

The Emergence of Order in Syntax

Jordi Fortuny Andreu

Departament de Lingüística General
Universitat de Barcelona

Programa de Ciència Cognitiva i Llenguatge
Bienni 2001-2003



Signatura de l'autor _____

Director _____

Jaume Solà i Pujols
Departament de Filologia Catalana
Universitat Autònoma de Barcelona

Prologue

The main idea of this study can be expressed in very few words: the syntactic component of the Faculty of Language is responsible for ordering categories and for ordering categories only. This would be a completely uninteresting thought, a truism, if one did not attempt to account for how and why the attested patterns emerge from the external requirements that the syntactic component has to satisfy.

Although saying that syntax is responsible for order does no more than to express the etymological meaning of the word 'syntax', the use of the term 'order' in this study may deserve some attention, since it does not subscribe to the common use in grammatical studies. Throughout the text, the term 'order' does not exclusively refer to the literal precedence relation among terminals (this is the common use of the term in grammar), but rather to the hierarchical properties that are attributed to syntactic representations, as will later become clear. Literal precedence is mapped from hierarchy. Thus, the object of inquiry of the discipline called syntax is how categories are ordered or how hierarchies are generated.

More precisely, this study poses two questions: what are the basic elements of the syntactic component? and why do syntactic patterns have the shape they seem to have? The first question is addressed in Part I and the second question in Part II. Part III summarizes the conclusions of the preceding two parts and discusses the possibility that Universal Grammar (Chomsky's Factor II) is a rewiring of elements that are in place independently (Chomsky's Factor III).

Chapter 1 suggests that the basic elements of the syntactic component are features and a combinatorial operation known as Merge. A feature is defined as an instruction for a particular level of interpretation of the Faculty of Language and Merge is defined as an operation that takes as input two categories or sets and yields as output the union of these two sets. The specific instructions that functional categories provide for the several levels of interpretation are described. It is argued that the hierarchical properties of syntactic objects derive from a derivational record, a set K (a nest) where the outputs of Merge operations are linearly ordered by the strict inclusion relation. Consequently, Kayne's Linear Correspondence Axiom is no longer an axiom one needs to postulate to account for the X' -theory; hierarchy is a product of creating structure successively and keeping the derivational information in a record. Applications of both internal and external Merge have been argued to be triggered by the requirement of matching [+type] categories and [+token] categories, without postulating any special device for the property of displacement; it has also been argued that suicidal greed, a device postulated to account for movement specifically, shows up three problems: the problem of generality, the problem of determinacy and the problem of consistency.

Part II (chapters 2, 3, 4 and 5) discusses the particular shape of syntactic objects by considering how syntactic derivations carried out by Merge are triggered by the morpholexical characterization of lexical items (clusters of features) and constrained by the Full Interpretation semantic legibility condition and by the Maximize Matching Effects principle of structural minimization.

I argue that there are three types of syntactic patterns: discontinuous, analytic and syncretic. That is, it is possible for a feature to be assigned to two projections, for there to be a one-to-one relation between features and projections, and for more than one feature to be assigned to one projection. Thus the terms ‘discontinuous’, ‘analytic’ and ‘syncretic’ are relative to how features are associated with projections, or to how features are coded in the spine of a tree. It may be worth stressing that the use of the terms ‘analytic’ and ‘syncretic’ in this study differs from the use of the terms ‘analytic’ and ‘synthetic’ in morphological typology, which are relative to how features are associated with words. This is also true for the term ‘discontinuous’, commonly relative to the morpholexical level. In a very illuminating work I became aware of a few days before writing this prologue, Huang (2005a, b, 2006) explores the idea that multiple parameters such as the head-parameter, the pro-drop parameter, the *wh*-parameter, the telicity parameter and the configurationality parameter may derive from the distinction between analytic and synthetic languages in the abovementioned traditional sense, a distinction that Huang conceives in terms of an analytic *vs.* synthetic macro-parameter.

Chapter 2 illustrates the relationship between C and Infl. The so-called C-Infl link can be materialized in the following semantic and morphosyntactic properties and generalized to several Infl-like features such as ϕ -features, tense, mood, modality and negation.

Property (I): complementizers can ‘replicate’ Infl-like features

Property (II): there is a correlation between the characteristics of features surfacing on Infl and the choice of C

Property (III): Infl-like features are involved in triggering V-to-C movement

Chapter 3 discusses the source of the C-Infl link and concludes that C and Infl are polarities of the same feature ([\pm clause typing]) and that Infl-like features are assigned to both polarities, directly accounting for the three properties of the C-Infl link. The reason for this simultaneous insertion is that the semantic instruction provided by Infl-like features (to trigger referential displacement) is orthogonal to the [\pm clause typing] distinction. The v-V link and the P-K link are also analyzed as two syntactic discontinuities. The proposal that C and Infl constitute a discontinuous syntactic object is preceded by a relatively intricate evaluation of a plausibility argument: Chomsky’s (2005b) Feature Inheritance Theory, a subcomponent of the theory of phases, which can be generalized to Infl-like features. Consider the Phase-Impenetrability Condition.

Phase-Impenetrability Condition (PIC)

Consider a Phase $PH = [\alpha, [H, \beta]]$, H being the phase head
Call α and H the edge of PH , and β the domain of H

The domain of H is not accessible to syntactic operations beyond PH , only the edge, $\{\alpha, H\}$, since Transfer sends β to C-I and A-P once H has terminated its work (adapted from Chomsky 2001b: 5-6).

Briefly, if the characteristics of Infl-like features depend on the choice of C (semantic property II) and that Infl-like features can surface and be syntactically operative not only in Infl but also in C (morphosyntactic properties I and III), and it can be independently argued that there is an important asymmetry as to the role that C and T/Infl play in syntactic derivations, then it becomes interesting to consider the possibility that C, the head responsible for defining the relatively complete fragments of structure to be transferred from the syntactic workspace to the external systems, is precisely the locus where Infl-like features are base-generated, appearing only derivatively on Infl heads due to inheritance mechanisms. Consider the following sketch of the Generalized Feature Inheritance Theory, which is based on Chomsky's recent insights.

Generalized Feature Inheritance Theory (GFIT)

- (i) C has a central role in the generation of syntactic objects
- (ii) Infl does not have a central role in the generation of syntactic objects
- (iii) Infl-like features are generated on C
- (iv) In the lexicon, Infl lacks Infl-like features
- (v) Infl-like features surface on Infl only derivatively
- (vi) The feature inheritance mechanism has the function of bringing semantic distinctions into a syntactic representation

As will be shown below, the classical empirical argument for strict cyclicity is no longer valid, due to parallel probing at the phase level. And after thoroughly revising the Phase-Impenetrability Condition, it must be concluded that it is both too weak and too strong. It is too weak because penetrating into the edge of H is as problematic as penetrating into the domain of H (as Chomsky's refinement of Huang's subject islands show) and because not only A'-movement but also A-movement is successive cyclic. It is too strong because probing into the domain of H is possible when there is no intervention effect, as in experiencer constructions in languages like Icelandic and Catalan. In order to account for subject islands, *wh*-islands and the requirement that both A'-movement and A-movement are successive cyclic, the Phase-Impenetrability Condition has been replaced by the Relativized Opacity Condition.

Relativized Opacity Principle

In a syntactic object $[_{\alpha_1 P} [_{\alpha_1} \dots [_{\alpha_2 P} \Delta [_{\alpha_2} \Gamma]]]]$,

- where: (i) α_1 and α_2 are two probes of the same type α each projecting an αP
(ii) Δ is SPEC- α_2 and Γ is α_2 -COMPL, and
(iii) $\exists \delta$: δ is a constituent of Δ and $\exists \gamma$: γ is a constituent of Γ ,

α_1 can probe Δ or Γ if they provide a suitable token for α_1 , but it cannot readily probe γ or δ

The factor that determines whether α_1 can readily probe a goal is the relative depth of such a goal in the α_2 -projection whose label is of the same type as the searching probe α_1 . $\alpha_2 P$ does not render the complete Δ and Γ opaque to α_1 , but search in them becomes difficult. For this reason, long A-movement must use intermediate SPEC-T_{def} as an escape hatch and long A'-movement must use intermediate SPEC-Cs and SPEC-vs to attain its final position. The source of subject islands and *wh*-islands is the same: a probe α fails to search too deep into the SPEC or the COMPL of a lower α -projection.

Therefore, it seems that there is no clear asymmetry between C and T/Infl, and thus the Generalized Feature Inheritance Theory loses plausibility. Unless further arguments are provided, the simplest account for the C-Infl link is that Infl-like features are simultaneously merged in both poles of the C-Infl discontinuity, without lowering or raising mechanisms.

Chapter 4 considers the idea that natural languages privilege invariable and richly articulated representations with semantically devoted positions (cartographies or analytic syntactic patterns). I begin by revising Cinque's (1999) preliminary position on how cartographies are related to semantic constraints. Although Cinque's arguments in this respect may be scarce and too weak, it is clear that the empirical results of his work on clausal structure pose an interesting problem for the syntax-semantic interface. I propose that several distributional restrictions derive from the Prohibition Against Tangled Structures (relative to discontinuous patterns) and the Prohibition Against Vacuous Quantification.

Prohibition Against Tangled Structures (PATs)

Two discontinuous syntactic templates $[A_1 \dots A_2]$, $[B_1 \dots B_2]$ cannot be broken off by yielding the tangled object $[A_1 \dots [B_1 \dots A_1] \dots B_2]$

Prohibition Against Vacuous Quantification (PAVQ)

For every occurrence of a variable x there must be a quantifier Q binding x , and for every quantifier Q there must be a variable x such that Q binds x

A syntactic object that violates the PAVQ contains elements that are not useful to the Conceptual-Intensional system, namely a free variable or a quantifier that does not bind any variable. Similarly, a discontinuous template that is broken off automatically becomes

useless. Thus, the PATS and the PAVQ are two instances of the more general Full Interpretation Condition.

Full Interpretation Condition

An object generated by a grammatical component must be constituted only of useful features for a particular level of interpretation when it attains such a level of interpretation

The partial considerations that have just been summarized suggest the possibility that cartographies do not constitute a primitive component of Universal Grammar but a cluster of properties emerging from legibility conditions.

Finally I note that discontinuous syntactic objects do not contradict the idea that the C-I system requires articulated syntactic representations with semantically devoted positions, but rather they adhere to it: when an Infl-like feature is introduced into the two poles of the C-Infl discontinuity, both occurrences of this Infl-like feature are minimally distinguished, and thus they are not the same element in two positions but two covariations. Discontinuous syntactic objects are a subtype of analytic syntactic objects, or cartographic effects.

Chapter 5 argues that several intricate phenomena can be better understood if a less rigid view on how features are coded in a syntactic representation is adopted. More precisely, it has been affirmed that syncretic syntactic patterns (or anticartographic effects) also exist: they are favored by the Maximize Matching Effects Principle, a principle of structural minimization.

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

The MMEP does not apply arbitrarily, but under special circumstances favoring structural condensation to the detriment of the C-I legibility requirement of creating analytic syntactic patterns and discontinuous syntactic patterns. Consider the following paragraphs as a simplified outline of the rather intricate course developed in this chapter.

If a richly inflected verb matches φ -features on T, there is no reason to claim that A-movement is necessary. If it is not necessary, it is impossible, under efficiency considerations. In this circumstance, SPEC-T is available to match other features, and it should not be used to match φ -features. This accounts for the A'-status of preverbal subjects in Null Subject Languages: matching φ -features with V-movement renders SPEC-T as an A'-position, to be filled by topical/referential subjects or by (topical/referential) CLLD elements (Solà 1992).

There is no reason to think that preverbal subjects in V2 languages have moved to a peripheral position. Thus, V should be on T, not on C. This is substantiated by Zwart's (1993) evidence from double agreement dialects: V shows up with the agreement mark characteristic of complementizers in inversion contexts and a different mark when the first

constituent is a subject; the mark in this second case is the same that V shows up with when it remains inside the VP (that is, in embedded clauses headed by a complementizer). Call the former type of agreement C-agreement and the latter Infl-agreement. Thus, the question is why V does not raise to C to check tense when the first constituent is a subject, yielding an inversion pattern. The proposed answer is, again, expressed in terms of the MMEP: syntax proceeds in an opportunistic way taking advantage of the discontinuous nature of tense; if tense is matched in TP as φ -features, all the computations take place in the same projection and CP is not initiated to match a feature that, due to its discontinuous nature, can be indistinctively matched in C or in T.

In English, *wh*-island effects are weaker when the embedded *wh*-item is nominative (Chomsky 1986, Chung & McCloskey 1983). This observation suggests that the nominative interrogative is in SPEC-T and that SPEC-C of the embedded clause is available as an escape hatch. Similarly, in Dutch, focal and topical 2SG subjects are followed by a V with Infl-agreement, not with C-agreement; therefore, V never raises to C in Dutch when the first constituent is a subject, regardless of whether such a subject has peripheral features or not; this suggests that focal/topical subjects in Dutch may also be in SPEC-T, as English nominative interrogatives. In these cases, SPEC-T becomes a mixed A'-A-position. The special circumstance favoring the application of the MMEP is that both φ -features and π -features (*wh*-feature and topic/focus-feature) are matched by the same category.

This chapter closes with a discussion of *that*-trace effects in English. If SPEC-T is the position where *wh*-features and φ -features are matched by a nominative interrogative, then SPEC-T may also be the position where the defective peripheral features triggering movement to intermediate positions are located. If both peripheral features and non-peripheral features are located in TP, there is no reason anymore to maintain the C-T distinction: TP has become a [\pm clause typing] category. Therefore, the complementizer cannot be inserted when the *wh*-phrase undergoing long-movement is nominative because there is no structural space. This reasoning also accounts for the more complicated patterns of *that*-trace effects discovered by Bresnan (1977) and also discussed by Rizzi (1997), among others.

