functions ushering to a file card, but rather marking an instance of non-monotone anaphora, namely "if an expression is a link, then its discourse referent Y is anaphoric to an antecedent discourse referent X such that X M Y" (p. 353). This move allows the authors to integrate this different notion of linkhood within Discourse Representation Theory, a framework which offers a simpler representational space than that of files, and is hence less problematic from an ontological point of view. With these provisos in mind, the reader is requested to take *my* use of the notion link thorough this work as a convenient descriptive label, without commitment to any theoretical position.

¹² Since Strawson's (1964), it is a widely accepted fact that the topic-focus articulation has an impact on way we interpret sentences. What remains as a matter of controversy for both pragmatists and semanticists is whether the topic-focus articulation has any effect on the truth-conditions of sentences. Strawson (1964) considers that the evaluation of the truth-value of a sentence is commonly topic-