Part III (chapter 6) provides a summary of the general line of argumentation and brings Chomsky's (2005a) three factors into the discussion. The idea put forward is that, with regard to the syntactic component, UG is a rewiring of elements that must be in place independently: features, Merge and principles of analysis and data processing.

I would like to express my gratitude to Víctor Acedo, Noam Chomsky, Bernat Corominas, Txuss Martín, Josep Quer, Joana Rosselló, Jaume Solà and Jeroni Tutusaus, who have been an endless source of inspiration. I would also like to thank the members of my dissertation committee, Itziar Laka, Carme Picallo, Gemma Rigau, Joana Rosselló and Jan-Wouter Zwart, for contributing to the improvement of this work with their comments. I am also indebted to Cedric Boeckx, for his support and insistence on making this text more readable, and to Emily Toder for her invaluable help in proofreading the manuscript. All the shortcomings in the conception, development and expression of the ideas are mine alone and reflect the limits of my current possibilities.

Table of contents

PART I. INTRODUCTION

Chapter 1. Elements of syntax	11
1.1 Instructions	11
1.2 Merge: nests	15
1.3 Merge: internal and external	19
1.4 Onset	24

PART II. PATTERNS

Chapter 2. The relationship between complementizers and inflectional heads	26
2.1 The C-Infl link	26
2.1.1 φ -features	26
2.1.2 Tense	30
2.1.3 Mood	33
2.1.4 Modality	37
2.1.5 Negation	40
2.2 Conclusion	45
Chapter 3. Discontinuous Syntactic Patterns	47
3.1 The source of the C-Infl link	47
3.2 Probe-goal relations in a phase	47
3.2.1 Subextraction (I): subject islands and φ -features	47
3.2.2 The Generalized Feature Inheritance Theory	50
3.3 Revising the Generalized Feature Inheritance Theory	53
3.3.1 The status of the phase head C and of Infl heads	54
3.3.1.1 Infl	54
3.3.1.2 C	55
3.3.2 Revising the Phase-Impenetrability Condition	55
3.3.2.1 Cyclicity	55
3.3.2.1.1 The classical empirical argument for strict cyclicity	57
3.3.2.1.2 Successive cyclicity	62
3.4 Discontinuities	68
3.5 Subextraction (II): relativized opacity for probe-goal relations	74
Chapter 4. Analytic Syntactic Patterns	77
4.1 On cartographies	77
4.2 The nature of ordering restrictions	80
4.2.1 Some allegedly non-primitive order restrictions	81
4.2.1.1 Epistemic modality and tense	81

4.2.1.2 Tense and aspect	83
4.2.2 Some allegedly primitive order restrictions	85
4.3 Toward a principled account for some order restrictions	86
4.3.1 Voice and tense	87
4.3.2 Quantificational aspectual adverbs and completive/prospective aspectual adverbs	88
4.3.3 Evidential mood and epistemic modality	89
4.4 The Full Interpretation Principle	90
4.5 Cartographic effects	92

Chapter 5. Sincretic Syntactic Patterns93

5.1 On structural minimization	93
5.2 Why and where V moves	96
5.3 A side-effect of V-to-T movement	108
5.3.1 SPEC-T becomes an A'-position when V has moved to T	109
5.3.1.1 Preverbal subjects in NSLs	109
5.3.1.2 Postverbal subjects in NSLs	110
5.3.1.3 Clitic Left dislocated elements in NSLs	111
5.3.1.4 Both preverbal subjects and CLLD elements match π -features in SPEC-T in NSLs	113
5.3.2 The source of π -features associated with SPEC-T when V instantiates φ -features on T	114
5.3.2.1 Subextraction (III): phase sliding	114
5.3.2.2 Maximize Matching Effects Principle	126
5.3.3 When SPEC-T is created in NSLs	127
5.4 Other conundrums to be solved on the basis of structural minimization	128
5.4.1 Double agreement dialects	129
5.4.2 From the Vacuous Movement Hypothesis to syntactic syncretism	138
5.5 Contraction	145
5.6 Conclusion	148

PART III. CONCLUSION

Chapter 6. On the emergence of order in syntax152

References156

The Emergence of Order in Syntax

PART I. INTRODUCTION

Chapter 1. Elements of syntax

This chapter discusses the most basic elements of the syntactic component of the Faculty of Language: features and a combinatorial operation called Merge. Features are defined as instructions, provided by the lexicon, for a level of interpretation of the Faculty (1.1). Merge is defined as an operation that takes as input two categories A and B, viewed as sets, and generates a new set C that is the union of A and B. The emergence of hierarchical properties derives from Merge (or union-formation) and from the existence of a derivational record, a set K (a nest) where the outputs of Merge are sets linearly ordered by the strict inclusion relation (1.2). Internal and external applications of Merge are argued to be triggered by the necessity of matching [+type] categories and [+token] categories, with no interpretable/uninterpretable dichotomy (1.3).

1.1 Instructions

An elementary metatheoretical notion relative to grammar –and to any other analytical discipline– is that of “level of representation” or “level of analysis”, understood as a set of concepts and devices used to construct a model for a phenomenon. Since any object of inquiry is too complex to be analyzed as a whole, one is forced to categorize the observed properties in several levels and subsequently study how they interact. It is well known that grammatical phenomena are classified at least as ‘semantic’, ‘syntactic’, ‘morphological’ or ‘phonetic’. It is also a truism that, although such splits are methodologically necessary, there are no ‘facts’ that belong to a unique level of representation, and researchers are forced to investigate a particular phenomenon by relating properties that are relative to different levels of representation.

A feature is a property of an object. A linguistic feature is particularly a property of a linguistic object, usually called a ‘linguistic expression’. If linguistic properties or features are classified into levels of representation or analysis, then a feature is ‘relative to’ a level of representation or analysis. This means that a ‘feature’ cannot exist unless there is a level of representation or analysis where such a feature can be represented or analyzed. Thus, we may say that there are features that are ‘relative to’ the levels of representation of phonetics, such as the [±plosive] feature or the [±voiced] feature, and others that are ‘relative to’ the level of semantics, such as the [±singular] feature or the [±negative] feature.

It may be interesting to shift from the metatheoretical notion of level of representation (a set of concepts and devices used to construct a grammatical model) to what could be called an empirical notion of level of interpretation. According to this empirical notion, the levels of interpretation are components of the Faculty of Language, viewed as a set of emergent properties of the brain, that interpret the features of a linguistic object as instructions. This implies that the sort of theory we shall be seeking is not only a model for the shape of linguistic products, but a model for the linguistic knowledge that generates those products. Such a move, inherent to the biolinguistic-generative grammar amalgam, has the virtue of reframing the usual purely ‘intuitive considerations’ of elegance and parsimony relative to the model in construction in terms of ‘empirical questions’ about the Faculty that try to discover to what extent it is an optimal solution to multiple usability conditions, such as

information coding and productive thinking. Therefore, instead of saying ‘this grammatical model is better than that one because the principles it needs are fewer and more general’, we will say ‘the Faculty of Language must be like this, and not like that, because this is all we need to satisfy usability conditions’. This approach to Language is what Noam Chomsky calls the Strongest Minimalist Hypothesis (SMT): ‘Language is an optimal solution to legibility conditions’.

According to the empirical notion of “level of interpretation”, linguistic features are useful instructions for a particular level of interpretation of the Faculty of Language.

A rather common working hypothesis in grammatical studies is that linguistic objects are comprised of categories, which are clusters of features. It is also customary to distinguish between substantive categories and functional categories, the former constituting an open class and the latter a closed class; accordingly, nouns and verbs would be considered substantive categories, and, for instance, person-number marks would be considered functional categories. However, it must be noted that the outlined distinction does not seem to be accurate, since the so-called substantive categories are not atomic, but made of functional categories and roots; if so, the right distinction may be in terms of functional categories and roots.

Let’s begin by considering functional categories, which will be central to our concerns. A first basic question is: for which levels of interpretation of the Faculty of Language do they contain features or instructions? Since several subdivisions may be necessary for the so-called semantic and phonetic levels of interpretation, I shall often refer to them as the Conceptual-Intentional (C-I) system and the Articulatory-Perceptual (A-P) system, adopting two familiar terms.

Functional categories contain instructions for the C-I system. For instance, conjunctions or clause typing categories are standardly considered to specify the force of a proposition ([±declarative], [±exclamative], etc.) and the [±embedded] distinction. A common semantic instruction of Inflection-like functional categories such as tense, mood, modality, negation and the person-number conglomerate is that they trigger referential displacement in different dimensions (temporal, modal, or individual).

More concretely, modal categories provide a relationship that determines the ‘force’ of a conclusion with respect to an evaluation parameter provided by a conversational background (see Kratzer 1981, 1986), indicating whether a proposition holds in some possible world ([+existential]) or in all possible worlds ([+universal]). Mood categories define an embedded proposition as [±actual] (see Quer 1998, 2003 for discussion and a new dynamic characterization of the mapping of mood morphemes in terms of ‘model shift’) and tense categories locate the time of the event in an actual time ([+present]) or in a non-actual time ([+past] or [+future]). As recently argued by Sigurðsson (2003, 2004) following a rather traditional view, the person feature relates the individuals of the event (the arguments) to the participants of the speech act, allowing reference to an individual present in the speech act ([+logophoric agent, -logophoric patient], [-logophoric agent, +logophoric patient]) or to an

individual absent in the speech act [-logophoric agent, -logophoric patient]. Hence the values for the person feature of an individual i can be defined as in (1):

- (1) Person of i : [1st] iff i = [+logophoric agent, -logophoric patient]
 Person of i : [2nd] iff i = [-logophoric agent, +logophoric patient]
 Person of i : [3rd] iff i = [-logophoric agent, -logophoric patient]

It is noteworthy that the values for the number feature (singular, plural, dual, etc.) do not strictly reflect properties of the logophoric participants. The [+plural] specification of the 1st and 2nd person pronouns in (2) does not necessarily convey that the logophoric agent or the logophoric patient are plural.

- (2) a. We are smoking
 b. You are smoking

This is also noted by Sigurðsson (2004: 19), among others (see references cited therein):

«*We*, for instance, is obviously not a plural of I in the sense that it denotes ‘more than one speaker’ (except perhaps under extremely rare and special circumstances). However, it is the plural of I in the sense that it denotes more than one potential linguistically active selves, where a linguistically active self is a speaking or a thinking (experiencing, feeling, ...) actor: ‘I and others that could be in my footsteps as speakers/thinkers/experiencers’».

Hence, the instruction provided by the values of the number feature relative to an individual i belonging to the set I of individuals that are involved in an event ψ is to define the cardinality of I (symbolically, $|I|$).

- (3) Number of i ($i \in I$): [singular] if $|I| = 1$
 Number of i ($i \in I$): [plural] if $|I| \geq 2$
 Number of i ($i \in I$): [dual] if $|I| = 2$

Finally, if negation in natural languages is not essentially different from the Boolean connective ‘ \neg ’, the semantic instruction it provides must be to shift from the truth value of χ to the truth value of $\neg\chi$ in such a way that $\neg\chi$ is false if and only if χ is true, and $\neg\chi$ is true if and only if χ is false.

In sum, the ability to trigger referential displacement seems to be a general feature or instruction provided by functional inflectional categories to the semantic level of interpretation: the semantic feature provided by modal categories, mood categories, tense categories and the person-number conglomerate is to shift from the actual to the non-actual, and the semantic feature provided by negation is to shift from the truth χ of to the falsity of $\neg\chi$, and from the falsity of χ to the truth of $\neg\chi$.

The morphological level of interpretation also requires features from functional categories. For instance, a tense category may be morphologically realized as a free particle, as an attached head or as a maximal projection (an adverbial phrase). Furthermore, it may be the case that two different categories are assigned to two different slots (analytic pattern), that

two different categories are assigned to a single slot (syncretic pattern) or finally, that a single category is assigned to two different slots (discontinuous pattern). Thus, a functional category provides ‘templatic’ features such as [\pm free], [\pm analytic], [\pm syncretic], [\pm discontinuous].

Functional categories also provide features to the A-P system. For instance, in English, an embedded finite assertive conjunction is spelled out as [$\delta\theta t$] and an indirect partial question conjunction as [$\underline{r}f$] or [$w\varepsilon\delta\theta$].

If there exists a grammatical component that orders categories, and this component is called syntax, one of the most elementary questions about syntax is where the necessary instructions for ordering pieces come from. In ordering pieces, syntax associates a category displaying a token of a particular type of feature with a position semantically specialized for that type of feature. This observation can be implemented in the theory of the syntactic component by postulating two sorts of features: types and tokens. If so, the syntactic level of interpretation interprets certain categories as types, which indicate which category can appear in a particular position, and others as tokens, which instantiate a particular type. According to this reasoning, syntactic operations are driven by the necessity of matching type categories with token categories.

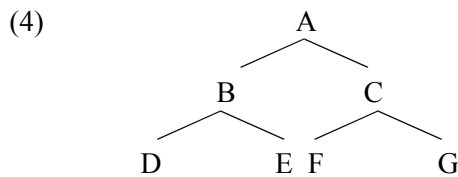
It may well be that categories are not grammatical primitives but grammatical by-products, or in other words, that there is no such a component as a unified lexicon with categories (clusters of features or instructions) listed, but a multiplicity of lexical components, each providing a particular subset of features. Since this issue is immaterial for our concerns, I shall adopt the handiest option, i.e., that phonetic features, morphological templatic features, the type/token syntactic feature and semantic features are coded in lexical entries dubbed categories. The lexicon, the locus of lexical entries, contains the instructions that activate grammatical computations.

The various grammatical components of the Faculty of Language (phonetics, morphology, syntax and semantics) are at the same time levels of interpretation and computational devices following the interpreted instructions. Syntax interprets objects selected from the lexicon and submits them to its own computations, constructing greater objects that are interpreted at the Logical Form (LF) level and at the morphological level. LF is not a final product, but the interface between syntax and the C-I system; similarly, morphology is also an interface between syntax and further devices of the A-P system. Indeed, syntax is also an interface, which relates the lexicon to morphology and LF. Both LF and morphology submit syntactic objects to its own computations creating new outputs, which will be repeatedly interpreted and computed until two final representations are attained, one (SEM) usable by thought systems and the other (PHON) by externalization systems. Whereas the lexicon is the first motor of grammatical computations, SEM and PHON are their products.

It is important to express these preliminary considerations here, since our efforts will be concretely guided to define the design properties of the syntactic component of the Faculty of Language as an optimal interface between the lexicon, the morphological component and the semantic component.

1.2 Merge: nests

If there is a continuous set of branches going down a phrase marker like (4) from a higher category to a lower category, it is said that the higher category dominates the lower one. Thus, in (4) *A* dominates the rest of categories, *B* dominates *D* and *E*, *C* dominates *F* and *G*, and *D*, *E*, *F* and *G* dominate no category. The relation of term-of is the inverse relation of dominance; hence, in (4) *D* and *E* are terms of *B*, *F* and *G* are terms of *C*, *B* and *C* are terms of *A*, all the categories are terms of *A* and *A* is not a term of any category.



On the basis of the dominance relation, two further structural relations have been defined: c-command and asymmetric c-command.

c-command

A category *X* c-commands another category *Y* iff *X* does not dominate *Y* and every category dominating *X* dominates *Y*.

asymmetric c-command

A category *X* asymmetrically c-commands another category *Y* iff *X* c-commands *Y* and *Y* does not c-command *X*.

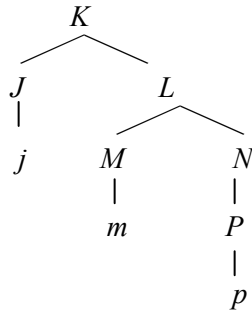
A very important proposal with far-reaching consequences developed by Richard Kayne (Kayne 1994) suggests that the hierarchical nature of syntactic objects is closely related to the defining properties of the linear order relation. More precisely, Kayne proposes that the terminals of a syntactic object can be ordered if and only if the non-terminals that dominate them follow an asymmetric c-command relation. This is expressed by the Linear Correspondence Axiom:

Linear Correspondence Axiom (LCA)

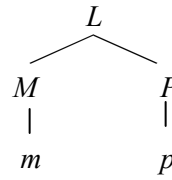
Given a set *T* of terminals of a phrase marker *P* and a total asymmetric c-command relation *A* among the non-terminals of *P*, the dominance relation from non-terminals to terminals yields a linear ordering on *T* (adapted from Kayne 1994: 3-6).

In the phrase marker (5), the non-terminal *M* asymmetrically c-commands the non-terminals *P*, and the non-terminal *J* asymmetrically c-commands the non-terminals *M*, *N* and *P*. By virtue of the LCA, the ordering of terminals must be $\langle j, m \rangle$, $\langle m, p \rangle$ and $\langle j, p \rangle$. However, in the phrase marker (6), the terminals *m* and *p* cannot be ordered, since neither *M* asymmetrically c-commands *P* nor *P* asymmetrically c-commands *M*. The tree-diagram (5) conforms to the *X'*-theory and provides a strict linear ordering relation on the terminals; diagram (6) does not.

(5)



(6)



The so-called X' -theory was “an early stage in the effort to resolve the tension between explanatory and descriptive adequacy” (Chomsky 1995: 241) that attempted to reduce multiple phrase structure rules belonging to particular languages and redundant in its form (Chomsky 1965, among many others) to the following unified syntactic template attributed to Universal Grammar (Chomsky 1981a, Chomsky & Lasnik 1993, among many others):

- (7)
- a. $X'' \rightarrow ZX'$
 - b. $X' \rightarrow X'Y$
 - c. $X' \rightarrow X^0W$

The lexical item inserted in X^0 determines the type or label of X' and X'' and restricts the choices of W and Z , called respectively complement and specifier.

Kayne’s general conclusion is that the X' -theory is not a primitive of Universal Grammar (UG), but a construct derived from the necessity of assigning a strict linear order to the terminals of a phrase marker.

It is crucial for Kayne’s proposal to assume a substantive (not merely notational) distinction between maximal, intermediate and minimal categories; however, as Chomsky has argued (1994a, 1995), these elements are stipulated conventions of the X' -theory: there is no independent motivation to think that the computations taking place at the C-I system or at the A-P system must be sensitive to whether a category is maximal, intermediate or minimal, a distinction that furthermore violates the inclusiveness condition (“any structure formed by the computation is constituted of elements already present in the lexical items” –Chomsky 1995: 228). Consequently, Chomsky proposes substituting the X' -theory with a bare phrase theory, where the unique legitimate objects are terminal categories or lexical items (clusters of features) and sets of lexical items formed by syntactic operations.

The LCA cannot be formulated in a bare phrase structure: if non-terminal categories are eliminated, there is no set A from which a dominance relation from non-terminals to terminals may yield a linear order among terminals. In what follows I shall develop a reformulation of the LCA such that it is compatible with bare phrase structure, postulating no entities but terminals and sets.

The simplest possible way of defining a combinatorial system is postulate the operation ‘Merge’ as a basic element. ‘Merge’ takes n objects and constructs a new object. The objects taken as input may be syntactic atoms (namely, lexical items) or objects previously constructed by Merge. Thus, Merge is a recursive operation: it can take objects that result from previous applications of itself as input. The output of Merge is an object that displays the property of ‘discrete infinity’: it is ‘discrete’ because it is ultimately made of atomic elements and ‘infinite’ because the number of applications of merge operations is potentially boundless.

More precisely, we shall define Merge as a set-formation operation that takes as input two sets A, B and yields as output a set C that contains all the elements of A and all the elements of B and that does not contain any other elements. Thus, Merge of A and B forms the union of A and B ($\text{Merge}(A, B) \rightarrow A \cup B$). When A or B (or both) is the output of an immediately preceding application of Merge, we may say that Merge is successive. Let’s examine the consequences of this definition.

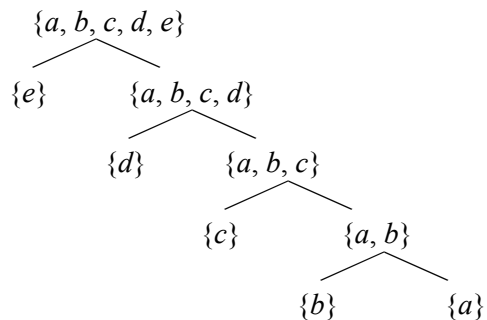
Consider the following derivation by successive Merge. Merge_1 is the first application of Merge, which means that \emptyset is the output of the immediately preceding application of Merge. The first application of Merge will thus yield a singleton as output.

$$\begin{aligned} \text{Merge}_1(\emptyset, \{a\}) &\rightarrow \{a\} \\ \text{Merge}_2(\{a\}, \{b\}) &\rightarrow \{a, b\} \\ \text{Merge}_3(\{a, b\}, \{c\}) &\rightarrow \{a, b, c\} \\ \text{Merge}_4(\{a, b, c\}, \{d\}) &\rightarrow \{a, b, c, d\} \\ \text{Merge}_5(\{a, b, c, d\}, \{e\}) &\rightarrow \{a, b, c, d, e\} \end{aligned}$$

Let the representation generated by the five applications of successive Merge be the derivational record of the five outputs of successive Merge, i.e., the set K where the five outputs are sets linearly ordered by the strict inclusion relation. We shall call a set of sets linearly ordered by inclusion a nest or a totally nested structure.

$$K = \{\{a\}, \{a, b\}, \{a, b, c\}, \{a, b, c, d\}, \{a, b, c, d, e\}\}$$

The sets that belong to K can be identified with the syntactic nodes of a tree-diagram and the elements of these sets as the the terminals dominated by the syntactic nodes.



Observe that K extends Kuratowski's (1921) definition of ordered pair ($\{\{a\}, \{a, b\}\} = \langle a, b \rangle$) to a 5-tuple:

$$\begin{aligned} \{\{a\}, \{a, b\}\} &= \langle a, b \rangle \\ \{\{a, b\}, \{a, b, c\}\} &= \langle \langle a, c \rangle, \langle b, c \rangle \rangle \\ \{\{a, b, c\}, \{a, b, c, d\}\} &= \langle \langle a, d \rangle, \langle b, d \rangle, \langle c, d \rangle \rangle \\ \{\{a, b, c, d\}, \{a, b, c, d, e\}\} &= \langle \langle a, e \rangle, \langle b, e \rangle, \langle c, e \rangle, \langle d, e \rangle \rangle \end{aligned}$$

Hence, by transitivity, $K = \langle a, b, c, d, e \rangle$, where a is the most deeply nested element and e the least deeply nested element. The representation K is an image of the syntactic derivation and a specular image of the literal precedence relation among terminals in an utterance.

Let us consider some possible complications.

If some terminals are equal to each other, Kuratowski's equivalence cannot be extended to n -tuples such that $n > 2$: Although $\langle a, b, b \rangle \neq \langle a, a, b \rangle$, we observe that $\{\{a\}, \{a, b\}, \{a, b, b\}\} = \{\{a\}, \{a, a\}, \{a, a, b\}\}$, by the Extensionality Principle. Note that it is crucial for Kuratowski's definition that $\{a\} \neq a$.

$$\begin{aligned} K_1 &= \{\{a\}, \{a, b\}, \{a, b, b\}\} = \{\{a\}, \{a, b\}, \{a, b\}\} = \{\{a\}, \{a, b\}\} \\ K_2 &= \{\{a\}, \{a, a\}, \{a, a, b\}\} = \{\{a\}, \{a\}, \{a, b\}\} = \{\{a\}, \{a, b\}\} \end{aligned}$$

This observation would raise a problem for the idea that hierarchical properties of syntactic objects emerge from the creation of nests if and only if both (i) and (ii) were correct:

- (i) Merge could be vacuous: the output of Merge_n could be equal to the output of Merge_{n+1} . For instance, in the hypothetical derivation that would generate K_2 , the output of Merge_1 would be $\{a\}$; Merge_2 would take as input $\{a\}$ and $\{a\}$ and yield $\{a\}$, and
- (ii) Some elements in a syntactic object, but not all, included themselves (a would include itself in K_2 , but not b).

There is no reason to think that Merge may apply vacuously or that some elements, but not all, contain themselves. Therefore, the problem under discussion does not arise in syntactic representations/derivations.

Consider the step Merge_3 in the Derivation 1, where the input of Merge is two objects previously constructed by Merge.

Derivation 1

$$\begin{aligned} \text{Merge}_1(\emptyset, \{\textit{novels}\}) &\rightarrow \{\textit{novels}\} \\ \text{Merge}_2(\{\textit{novels}\}, \{\textit{reads}\}) &\rightarrow \{\textit{novels}, \textit{reads}\} \\ \text{Merge}_3(\{\textit{novels}, \textit{reads}\}, \{\textit{the}, \textit{gardener}\}) &\rightarrow \{\textit{novels}, \textit{reads}, \textit{the}, \textit{gardener}\} \end{aligned}$$

$$K_1 = \{\{\textit{novels}\}, \{\textit{novels}, \textit{reads}\}, \{\textit{novels}, \textit{reads}, \textit{the}, \textit{gardener}\}\}$$

The derivational record K_1 is a nest (its subsets are linearly ordered by the inclusion relation), but a linear order among terminals is not obtained unless the derivational record K_2 is accessed: *novels* is more embedded than *reads*, both *novels* and *reads* are more embedded than *the* and *gardener*, but no ordering exists in K_1 between *the* and *gardener*.

Derivation 2

$\text{Merge}_1(\emptyset, \{\textit{gardener}\}) \rightarrow \{\textit{gardener}\}$

$\text{Merge}_2(\{\textit{gardener}\}, \{\textit{the}\}) \rightarrow \{\textit{gardener, the}\}$

$K_2 = \{\{\textit{gardener}\}, \{\textit{gardener, the}\}\}$

In the X' -theoretic jargon, this means that a complement is more deeply embedded than a head, and that they are both more deeply embedded than a specifier, whose terminals are independently ordered because they are merged in a different (sub-)derivation. Thus, specifiers and non-branching complements are not a problem for linearization anymore. Note that, if Merge were defined as set-formation (as Chomsky 2005b suggests), instead of union-formation, specifiers would still pose a problem for linearization: $\{\{\textit{the}, \{\textit{gardener}\}\}, \{\textit{read}, \{\textit{novels}\}\}\}$.

We are now in a good position for reformulating the LCA in bare phrase structure terms, replacing the total asymmetric c-command relation among non-terminals, the set A of Kayne's formulation, with a linear strict inclusion relation among sets, a nest.

LCA

A strict linear order on a set of terminals T is yielded by a totally nested syntactic structure on T .

However, the above sentence is not an axiom. It is not a proposition one needs to take for granted to derive other propositions, namely the X' -theory. Indeed, all we need in the theory of syntax to account for hierarchy is a successive operation of union-formation and the existence of a derivational record, a nest. In this case, the history of the theory of phrase structure is no more than an instance of how the construction of a model for a particular phenomenon is guided by the desiderata of minimal design.

1.3 Merge: internal and external

In the development of the minimalist inquiries, it has been proposed that the following two properties of language are related:

- (I) Functional features are interpretable on some lexical items and uninterpretable on others
- (II) The surface phonetic position of a category is dissociated from the deep semantic position

The person-number conglomerate (referred to as ‘ φ -features’, as it is customary in the minimalist tradition) offers the most compelling evidence for the existence of property (I). As illustrated in the following Catalan paradigm, φ -features may surface on nominals (nouns, pronouns and determiners), but also on verbs and adjectives, although the semantic instruction that each occurrence provides is relative to nominals exclusively.

- (8)
- | | | |
|--------------------------|-----------------------|-----------------|
| a. Vindr  | jo | |
| come.FUT.1SING | 1SING | |
| ‘I will come’ | | |
| b. Vindrem | nosaltres | |
| come.FUT.1PLUR | 1PLUR | |
| ‘We will come’ | | |
| a. Vindr s | tu | |
| come.FUT.2SING | 2SING | |
| ‘You will come’ | | |
| b. Vindreu | vosaltres | |
| come.FUT.2PLUR | 2PLUR | |
| ‘You will come’ | | |
| a. Vindr  | l’estudiant | |
| come.FUT.3SING | The.SING student.SING | |
| ‘The student will come’ | | |
| b. vindran | els | estudiants |
| come.FUT.3PLUR | The.3PLUR | estudiants.PLUR |
| ‘The students will come’ | | |
- (9)
- | | | |
|------------------|----------|-----------|
| a. Un | cotxe | r pid |
| a.SING | car.SING | fast.SING |
| ‘A fast car’ | | |
| b. Uns | cotxes | r pids |
| a.PLUR | car.PLUR | fast.PLUR |
| ‘Some fast cars’ | | |

An occurrence of φ -features on a nominal is said to be [+interpretable] because it provides semantic instructions relative to such a nominal, whereas an occurrence of φ -features on a verb or an adjective is said to be [-interpretable] because it does not provide instructions relative to events, but rather reflects semantic properties of the nominal with which it ‘agrees’.

Property (II), usually conceived of as a displacement operation of a category from one position to another, is broadly attested in natural languages and displays a relatively rich phenomenology: there is movement to case positions and to peripheral positions, movement of heads and movement of projections, apparently long-movement and local movement, movement with pied-piping and movement with stranding.

Both properties (I) and (II) have been viewed as striking, or in fact as ‘imperfections’: why should lexical items contain uninterpretable features?; why should a category be pronounced in a position where it is not semantically interpreted?, or in other words, why is it that the occurrences phonetically interpreted at the A-P system and semantically interpreted at the C-I system are not the same, and hence PHON (the linguistic object to be used by externalization systems) diverges from SEM (the linguistic object to be used by thought systems)?

In Chomsky (2001a, 2001b), and related work by many others, the working hypothesis that these two apparent imperfections of narrow syntax may be related is investigated. According to this view, uninterpretable features are the trigger of displacement. Consider the following prototypical examples of raising constructions.

- (10) a. [A proof] is likely [to have been discovered [~~a proof~~]]
b. [A proof] seems [to have been discovered [~~a proof~~]]

It is claimed that matrix T has a complete set of uninterpretable φ -features; embedded T has a defective set of φ -features which cannot match the complete set of φ -features of a referential DP like ‘[a proof]’. T matrix forces the DP to move to its SPEC, causing ‘[a proof]’ to be pronounced in the clause of ‘be-likely’ or ‘seem’, and thematically interpreted in the VP of ‘discover’. According to this analysis, the traditional case filter effects are not attributed to the necessity for a DP to receive case, but rather to the necessity for the uninterpretable φ -features of matrix T to get deleted. Thus the source of movement, property (II), is precisely property (I): uninterpretable features trigger movement of a category with interpretable features in order to satisfy their need to get deleted. Call this need, the device that triggers displacement, suicidal greed. This deletion operation must be carried out before the syntactic object arrives at the C-I system; otherwise, the derivation will yield a gibberish object, containing categories without useful instructions for the C-I system and thus violating the Full Interpretation Condition:

Full Interpretation Condition

An object generated by a grammatical component must be constituted only of useful elements for a particular level of interpretation when it attains such a level of interpretation

Recently, Chomsky has observed that the existence of movement in natural languages is not a striking property requiring special motivation:

«For over 40 years, there have been efforts to motivate displacement. That seems to have been a mistake. Recourse to any device to account for the displacement phenomena also is mistaken, unless it is independently motivated (Chomsky 2001b: 8, note 29)».

The argument is simple: Merge, the recursive operation responsible for putting pieces together by creating greater syntactic objects, comes for free in a combinatorial system like the syntactic component. Merge takes as input *A* and *B*. If these two objects are separate objects, Merge is ‘external’, whereas, if *A* is a term of *B*, Merge is ‘internal’; in the second case, *A* is taken from *B* and re-merged to *B*, yielding the property of displacement. “Accordingly, displacement is not an imperfection of language; its absence would be an imperfection (Chomsky 2001b: 8)”.

Property (I) must be studied more thoroughly. It is currently understood that displacement, property (II), cannot be seen as an imperfection; however, it seems that it does not apply vacuously, since all instances of displacement yield semantic effects of multiple types. One might consider whether suicidal greed is the device responsible for constraining or triggering the applications of internal Merge: an uninterpretable feature located in a position of the hierarchy needs to be deleted, and hence triggers movement of a category with a matching interpretable feature. Accordingly, a category would be associated with multiple interpretations due to uninterpretable features that would be present in different positions of a hierarchy just to be deleted.

Such a theoretical device presents some problems. I shall point out three of them, which suffice to show the inadequacy of suicidal greed as the trigger of displacement. The first one may be viewed as a problem of generality, the second one as a problem of determinacy, and the third one as a problem of consistency.

As for the problem of generality, there are many functional features triggering displacement. Thus, if uninterpretability is the real trigger of displacement, it is expected that the uninterpretable status of triggering features can be determined not only for φ -features but for a plentitude of functional features. But once we move to functional features other than φ -features of T, there are no convincing reasons for uninterpretability. For example, it has always been intuitively unclear that features triggering A'-movement are uninterpretable, since the claim that peripheral features (force features or topic/focus features) are uninterpretable on its locus, C, is entirely stipulative. Indeed, π -features (peripheral features) on C are more naturally characterized as positional indications for where a π -constituent must be merged in order to be properly interpreted. This is usually expressed in terms of Rizzi's (1991, 1997) or Haegeman's (1995) criteria, but these representational conditions cannot be the derivational trigger of syntactic operations. More plausibly, as argued above, force features are not the only type of triggering features that are clearly interpretable: the class of inflectional-like features (mood, tense, φ -features and negation) provide the semantic instruction of triggering referential displacement and also trigger movement operations. Hence, suicidal greed does not seem to be the general source of displacement. This problem of generality casts some doubt on the adequacy of the suicidal greed device for the particular case of φ -features: a model that accounted for all instances of movement with a unique device is theoretically superior to a model with a plurality of devices.

Let's now shift to the problem of determinacy. We must more closely examine whether the occurrences of φ -features triggering A-movement are uninterpretable. Assume that φ -features

on a verb are [-interpretable], for they do not provide semantic instructions relative to V but rather relative to a nominal. Note that the (un-)interpretable status of φ -features on T does not derive from the premise that they are uninterpretable on V. The interpretability of a feature is relative to a category, and thus one cannot determine the uninterpretable status of φ -features on T from the uninterpretable status of φ -features on V. In the absence of specific arguments for the uninterpretability of φ -features on T, the neutral hypothesis is that they are interpretable, as π -features of C. In fact, we may extend Rizzi's view on π -features to φ -features of T: φ -features of T are responsible for determining the binder of a subject-oriented anaphor, as π -features of C are responsible for determining the scope and the quantificational force of a π -phrase. Again, this suggests that φ -features of T, the trigger of A-movement, are as interpretable as π -features of C, the trigger of A'-movement.

Finally, we address the problem of consistency. Although there is no reason to think that the uninterpretability of φ -features on V entails the uninterpretability of φ -features on T (the determinacy problem), which would indeed be an exceptional case among triggering features (the problem of generality), continue assuming that φ -features on both V and T are uninterpretable. As will be argued in detail in chapter 5, V moves to T precisely in those languages where V shows up with an overt occurrence of uninterpretable φ -features; if V does not show up with an overt occurrence of uninterpretable φ -features, it remains inside the vP. This presents an inconsistency for the hypothetical suicidal greed device, since uninterpretable features on V delete uninterpretable features on T. It is also important to note that Null Subject Languages display rich inflectional morphology on V, which suggests that an interpretable instance of φ -feature (a DP) and an uninterpretable instance of φ -features (an inflected V) compete to delete an uninterpretable feature on T (see chapter 5); on these grounds, it could be stipulated that φ -features on T provide useful semantic instructions (i.e., they are interpretable) when V is inflected for (uninterpretable) φ -features, whereas they are uninterpretable when V is not inflected for φ -features, in which case a DP with interpretable φ -features must move to delete the uninterpretable φ -features of T. It is not worth considering such an inelegant possibility, whereby the interpretability of features is decided in a purely stipulative way.

The problems of generality, determinacy and consistency strongly push the device of suicidal greed into an unacceptable situation and hence manifest its inadequacy. In this case, what is the trigger of internal Merge of a category? Since the property of displacement dissolves into the basic operation of Merge, there can be only one answer: the same that the triggers external Merge of a category. Any device specially constructed for displacement is suspicious.

Indeed, an accurate way of expressing the asymmetry between hosting categories and hosted categories has already been proposed in this study. A functional category X can enter the derivation specified as a type or as a token. If X is a type, it is interpreted as a hosting category in syntax. There are two possible strategies to instantiate a [+type] feature with a suitable [+token] feature: by external Merge or by internal Merge. Thus, both internal and external Merge of categories with [+token] features are triggered by the necessity of matching [+type] features present in a syntactic hierarchy and [+token] features, which yields

a positional system where particular semantic instructions are associated with particular positions.

1.4 Onset

Therefore, the most basic elements of the syntactic component are features provided by categories or lexical units and a combinatorial operation called Merge that, in matching types and tokens, automatically yields nests or hierarchically structured representations.

In the next part, I shall examine how syntactic derivations by successive (internal and external) Merge are driven by the morpholexical characterization of lexical items and constrained by semantic legibility conditions that apply as well-formedness conditions and by general efficiency principles that reduce the necessary span of structure to match types and tokens. Or in other words, I shall attempt to understand why syntactic objects have the shape they seem to have by inspecting how the syntactic component interfaces the morphological component and the semantic component in an efficient way.

PART II. PATTERNS

Chapter 2. The relationship between complementizers and inflectional categories

This chapter argues (i) that complementizers can ‘replicate’ Infl-like features as φ -features, tense, mood, modality and negation, (ii) that there is a correlation between the characteristics of Infl-like and the choice of C, and (iii) that Infl-like features can trigger movement both to Infl and C. These three properties materialize the link between C and Infl.

2.1 The C-Infl link

There is a strong correlation between complementizers (or [+clause typing] categories) and inflectional categories (or [-clause typing] categories), a traditional observation adhered to in several classical works in generative grammar (Rosenbaum 1965, Bresnan 1972, den Besten 1977/1981, Koster 1975, 2003, Stowell 1981). This point is also noted in cartographic studies as Rizzi (1997) or Cinque (1999a: note 53 and relative text in p. 18; note 20 and relative text in p. 84). For example, Rizzi (1997: 283) observes that:

«It is a traditional observation that the choice of the complementizer reflects certain properties of the verbal system, an observation formalized by, e.g., “agreement” rules between C and I, responsible for the co-occurrence of “that” and a tensed verb, or “for” and an infinitive in English (Chomsky and Lasnik 1977), etc. A straightforward manner to account for these dependencies would be to assume that C contains a tense specification which matches the one expressed on the lower inflectional system (an idea that goes back at least to den Besten (1977))».

I shall discuss a sample of phenomena that substantiate the idea that several features generally associated with [-clause typing] positions are linked to [+clause typing] positions, a topic profusely studied in grammar. I shall use the label ‘Infl’ instead of the currently more common ‘T’, since it would be confusing to label [-clause typing] occurrences of features such as negation mood or modality as ‘T’. Chapter 4 shall consider the source of the generalized C-Infl link.

2.1.1 φ -features

In several languages, not only verbs but complementizers as well are inflected for φ -features. Certain Germanic languages offer well-known examples of this morphosyntactic phenomenon (see, among others, Bayer 1984, Haegeman 1992, Hoekstra & Marác 1989 and Zwart 1993). Consider the following paradigm from West Flemish (Haegeman 1992), where complementizers show up with two different subject agreement marks successively attached:

- | | | | |
|-----|----|--------------------|--------------|
| (1) | a. | da-n-k | ik komen |
| | | that-1SG-I | I come-1SG |
| | b. | da- \emptyset -j | gie komt |
| | | that-2SG-you | you come-2SG |
| | c. | da-t-j | ij komt |

	that-3SG-he	he come-3SG	
d.	da-∅-se	sij komt	[∅ < t]
	that-3SG-she	she come-3SG	
e.	da-∅-me	wunder komen	[∅ < n]
	that-1PL-we	we come-1PL	
f.	da-∅-j	gunder komt	[∅ < t]
	that-2PL-you	you come-2PL	
g.	da-n-ze	zunder komen	
	that-3PL-they	they come-3PL	

It could be argued that ϕ -features are base-generated in the Infl-area, but languages may vary with regard to ‘replicating’ these instructions in the C-area. In that case, there would be a rather arbitrary morphological device for replicating functional features where they are not interpreted. Such a process may be operative in natural languages, as in multiple auxiliaries that agree with a nominal in Swahili (2.a) or Arabic (2.b), multiple determiners or adjectival categories that agree with N in Catalan (3), or multiple elements of the clause that may show an optional plural mark when the subject is plural in Korean (4), a phenomenon known as spurious agreement.

- (2) a. Juma *a-li-kuwa* *a-ngali a-ki-fanya kazi*
 Juma SA-PST-be SA-still SA-PROG-do work
 ‘Juma was still working’

- b. *al-bint-aani* *kaan-ataa* *ta-ktub-aani* *darsa-humaa*
 the-girls(F)-3D be-PST-3FD 3f-write-D lesson-FD
 ‘The two girls were writing their lesson’

(Carstens 2001: 395)

- (3) a. *La* *meva* *casa* *verda*
 the.FEM.SING my.FEM.SING house. FEM.SING green.FEM.SING
 ‘My green house’

- b. *El* *meu* *cotxe* *verd*
 the.MASC.SING my.MASC.SING car.MASC.SING green.MASC.SING
 ‘My green car’

- (4) *ai-tul-i* *sensayngnim-kkey-tul* *yelsimhi-tul* *cilmun-ul*
 child.PL.NOM teacher.DAT.PL intently.PL questions.ACC
 ha-ko-tul *iss-ta*
 do-comp-PL be
 ‘The children are asking questions to the teacher intently’

(Kim 1994: 303)

In (4) plural features of N spread not only to the verb, but also to an indirect object and an adverb. Interestingly, as Kim (1994) illustrates, they may also spread to the complex [V-complementizer] of an embedded clause, but not to arguments or adjuncts of such an

embedded clause. In (5) the plural feature of the matrix subject may spread to the embedded [V-complementizer], but not to the locative adjunct of the embedded clause, whose subject is singular.

- (5) a. *sensayngnim-tul-un* [kim-I emma-lul hakkyo-ey(*-tul) mosie olila-ko]-tul
 teacher-PL-TOP Kim-NOM mother-ACC school-LOC-PL bring-COMP-PL

sayngkak hayessta
 thought

‘Teachers thought that Kim would bring his mother to the school’

But the sort of correlation between [+clause typing] positions and [-clause typing] positions seems to be deeper than this morphological process of spreading features to several constituents, since there is a meaningful relationship between the value of ϕ -features showing up on Infl and the choice of C. The most basic observation is that the finiteness of a clause is coded not only on Infl but also on C in such a way that the Infl head of an embedded clause in several languages has a complete set of ϕ -features (V is finite and a full DP can be licensed) only if the embedded C is finite; in Catalan, for instance, the complementizer *de* correlates with non-finite embedded verbs, whereas the complementizer *que* correlates with finite embedded verbs, and only in finite clauses can a referential subject DP be licensed.

- (6) a. Diu [que/*de vinguin]
 say.3SG C[+FIN]/[-FIN] come.SUBJ.3PLUR
- b. Diu [*que/de venir]
 say.3SG C[+FIN]/[-FIN] come[-FIN]
- (7) a. Diu [que vinguin els estudiants]
 say.3SG C[+FIN] come.SUBJ.3PLUR the students
- b. *Diu [de venir els estudiants]
 say.3SG C[-FIN] come[-FIN] the students

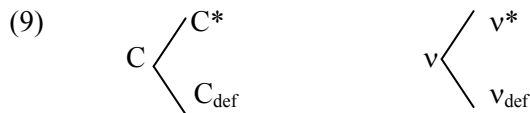
In fact, the [\pm fin] contrast in languages like Catalan is essentially the same as the richer paradigm of West Flemish: in Catalan, complementizers display [\pm fin] distinctions, whereas in West Flemish, complementizers also display person and number distinctions.

Chomsky (2000, 2001a, 2001b) has argued that T has a complete set of ϕ -features associated with nominative case only if it is selected by C, as V has a complete set of ϕ -features associated with accusative case when it is selected by v^* (this is essentially Burzio’s 1986 generalization). If T is selected by V (instead of C) and V by v (instead of v^*), T and V, respectively, lack ϕ -features. Thus, there is a one-to-one relationship between ϕ -completeness and the presence of a phase head, C and v^* .

- (8) T is [φ -complete] if selected by C
 T is [φ -defective] if selected by V (C = \emptyset)
 V is [φ -complete] if selected by v*
 V is [φ -defective] if selected by v

Note that there is a suspicious asymmetry between the phase head C, associated with the properties of a proposition, and the phase head v*, associated with argument structure. v* is one of the possible values for v, the selector of the root and the locus of the agentive/non-agentive distinction. Consequently, when the phase head v* is not present, the archicategory v must be present to define its complement as verbal. However, when there is no phase head C, there is no archicategory to define the clause type of proposition or to encode the [\pm embedded] distinction.

In order to overcome this vagueness for propositions, I shall go a step further in the analogy between the two phases, by expressing that, as v is always present containing argument structure information, C is always present containing clause typing information. Both v and C can have several values; more concretely, they can be v* and C*, in which case they select a full set of φ -features, or C_{def} and v_{def}. Hence, v* is a specific value for v that selects an external argument and a complete set of φ -features that offers an A-position for accusative DPs, as C* is a particular value for C that selects a complete set of φ -features that offers an A-position for nominative DPs.



- T is [φ -complete] if selected by C*
 T is [φ -defective] if selected by C_{def}
 V is [φ -complete] if selected by v*
 V is [φ -defective] if selected by v_{def}

The notion of defectiveness is very vague. What does it mean for a set of φ -features to be defective or complete? Let us say that φ -features of an embedded T can be free or bound by the occurrence of φ -features of the matrix T. If they are bound, they must have the same value as the matrix occurrence, and hence the subject of both sentences must be correferential. As is well known, referential DPs must be free, and thus a set of bound φ -features cannot match the unbounded φ -features of a DP. Thus, the property of defectiveness/completeness of φ -features of T can be properly understood in terms of binding. If this is correct, the set of φ -features of control and raising infinitivals are identical; the difference lies in how the complete set of φ -features of matrix T is matched: by external Merge in control structures and by internal Merge in raising structures.

- (10) T is [φ -free] if selected by C* ([+fin]): [T_i [C_{fin} [T_j]]]
 T is [φ -bound] if selected by C ([-fin]): [T_i [C_{-FIN} ... [T_i ...]]]

Therefore, φ -features can surface on C in the form of finiteness distinctions or in the form of richer paradigms of person-number distinctions, as in the case of West Flemish. Moreover, there is a meaningful relation between the type of φ -features active on T and the choice of C: if C is [+fin] φ -features are free, while if C is [-fin], φ -features must be bound by a higher occurrence of φ -features. If these two properties are gathered, it becomes clear that the morphosyntactic phenomena match the semantic relation, or in other words, φ -features appear on both C and T, and thus the inflectional properties of V and the ability to license a subject are related to the properties of C. That is, the semantic relation between C and T is no more than an agreement relation between two occurrences of φ -features, which may or may not become morphologically overt or syntactically active: $[[\varphi]_C \dots [\varphi]_T]$.

2.1.2 Tense

In the generative tradition, den Besten (1977/1981) might have been the first to consider the C position as a [+Tense] position. In his revision of the standard analysis of root-phenomena in the 70s, he argued that all Complementizer Attraction Transformations were structure-preserving rules of the following type:

$$(11) \quad \begin{array}{cccccc} X - [+F_i] - Y - [C +F_i] - Z & & & & & \\ 1 \quad 2 \quad 3 \quad 4 \quad 5 & \rightarrow & & & & \\ 1 \quad 4 \quad 3 \quad e \quad 5 & & & & & \end{array}$$

where C is some constituent and
 F_i is some morphosyntactic feature

Den Besten states that the presentential domain must contain features other than [+wh]. Consider the following base rule (12), where [+wh] defines the landing site for interrogatives, according to the scheme rule (11).

$$(12) \quad S' \rightarrow [+wh] \quad [+T] \quad S$$

In this rule, den Besten argues, “COMP is rendered as [+T], i.e. as [+Tense]. [+T] lexicalizes as dat ‘that’ or of ‘whether, if’, and [-T] as om ‘for’. These features are not counterintuitive because in a sense there is agreement between om and a te-infinitival and between dat, of and the finite verb”.

Viewing COMP as [+T] led to a welcome consequence: it could obviate the obligatory rule of Complementizer Deletion when V was fronted, by respecting the general rule scheme (11); in the words of Besten:

«The position [+T] allows for another instantiation of rule scheme [11]. The resulting rule is Move Tense, i.e., Verb Fronting (Verb Second). This new formalization of the rule of Verb Fronting predicts that there will be Verb Fronting only if the corresponding lexical complementizer (dat or of) is absent –since the proposed finite verb occupies the

complementizer position. This way, the obligatory rule of complementizer deletion which was necessary under the formalization for Verb fronting [...] is obviated».

As illustrated in (2.1.1), complementizers can display not only [\pm fin] distinctions, but also more fine-grained person-number distinctions, arguing against the possibility that the occurrence of φ -features on C is essentially poorer than the one surfacing on Infl. As with the complementizer agreement paradigms, Irish offers a rather rich system of complementizers, marking a clause not only as subordinate, relative or interrogative, but also as being [\pm past] (see Chung & McCloskey 1987 and Hendrick 2000).

(13)		Nonpast	Past	(Chung & McCloskey 1987: 218)
	Subordinating	go	gur	
	Direct relative	a	a	
	Indirect relative	a	ar	
	Interrogative	an	ar	
	Matrix negative	ní	níor	
	Embedded negative	nach	nár	

The characteristics of φ -features tend to correlate with the characteristics of tense: usually φ -features are complete (they can license a referential DP) only when tense is complete. Both features depend on finiteness, and therefore on the choice of C (see Bianchi 2001a, b, Holmberg & Platzack 1995, Platzack 1995 and Zagana 1998 on the role of finiteness in defining the characteristics of φ -features and tense).

- (14) φ -features and tense are complete if they are selected by C* ([+fin])
 φ -features and tense are defective if they are selected by C_{def} ([-fin])

There are several well-known counterexamples to this tendency, which suggests that, although the characterization of tense and φ -features tend to correlate, there is no necessary one-to-one relation between them. Rigau (1995), for instance, has argued that the temporal characterization of prepositional adjunct infinitival clauses in Catalan and Spanish is defective, although they display a complete set of φ -features licensing a full DP. Interestingly, φ -features are morphologically realized in European Portuguese adjunct infinitivals (Raposo 1987). See also Rigau (1998) for the analyses of several issues related to these constructions. Examples in (15) are extracted from Rigau (1995).

- (15) a. Tothom va aplaudir en acabar el concert (Catalan)
 everybody applauded in.the to.finish the concert
 ‘Everybody applauded when the concert was finished’
- b. Todo el mundo aplaudió al acabar el concierto (Spanish)
 everybody applauded in.the to.finish the concert
 ‘Everybody applauded when the concert was finished’
- c. Ao entrares tu, a Maria saiu (European Portuguese)
 in.the to.enter-2SG you Maria left

‘When you went in, Maria left’

Since [\pm fin] Cs tend to correlate not only with the characteristics of tense but also with those of φ -features, it may be generally immaterial whether the former or the latter is the trigger of V-to-C movement in constructions like (16). That being the case, we shall consider for the moment that the trigger of such an operation is a featural conglomerate made of φ -features and tense, i.e., finiteness, as argued in Holmberg & Platzack (1995) and Platzack (1995). See section 5.4.1 of chapter 5 for a more detailed discussion on the trigger of V-movement in Germanic languages.

- (16) a. Diesen Buch las ich letztes Jahr
 this book read.PAST I last year
 ‘I read this book last year’
- b. Diesen Buch habe ich letztes Jahr gelesen
 this book have.PRES I last year read.PART
 ‘I have read this book last year’

In (2.1.1), the inability of φ -features to host a full DP is associated with binding relations; concretely, a set of φ -features selected by C_{def} needs to be bound by a higher set of φ -features. In both raising and control constructions C is [-fin] and φ -features are anaphoric. It can also be argued that tense is anaphoric when C is [-fin]. Consider the interpretation of the following control infinitivals, where the value of the embedded tense $TENSE_E$ is defined by the matrix tense $TENSE_M$.

- (17) a. John remembered [C_{def} to lock the door]
 $TENSE_M$: PAST \rightarrow $TENSE_E$: PAST
- a’. John remembers [C_{def} to lock the door]
 $TENSE_M$: PRESENT \rightarrow $TENSE_E$: PRESENT
- b. Peter tried [C_{def} to give up smoking]
 $TENSE_M$: PAST \rightarrow $TENSE_E$: PAST
- b’. Peter tries [C_{def} to give up smoking]
 $TENSE_M$: PRESENT \rightarrow $TENSE_E$: PRESENT

If the event of remembering occurs prior to the speech time ($TENSE_M$: PAST), then the event of locking the door must also occur prior to the speech time ($TENSE_E$: PAST). The reading of John locking the door simultaneously or after the speech time is not available when $TENSE_M$: PAST. When the matrix event is present, the embedded event can occur neither before nor after the speech time. Note that a habitual reading is prominent in (17.a’) (‘It is habitual that John remembers to lock the door’), unless an adverb like ‘now’ or ‘right now’ is present (‘John remembers right now to lock the door’).

want.3PL SUB resign.3SG
 ‘They want her/him to resign’

c. ja zhelaju cto-by ona ushla (Russian)
 I wish.1SG that-BY she go.PAST
 ‘I wish that she should go’

In other languages mood inflectional particles co-occur with mood complementizers (Quer in press).

(22) a. Ion vrea ca sa∫ manince (Romanian)
 Ion want-3SG that.SUB SUB eat.3SG
 ‘Ion wants him/her to eat’

b. Jani do që të hajë (Albanian)
 Jani want.3SG that.SUB SUB eat.3SG
 ‘Jani wants him/her to eat’

So there is no doubt that, like φ -features and tense features, mood features can be marked on both C and Infl.

The classic Latin system of complementizers offers a rather clear and rich instance of the correlation between the choice of complementizer and the verbal mood ending. See table (23), extracted from Valentí i Fiol (1979/1992: 156), an introduction into Latin syntax where the correlation between complementizers and verbal mood endings is very explicitly drawn. Examples (24-27) are also extracted from this work. The English translations are mine.

(23)

Completive clauses	Assertive	Infinitival phrases	
		Introduced by conjunctions	By <i>quod</i> : indicative verbal mood By <i>ut, ne, quin, quominus, quod</i> : subjunctive verbal mood
	Indirect questions	Introduced by interrogative pronouns or adverbials (<i>quis/quid, ut, ubi</i> , etc): subjunctive verbal mood	

In Latin, indirect questions always display subjunctive mood.

(24) a. Ego instare [sic] ut mihi responderet quis esset (Cic., *Verr.*, 2, 77)
 I insist.INF that I.DAT told.SUBJ who.NOM was.SUBJ

‘I insisted that he told me who he was’

b. Non properaui scire quis illis esset rerum mearum
 NEG hasten.PAST know.INF which.NOM them.ABL was.SUBJ wells.GEN my.GEN

status, quid afferent (Sen., *Lucil*, 77, 3)
 state.NOM what.ACC produced.SUBJ

‘I did not hasten to know which the state of my wells was in them, nor what they produced’

Assertive embedded clauses may display indicative or subjunctive mood depending on the selecting category. Interestingly, the choice of complementizer tends to correlate with the value of the mood ending; generally, ‘quod’ correlates with the indicative mood (25) and ‘ut/ne’, ‘quin’, ‘quominus’ with the subjunctive mood (26).

(25) a. Eumeni inter Macedonas viventi multum detraxit quod
 Eumenius.DAT among Macedonians-ACC living much harmed that

alienae erat civitatis (Nep., 18, 1)
 foreigner.GEN was-IND city-GEN

‘That he was from a foreigner city harmed Eumenius very much while he was living among Macedonians’

(26) a. Peto igitur ab te ut omnibus rebus quod sine
 beg.1SG thus from you.ABL that all-ABL things-ABL that without.ABL

molestia tua facere possis ei commodes (Cic., *Fam.*, 13, 35)
 nuisance.ABL your.ABL do.INF can.SUBJ he.DAT please.SUBJ

‘I thus beg you to please him in all the things that you can without [causing] your nuisance’

b. Quibus non humana ulla neque divina obstant quin socios,
 them.DAT not human any nor divine avoid.3PL that allies-ACC,

amicos, procul, iuxta sitos, inopes potentisque trahant, excidant
 friends.ACC, distant close positions, weak strong.AND push.SUBJ destroy.SUBJ
 (Sall., *Mithr.*, 17)

‘No human or divine law prevents them from pushing and destroying their allies or their friends –close or distant, weak or strong’

c. Neque impedio quominus susceptum negotium
 Not oppose-1S that assumed.ACC obligation.ACC

gerere possis (Cic., *Fam.*, 13, 5)

carry out-INF can-SUBJ

‘I am not against that you may carry out the assumed obligations’

However, this correlation is not perfect: although ‘ut’, ‘ne’, ‘quin’, and ‘quominus’ never co-occur with indicative mood, ‘quod’ can co-occur with both subjunctive and indicative verbal endings in certain cases, especially when the embedded clause is selected by affective verbs like ‘laudo’ (27.a vs. 27.b). This is usually attributed to the ‘subjunctive’ status of the predicate embedded under the affective verb.

- (27) a. Laudat Panaetius Africanum quod fuerit abstinens (Cic., *Off.*, 2, 22)
laud.3SG P.NOM A.ACC that was.SUBJ neutral
‘Panaetius lauds the African for keeping neutral’
- b. Quod viris fortibus honos habitus est, laudo (Cic., *Rosc. Am.*, 47)
that men-ABL strong.ABL honor customary is.IND laud
‘That honor is customary in strong men, I laud’

Despite the apparent exceptions, the correlation between the choice of complementizer and the value of mood is prevalent:

- (28) C = [+ embedded, -assertive] → mood: [subjunctive]

C = [+embedded, + assertive]:

(a) ut/ne, quin, quominus → mood: [subjunctive]

(b) quod → mood: [indicative]*

*‘quod’ can select subjunctive mood in some special cases, as when it is selected by an affective verb

Finally, some cases of V-to-C movement are sensitive to mood distinctions; particularly, in formal English and formal French, V is fronted when the conditional complementizer is absent and the verb shows up with non-factual overt morphology. In formal English, for instance, ‘had’ and ‘should’, but not other auxiliaries like ‘have’ or ‘is’, can be fronted in the absence of a complementizer.

- (29) a. Had I known you were coming, I would have stayed longer
a’. If (*had) I (had) known you were coming, I would have stayed longer
b. Should you see him, tell him that his letter has already arrived
b’. If (*should) you (should) see him, tell him that his letter has already arrived
- (30) a. *Have you failed the exam, you will be punished
b. *Is it raining, let’s stay at home

Similarly, in formal French, according to Rizzi & Roberts (1989: 4), a conditional clause can be introduced either by a complementizer or by a verb in the conditional mood.

- (31) a. Si tu avais fait cela, ...
 If you had done that, ...
- b. Aurais-tu fait cela, ...
 Had you done that,
- c. *Si aurais/avais-tu fait cela, ...
 If had you done that, ...

From this summary we may conclude that mood distinctions may be displayed on complementizers and also as free inflectional particles and verbal endings; the value of mood seems to correlate with the choice of C, as the example of Latin clearly illustrates, and mood features can be syntactically active on C triggering V-to-C movement.

2.1.4 Modality

To the best of my knowledge, modal distinctions (epistemic necessity, deontic possibility, etc.), unlike ϕ -features, tense and mood, are not marked on complementizers and there are no instances of V-to-C movement triggered by modal features. However, before arriving at any conclusion, we must consider whether there is a semantic relation between the choice of clause typing features and the modal value of a proposition.

The idea that modal features are related to the C-system is not new in the literature. Here I shall focus on two possible pieces of evidence, those that seem to me more compelling. Different analytical arguments for this idea are given in Drubig (2001) and Rochette (1988), who attribute to C the emergence of opacity effects in certain subjunctive embedded clauses in certain Romance languages when they contain a modal verb (Picallo 1985), assuming that C is the site where sentential operators must appear at LF. The emergence of opacity effects does not necessarily indicate that the opacity trigger is associated with C (although this is a rather common view), but I prefer to leave it as an open question.

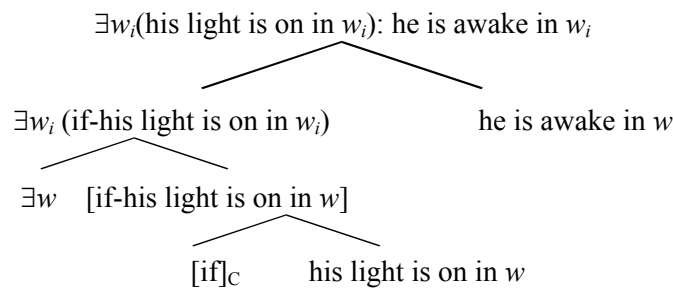
One may consider the possibility that the complementizer of conditional sentences ('if' in English) encodes a modal feature. A possible analysis might consist in viewing 'if' as a modal quantifier; accordingly, in 'if his light is on, he is awake', the semantic role of 'if' is to define the set of possible worlds where the proposition 'his light is on' holds as a subset of the set of possible worlds where the proposition 'he is awake' holds. Or, equivalently, the clause introduced by 'if' ('his light is on') would introduce the restriction for 'if' and the matrix clause ('he is awake') the nuclear scope. Generally, 'if p , q ' would mean:

$$\{w: p(w)\} \subseteq \{w: q(w)\}, \text{ or equivalently, } \\ \forall w [p(w)]: q(w)$$

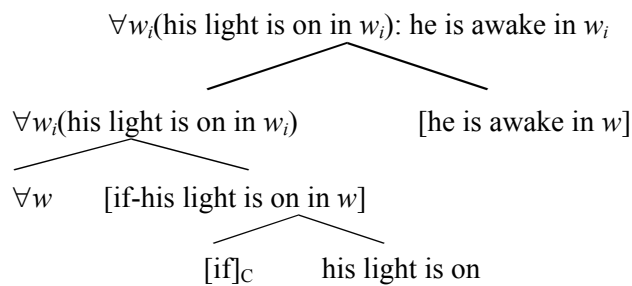
In this case, the only difference between the meaning of 'if' and the universal epistemic quantifier 'must' would be that the restriction of the former is overt: $p(w)$. However, this treatment seems inadequate when dealing with the interaction of 'if'-clauses and

quantificational adverbs ('always if...', 'rarely if...', 'often if...'), as observed by Lewis (1975), who argues that, in these particular cases, 'if'-clauses bear the restriction for the quantificational adverb. This insight is extended to 'if'-clauses in general by Kratzer (1986) in her claim that "the history of the conditional is the story of a syntactic mistake. There is no two-place *if...then* connective in the logical forms for natural languages. *If*-clauses are devices for restricting the domains of various operators (p. 656)". Adopting Kratzer's and Lewis' ideas, the (simplified) LF representations of conditional sentences should be as follows:

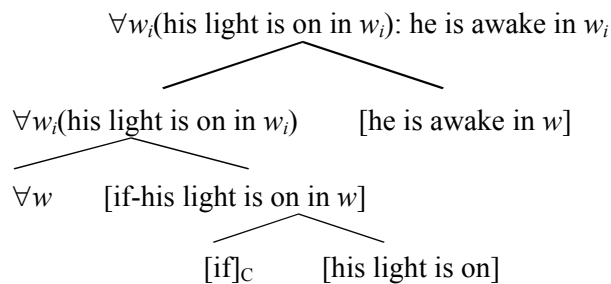
- (32) a. Sometimes if his light is on he is awake (quantificational adverb)



- b. If his light is on he is awake (covert universal epistemic quantifier)



- c. If his light is on, he must be awake (overt universal epistemic quantifier)



An 'if'-clause is not strictly a modal proposition, but it serves as a restriction for a modal operator. Kratzer's analysis suggests that previous analyses of the meaning of 'if' were indeed dealing with the meaning of the operator that the 'if'-clause restricts. Consequently,

we may define ‘if’ as a [+clause typing] category responsible for defining a proposition as an overt restriction that contains a variable ranging over possible worlds, a useful instruction for the levels of interpretation of the C-I system.

There is another case where the characterization of C determines the modal interpretation of a clause. Non-finite questions and non-finite relatives have an obligatory modal reading although there is no overt modality feature. Unlike conditionals, these infinitival clauses are not interpreted as restrictions for modal operators, but exactly as modal propositions.

Infinitival questions

- (33) a. She knows [_{CP} where to go]
b. I finally found out [_{CP} what to do]
c. [_{CP} How to learn French in five days]

Infinitival relatives

- (34) a. The people [_{CP} to talk to at the party]
b. The book [_{CP} to read for tomorrow]
c. The man [to fix the sink] (from Bhatt 1999)

The presence of a covert modal semantic instruction depends on the choice of C, since modal readings are not obtained in all non-finite contexts. For instance, they are absent in the complement clauses of raising and control verbs.

- (35) a. He prefers [_{CP} to talk at the party]
b. He seems [_{CP} to talk at the party]

This suggests that the covert modal feature is derivative from C. If C = [-FIN, +question] or if C = [-FIN, +relative], the proposition must be interpreted modally.

- (36) C [-fin, +qu/+rel] → C = [+modal]

The idea that the source of covert modality in infinitive clauses is C is developed in detail in Bhatt’s dissertation (Bhatt 1999). Bhatt (1999: chapter 3), observes that infinitival relatives with an object gap are interpreted necessarily with a modal value, whereas infinitival relatives with a subject gap need not be interpreted modally under special licensing conditions, e.g., under the scope of modifiers such as superlatives, ordinals or the nominal *only*.

Infinitival relatives with an object gap (Bhatt 1999: 12)

- (37) a. a book [Op_i [PRO to read t_i]]
b. a knife [Op_i [PRO to cut the bread [with t_i]]]
c. the person [Op_i [PRO to believe [t_i to be innocent]]]

Infinitival relatives with a subject gap (Bhatt 1999: 46)

- (38) a. This is the best argument to be considered by the committee
 ≈ (... which has been considered...) (non-modal)
 ≈ (... which is to be considered ...) (modal)
- b. Joseph was the first/only/tallest man to walk on the moon last year
 ≈ Joseph was the first/only/tallest man who walked on the moon last year (non-modal)
 ≈ Joseph was the first/only/tallest man who was supposed to walk on the moon last year (modal)

In light of the asymmetry between object and subject gaps in infinitival relatives, Bhatt proposes that “in general the infinitival clause that appears in subject infinitivals is ambiguous between an infinitival clause that is interpreted modally and an infinitival that is interpreted non-modally. The infinitival clause that is interpreted modally does not have special licensing conditions and can appear wherever subject infinitival relatives appear. [...]. The infinitival clause that is interpreted non-modally, however, needs to be licensed and can therefore appear only in certain environments (p. 47)”.

Note that covert modality in non-finite questions and relatives (37-38) cannot be epistemic. Thus, the modal structure of finite clauses is more full-fledged than the modal structure of non-finite clauses. This indicates that modality is derivative from C not only in non-finite clauses but also in finite clauses.

- (39) C [+fin] → modality is overt and can be epistemic or deontic
 C [-fin, +qu/rel] → modality is covert and cannot be epistemic

In sum, a modal feature is encoded in the complementizer ‘if’ of conditional clauses, which introduces a restriction of possible worlds for a quantifier, and the characteristics of modality are related to the choice of C. Concretely, if C = [+fin], modality is full-fledged, and if it is [-fin, question/relative], it cannot be epistemic; the modal interpretation is forced in non-subject infinitival relatives and optional in subject infinitival relatives when a special licenser appears.

2.1.5 Negation

Complementizers are known to encode negative features in several languages. For instance, in Latin the subordinator ‘ne’ is the negative counterpart of ‘ut’.

- (40) a. Imperat ut mihi respondeas
 ‘He/she orders that you reply me’
- b. Imperat ne mihi respondeas
 ‘He/she orders that you do not reply me’

Celtic languages have negative particles in both main and embedded sentences, which are usually analyzed as clause typing categories. See table (13) above, and the following constructions, from Hendrick (2000: 22-23). Note that Breton (44) presents two negative markers: a preverbal one, typing a clause as [\pm embedded], and a postverbal one. This is similar to the constructions in (22), where inflectional mood particles co-occur with subjunctive complementizers.

- (41) Irish
- a. *Ní* heo é an bord
 NEG here it the table
 ‘This here is not the table’
- b. Deir sé [*nach* ‘eo é an bord]
 say he NEG here it the table
 ‘He says this here is not the table’
- (42) Gaelic
- a. *Nach* robh còta aice
 NEG was coat at3SF
 ‘She didn’t have a coat’
- b. Bha Ceiteag ag ràdh *nach* robh còta aice
 was kaitei prog say NEG was coat at3SF
 ‘Katie was saying that she didn’t have a coat’
- (43) Welsh
- a. *Ní* welodd y bachgen y dyn
 NEG saw the boy the man
 ‘The boy did not see the man’
- b. Dywedodd *nad* oedd y bachgen yn dod
 say NEG was the boy prog coming
 ‘He said the boy was not coming’
- (44) Breton
- a. *Ne* lennan *ket* al lizher
 NEG read1S NEG the letter
 ‘I don’t read the letter’
- b. Gouzout a ran mat *ne* lennan *ket* al lizher
 knowPRT do.1S well NEG read1S NEG the letter
 ‘I know I don’t read the letter’

As shown above, the properties of φ -features, tense and modality depend on the choice of C; in a sense these features are ‘shrunk’ in non-finite contexts: φ -features and tense must be

anaphoric and modality must be non-epistemic when C = [-fin, qu/rel]. Negation differs from these Infl-like features in that it is not shrunk in infinitival contexts (45).

- (45) a. He prefers not to go out at night
b. He told me where not to buy cheese
c. He seemed not to be able to manage that problem
d. Hafdis knows where not to get gas (Bhatt 1999 158-161)

Clearly, negation does not appear universally on the C-area: like other Infl-like categories, it can also show up on the Infl-area. The chameleonic status of sentential negation is again present in grammatical studies of negation. On the one hand, several comparative studies, such as Pollock (1989) and Zanuttini (1997), have focused on the position of negation in the Infl-area, and on the other hand, Klima (1964) has attempted to substantiate the idea that negation is in a presentential position. In fact, Lasnik (1972) and Laka (1990) have each defended, on different grounds, that negation is generated both on the C-area and on the Infl-area. Here I shall review a sample of syntactic properties of sentential negation that demonstrate how this feature shares properties with *wh*-features. See Horn (1989) and Horn & Cato (2000) for a review of the syntactical and logical analyses of negation.

In English, negative phrases can move to the C-area and trigger subject-auxiliary inversion, or Aux-to-C movement. This syntactic operation, which shows that the type feature that defines the landing site of negative phrases operates from a position of the C-area, is not general among nominal quantifiers. Observe the following minimal pair, extracted from Rizzi (1990: 53).

- (46) a. With no job would Bill be happy
b. *With some job would Bill be happy

As observed by Rizzi (1990: 53), who follows Klima (1964), this is not a characteristic property of negative phrases, but rather of the so-called affective quantifiers (47), which are defined by the ability of licensing negative polarity items (48).

- (47) a. Few people did anything
b. *A few people did anything

- (48) a. With few jobs would Bill be happy
b. *With a few jobs would Bill be happy

These syntactic patterns suggest that negation is a value of a broader feature which also includes quantifiers such as 'few', and not an isolated feature. This feature, which we call affectiveness, seems to be active on C in English, as *wh*-features.

Chomsky's (1957) analysis of the insertion of 'not'/'n't' into the auxiliary verb phrase in English suggests that negation is C-related. Chomsky claims that if negation is not included in the kernel sentence but is a transformation that applies to it, all instances of auxiliary support are accounted for in a very simple and uniform way. Chomsky (1957) postulates that

a transformation T_{not} is responsible for negation and that a transformation T_A is responsible for emphatic affirmation. The first transformation denies the kernel sentence (49) and the second emphatically asserts it and requires an extra heavy stress (50).

- (49) a. John doesn't arrive
 b. John can't arrive
 c. John hasn't arrived

- (50) a. John DOES arrive
 b. John CAN arrive
 c. John HAVE arrived

Laka (1992) has shown that very similar patterns are observed in Basque, where sentential negation and emphatic affirmation trigger auxiliary-movement (51). Laka postulates that a more abstract feature is responsible for affirmation and denial, which she terms Σ , a label that attempts to suggest the notion of speech act. Σ heads the projection where V moves to to match sentential negation or emphatic affirmation.

- (51) a. Mari joan da (declarative sentence)
 Mary left has
 'Mary has left'
- b. Mari ez da joan (negative sentence)
 Mary NEG has left
 'Mary hasn't left'
- c. *Mari da joan (declarative sentence)
 Mary has left
 'Mary has left'
- d. Mari da joan (emphatic affirmative sentence)
 Mary has left
 'Mary has left'

As Chomsky shows, the advantage of the transformational analysis becomes clear once we observe that other instances of transformations that apply to kernel sentences trigger auxiliary-insertion. Transformations T_q and T_w , responsible for total (52) and partial (53) questions, require auxiliary insertion and were considered to apply to kernel sentences.

- (52) a. do they arrive?
 b. can they arrive?
 c. have they arrived?
 d. are they arriving?
- (53) a. what did John ate?
 b. where can they arrive?

- c. when have they arrived?
- d. why are they coming?

If the particle ‘so’ is the affirmative counterpart of ‘neither’, then emphatic affirmation, one of the two possible values of Σ , can also surface on C, like Affective-features, and trigger auxiliary insertion and auxiliary-subject inversion. This possibility is also considered by Chomsky (1957), who postulates a T_{so} , analogous to T_{neg} , T_A , T_q , and T_w . Chomsky does not explicitly treat the particle ‘so’ as an affirmation particle, but he postulates a special transformation T_{so} . Bearing in mind the opposition ‘so’/‘neither’, it seems that we can account for the syntax of these particles simply in terms of T_{neg}/T_A , with no empirical loss and a gain in simplicity.

- (54) a. John arrives and so do I
b. John can arrive and so can I
c. John has arrived and so have I
- (55) a. John doesn’t arrive and neither do I
b. John can’t arrive and neither can I
c. John hasn’t arrived and neither have I

These observations reveal that affirmation, negation and other affective categories indeed constitute a sole feature, which may be called Σ , adopting Laka’s terminology. Σ can appear in a [+clause typing] position or in a [-clause typing] position, and can trigger movement of Σ -phrases and of auxiliaries to the C-area, as well auxiliary-insertion in the Infl-area.

It could be argued that Chomsky’s analysis of auxiliary-insertion is no longer valid, as the distinction between obligatory and optional transformations has been long abandoned. In the model where the notion of “kernel sentence” was formulated (see Chomsky (1955/75-1957)), the proposed architecture of grammar consisted of (a) a context-free rewriting base which generated phrase markers, and (b) a transformational component that took phrase-markers as input (generated by the base) and yielded new phrase markers. In this framework, a kernel sentence was understood as a structure generated solely by obligatory transformations; optional transformations took kernel sentences as input. A mono-clausal affirmative active sentence was a kernel sentence, and the transformations of ‘Passive’ and ‘Question’ were instances of optional transformations. Putting aside the accuracy of the technical distinction between optional and obligatory transformations, and the reasons for which the sketched model was abandoned, it seems that the syntactic similarities between Σ -features and *wh*-features are real, or even that these two classes of instructions may belong to a broader class.

One way to conceive of the similarity between Σ -features and *wh*-features is to view denial, (emphatic) affirmation and query as three different values of a force feature relative to a proposition. Thus, the illocutionary force of a proposition can be emphatically affirmative, negative or interrogative. This allows us to account for the syntactic similarities between Σ -features and *wh*-features without needing to resort to the dichotomy between optional and transformational operations.

Another legitimate qualm about the proposed version of Chomsky's (1955/75-1957) analysis is that negation and emphatic affirmation trigger auxiliary-insertion (47-48) into the Infl-area, whereas *wh*-features trigger auxiliary insertion into the C-area, an areal distinction not present in Chomsky's (1955-1957) framework. This is not a real problem, since negation and affirmation, like *wh*-features (52, 53), can surface on C (40-44) and can trigger auxiliary raising to C (46, 51.b, 54, 55).

One might ask why in English the negative marker cannot appear in the C-area, unlike negative phrases, if it conveys a force indication. This may be due to a cross-linguistic parameter relative to the characterization of negative markers: (a) certain languages, such as Celtic languages, present a negative marker lexically defined as [+embedded] and another one as [-embedded], and thus negative markers surface on C in both embedded and matrix clauses; (b) other languages, such as Latin, display negative markers specified as [+embedded], but not as [-embedded]; in this case, a negative marker surfaces on C only in embedded clauses ('ne' in Latin); (c) in other languages, such as Romance languages or English, negative markers are not specified as [\pm embedded], and thus they are forced to occur in Infl positions; and finally, (d) other languages display a negative marker defined as [+embedded], a negative marker defined as [-embedded] and a negative marker not defined as [\pm embedded]; such is the case of Breton. As expected, syntactic variation is reduced to the morpholexical specification of functional categories.

From this we may conclude that the English negative marker has a hybrid status: it is a head that conveys a value of the illocutionary force feature but does not convey the [\pm embedded] distinction. This second property forbids the negative marker from appearing on C (only heads specified as [\pm embedded] can surface as complementizers), and thus forces the type feature responsible for hosting the negative marker to spread to an Infl-position in English. As argued above, this morpholexical property distinguishes English negative markers from Celtic negative markers. The first property accounts for auxiliary-insertion, a syntactic operation triggered by *wh*-features and Σ -features, two subtypes of force features.

In sum, complementizers may display negative or affirmative forms, and negation, along with other affective categories, can trigger V-to-C movement. If Σ -features (affectiveness, emphatic affirmation and negation) are viewed as a class of illocutionary force features, along with *wh*-features, all instances of auxiliary-insertion can be attributed to a homogeneous class of features.

2.2 Conclusion

The so called C-Infl link may be materialized in the following three properties and generalized to various Infl-like feature, φ -features, tense, mood, modality and negation:

Property (1): complementizers can 'replicate' Infl-like features

φ -features, tense, mood and negation can surface on Infl-positions or on C-positions. The complementizer of conditional clauses ('if') has been argued to encode a modal value for the

proposition it introduces. Concretely, this complementizer marks the clause it heads as a restriction for a higher quantifier.

Property (II): there is a correlation between the characteristics of features that surface on Infl and the choice of C

The characteristics of ϕ -features, tense and modality correlate with the choice of C:

If C = [-fin], ϕ -features and tense are anaphoric

If C = [-fin, qu/rel], modality is not epistemic

If C = [+fin], ϕ -features and tense are free, and modality is full-fledged.

The classic Latin system of complementation has been argued to display a clear and rich correlation between the mood of verbal endings and the choice of C.

Property (III): Infl-like features are involved in triggering V-to-C movement

Tense, mood and negation are involved in triggering V-to-C movement. It has been argued that negation is not an isolated feature, but rather a value of a broader head Σ , along with affective categories and emphatic affirmation. Furthermore, Σ -features have been associated with *wh*-features through the claim that they both convey illocutionary force distinctions. Consequently, all operations of auxiliary-insertion in English are triggered by force features. No evidence has been found for the possibility that modality distinctions trigger V-to-C movement.

Chapter 3. Discontinuous Syntactic Patterns

This chapter begins by considering the possibility pointed out by Chomsky (2005b) that Infl-like features are base-generated on C and inherited by T/Infl; accordingly, the C-Infl link would derive from inheritance mechanisms. This is called the Generalized Feature Inheritance Theory, whose plausibility depends on the adequacy of the theory of phases: there is an important asymmetry regarding the role that C and T/Infl play in syntactic derivations; more precisely, as the Phase-Impenetrability Condition states, the former head, but not the latter, is a phase head and defines its complement as inaccessible to further probes. After reviewing the Phase-Impenetrability Condition, we conclude that there is no clear reason for such an asymmetry, and thus the Generalized Feature Inheritance Theory loses plausibility. The Phase-Impenetrability Condition is replaced by the Relativized Opacity Condition to account for subject islands, wh-islands and the requirement whereby both A'-movement and A-movement are successive cyclic. The C-Infl link is derived from the assertion that C and Infl constitute a discontinuous syntactic pattern: C and Infl display different polarities of the same lexico-semantic feature ([±clause typing]). The semantic instruction provided by Infl-like features (to trigger referential displacement) is orthogonal to the [±clause typing] distinction, and thus they are simultaneously merged in both poles of the discontinuity, without lowering or raising mechanisms. This reasoning is extended to account for the ν -V link and the P-K link.

3.1 The source of the C-Infl link

In the preceding chapter, I addressed the illustration of a traditional topic in grammar: the relationship between complementizers and inflectional heads as to ϕ -features, tense, mood, modality and negation. The principal objective of the present chapter is to define the source of the C-Infl link. I shall take as a starting point Chomsky's feature inheritance theory.

3.2 Probe-goal relations in a phase

Chomsky (2005b) has recently provided interesting evidence for the idea that C and T/Infl constitute a unit concerned with the timing of A'- and A-movement in the C phase. One of the many implications of Chomsky's findings is that a phase not only defines the syntactic domain that becomes impenetrable (due to transfer operations), but also the level at which syntactic operations apply, since both ϕ -features and peripheral features (π -features) of a phase probe their goal in parallel.

3.2.1 Subextraction (I): subject islands and ϕ -features

Chomsky's reasoning begins by addressing subject islands; as observed by Huang (1982), subextraction of a *wh*-phrase from a subject DP contrasts with the subextraction of a *wh*-phrase from an object DP.

- (1)
- a. *about who did [stories ~~about who~~] terrify John
 - b. about who did John read [stories ~~about who~~]
 - a'. *of who did [a critic ~~of who~~] see you

- b. of who did you see [a critic ~~of who~~]

Chomsky's new observation is that the surface subject position is not the island, but the base-position SPEC- v^* , since island effects are attested solely with agentive subjects, and not with subjects of passive or unaccusative verbs. That is, subextraction of a PP from a non-agentive subject (3) patterns like subextraction from an object (2), and not like subextraction from an agentive subject (4).

Subextraction from an object (Chomsky 2005b: 14)

- (2) a. it was the CAR (not the TRUCK) of which they found the (driver, picture)
b. of which car did they find the (driver, picture)

Subextraction from the subject of a passive verb (Chomsky 2005b: 14)

- (3) a. it was the CAR (not the TRUCK) of which the (driver, picture) was found
b. of which car was the (driver, picture) awarded a prize

Subextraction from an agentive subject (Chomsky 2005b: 14)

- (4) a. *it was the CAR (not the TRUCK) of which the (driver, picture) caused a scandal
b. *of which car did the (driver, picture) cause a scandal

The inactivity of SPEC-T is not a sufficient condition to account for these patterns; if the key were the inactivity condition, subextraction from a subject would always be unavailable, regardless of its base-position. Indeed, if π -features subextracted the PP from the DP when it is in SPEC-T, the contrast between agentive and non-agentive subjects should be neutralized. To account for the contrast between subextraction from agentive and from non agentive subjects, φ -features and π -features must try to extract their respective tokens in parallel and when they are in the base-position.

More precisely, v^* can subextract a *wh*-phrase (by pied-piping, a PP) from an object DP because such a DP is in the search domain of v^* ; however, a *wh*-phrase cannot be subextracted from an external argument by v^* because the external argument is outside the search domain of v^* . Consequently, when a type *wh*-feature of C searches a suitable token, it can see a *wh*-phrase at the edge of v^* (previously subextracted from an object DP by v^*), but not a *wh*-phrase that is a constituent of an external argument when the phase head C is active. The work done by the v^* phase head is thus crucial to explain the asymmetry between subextraction from accusative DPs and nominative agentive DPs.

Subextraction from a non-agentive subject (3) may be accounted for in two different ways:

- (i) C can subextract the *wh*-phrase without problems from the domain of v_{def} because v_{def} does not constitute a phase boundary, and hence, intermediate movement to the edge of $v_{\text{def}}P$ is not necessary

- (ii) v subextracts the *wh*-phrase from the DP and relocates it in an edge position, and thus C can have access to it.

Chomsky tends to adopt option (i) in his work, keeping as close as possible to the ideal that those units with a complete set of φ -features are the units transferred to the outer systems. However, the existence of an edge of $v_{\text{def}}\text{Ps}$ has been independently motivated by Legate (2003), who observes that passive, unaccusative and raising $v\text{Ps}$ offer reconstruction sites for *wh*-phrases. This being the case, we must conclude to conclude that v_{def} subextracts a *wh*-phrase from its search domain and relocates it in the edge of $v_{\text{def}}\text{P}$, exactly as v^* does.

Interestingly, a *wh*-phrase can be subextracted from an agentive subject in raising and ECM constructions. This reveals that a probe can penetrate into a SPEC that is in its search domain. However, a probe cannot penetrate into a SPEC that belongs to the edge of a passed phase (SPEC- v^*). Thus, the source of subject islands “is search that goes too deeply into a phase already passed, not the difference between base and surface position (Chomsky 2005b: 20)”.

- (5) a. it is the CAR (not the TRUCK) of which [the (driver, picture) is likely [t to cause a scandal]]]
 b. of which car is [the (driver, picture) likely [t to [cause a scandal]]]
 c. of which car did they believe the (driver, picture) to have caused a scandal

In the derivation behind sentence (5), matrix C detects ‘which car’ when the DP ‘the (driver, picture) of which car’ is in SPEC- T_{def} . Since no phase is penetrated, C can subextract the *wh*-phrase.

According to the above considerations, subextraction from SPEC-C should be degraded. This prediction is verified by the following English and Catalan constructions. See Gallego (2006) and references cited therein for further evidence.

- (6) a. *[_{CP} [which car] did you ask [_{CP} [which driver of ~~which car~~] Peter saw]]
 b. *[_{CP} [which director] did you ask [_{CP} [which film of ~~which director~~] Peter bought]]
 c. *[_{CP} [who] did Mary ask [_{CP} [which novel of ~~who~~] you read]]
- (7) a. *[_{CP} [de quin cotxe] em vas demanar [_{CP} [quin conductor ~~de quin cotxe~~] ha vist en Pere]]
 b. *[_{CP} [de quin director] em vas demanar [_{CP} [quina pel·licula ~~de quin director~~] ha comprat en Pere]]
 c. *[_{CP} [de qui] em va demanar la Maria [_{CP} [quina novel·la ~~de qui~~] vas llegir]]

We ought to note that there is a potential problem for the idea that a subconstituent at the edge of C cannot be readily probed: subextraction of a *wh*-phrase₂ from a *wh*-phrase₁ is less degraded if *wh*-phrase₁ is a subject, a phenomenon observed by Torrego (1985) founded on Spanish data. See also Chomsky (1986a).

- (8) a. ??/*esta es la autora de la que varias traducciones han ganado premios internacionales

‘This is the author by whom several translations have won international awards’

b. de que autora no sabes qué traducciones han ganado premios internacionales
‘By what author don’t you know what translations have won international awards’

This issue has recently been reconsidered by Uriagereka (2004) and Gallego (2006), extending Torrego’s original contrast and bringing several intricate factors into consideration; Uriagereka argues that subextraction in (8.b or 9) is possible because the embedded *wh*-phrase, being topicalized and having a *de re* reading, is not a SPEC but an adjunct, which, in Uriagereka’s model, allows for subextraction; if a *de dicto* reading is forced, as in (9), subextraction is degraded. Uriagereka’s suggestion that topics are adjuncts, and not SPECs, and that adjuncts allow subextraction is not uncontroversial.

(9) de que autora no sabes qué traducciones están a la venta ...
of what author not know.2SG what translations are-3PL at the sale
‘Of what author don’t you know that what translations are on sale

... # y de hecho dudas que haya ninguna
and of fact doubt that there is.SUBJ.3SG none
... and actually you doubt that there is any?’

Gallego, however, argues that the key factor is not the topical status of the *de re* embedded *wh*-phrase, but the presence of the negative marker, which forces a presuppositional/D-linked/*de re* interpretation. According to Gallego, if the negative marker of (10) is eliminated, subextraction is degraded.

(10) De que autora ^{??/*} (no) sabes qué traducciones están a la venta?

It is worth stressing that Torrego’s original contrast, or its Catalan or English counterparts, is quite weak, with some disagreement among speakers. The same consideration applies to (9-10): although it may be true that in general the *de re* reading of the embedded *wh*-phrase makes subextraction easier or that the presence of negation tends to facilitate the *de re* reading, I would resist the claim that they are necessary conditions that determine whether or not extraction from a domain is possible.

3.2.2 The Generalized Feature Inheritance Theory

Chomsky (2005b) interprets the parallel probing effects and the observation that the characterization of φ -features on T correlates with C through what may be called a Feature Inheritance Theory (FIT), which can be summarized as follows:

Feature Inheritance Theory

- (i) both φ -features and π -features are generated on C
- (ii) both φ -features and π -features operate in parallel

- (iii) φ -features surface on T only derivatively
- (iv) The feature inheritance mechanism has the function of bringing semantic distinctions (arguments vs. operator-variable constructions) into a syntactic representation
- (v) C, as a phase head, has a central role in the generation of syntactic objects
- (vi) In the lexicon, T lacks φ -features

Note that the conclusion that φ -features and π -features of a phase work in parallel (ii) does not necessarily lead to the idea that φ -features are generated on C; the attested patterns of subextraction do not have any implication on where features are generated but only on the ordering of probing operations. The possibility that φ -features were generated on T would be compatible with the evidence of subextraction provided above, with the proviso that φ -features operated in parallel with π -features, since they belong to the same domain where syntactic operations apply, i.e., to the same phase. In fact, the FIT is a consequence of the theory of phases. If C, and not T, has a central role in the generation of syntactic objects (v), and the value of T as to φ -features correlates with the choice of C, the simplest explanation is that φ -features, like π -features, are base-generated on C (i). Accordingly, the probes of both A'- and A-movement are generated in the same category, the phase head, and φ -features surface only derivatively on T (iii), T being a lexical item generated without φ -features (vi); the derivational device responsible for spreading φ -features from C to T is an inheritance mechanism, which yields semantically specialized positions: according to Chomsky, SPEC-T becomes specialized for arguments and SPEC-C for operators (iv). Thus, the source of feature spreading is a C-I requirement, in the spirit of the SMT.

There are several other conceivable implementations for the C-T link, such as selection rules or coindexing algorithms; however, if there is independent reason for the idea that C, and not T, has a special generative status, these alternatives are just unnecessarily complicated ways of saying that φ -features are generated on C and spread to T in order to yield semantically dedicated positions.

The v-V system appears to have the same properties as the C-T system. Ideally, φ -features inheritance should be a property of phase heads, and thus v* should also be generated with φ -features, which should be inherited by V. If SPEC-V, and not SPEC-v*, offers an A-position for an accusative DP, as T offers an A-position for a nominative DP, then, as Chomsky (2005b) observes, “the intriguing but puzzling conclusions about raising of object to SPEC-V (p. 15)” can be accounted for automatically. Chomsky refers to an array of observations by Postal (1974) and further discussed by Lasnik & Saito (1991) and Lasnik (2001, 2002), whereby a verb precedes an object which in turn c-commands an adjunct. Patterns in (26, 27), taken from Lasnik (2001), illustrate that ECM constructions provide a DP with a position high enough to bind an anaphor and to license an NPI into a VP adjunct; as expected, when the relevant DP is in SPEC-T of the embedded finite sentence, it fails to bind an anaphor or to

license an NPI into the adjunct of the upper VP. These contrasts imply that the relevant DP moves in ECM constructions to a position higher than the VP adjunct of the matrix clause.

Condition A satisfaction

- (11) a. The DA proved [two men to have been at the scene of the crime] during each others’s trials (ECM)
- b. [?]*The DA proved [that two men were at the scene of the crime] during each other’s trial (finite embedded sentence)

NPI licensing

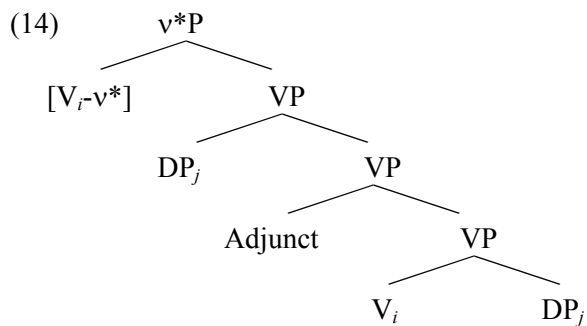
- (12) a. The DA proved [noone to have been at the scene] during any of the trials
- b. [?]*The DA proved [that noone was guilty] during any of the trials

Note that DP objects of transitive verbs, which do not involve raising from the embedded clause to the matrix clause, can also bind an anaphor or license an NPI inside a VP adjunct.

- (13) a. The DA accused two men during each other’s trials
- b. The DA cross-examined none of the witnesses during any of the trials

This phenomenon supports the idea that the verbal phrase contains a position for case licensing. This position would host the object DP of a transitive predicate as well as the DP moving from the non-finite embedded clause, which can be agentive, in an ECM construction. This accounts for the observation that these DPs occur higher than a VP adjunct.

The exact position that objects occupy has not yet been defined. The two possibilities are an extra SPEC-v* and SPEC-V, both higher than the VP adjunct. Since a verb precedes the accusative DP and it cannot be as high as T, the most reasonable account is that V raises to v* and the DP to SPEC-V, in virtue of the inheritance mechanisms previously formulated for the C-T system, which are responsible for offering an A-position for a DP. Thus, V inherits ϕ -features from v* as T inherits ϕ -features from C*.



This analysis is in line with Koizumi's VP-split (Koizumi 1995), according to which the AGR-Obj projection is tucked into two VPs, a higher one, which hosts the external argument, and a lower one, in which internal arguments are base-generated.

The FIT developed by Chomsky (2005b) for ϕ -features can be extended to several Infl-like features, as Chomsky himself points out, since the value of several Infl-like features correlates with the choice of C, as illustrated in chapter 2. This leads to what we may call the Generalized Feature Inheritance Theory:

(15) *Generalized Feature Inheritance Theory (GFIT)*

- (i) C has a central role in the generation of syntactic objects
- (ii) Infl does not have a central role in the generation of syntactic objects
- (iii) Infl-like features are generated on C
- (iv) In the lexicon, Infl lacks Infl-like features
- (v) Infl-like features surface on Infl only derivatively
- (vi) The feature inheritance mechanism has the function of bringing semantic distinctions into a syntactic representation

The GFIT, like the FIT, is a sub-theory of the broader theory of phases: if there are independent reasons to think that C, and not Infl, defines a phase boundary (*i-ii*), as it can be demonstrated that C, and not Infl, has a central role in the generation of syntactic objects (the essence of the theory of phases), and it is observed that Infl-like features are linked to C (they can surface on C and on Infl, their characteristics correlate with the choice of C, and they can trigger V-to-C and V-to-Infl movement), the simplest account for the C-Infl link is that the head argued to have a central role in the generation of syntactic objects is also the locus where Infl-like features are generated (*iii*). If this is the case, two syntactic properties, namely the locality of syntactic operations and the C-Infl link, can be attributed to the phase head, which defines both the syntactic domains to be transferred and the intermediate steps in long movement, as well as the locus where functional features are generated. Thus, Infl-like features are inherent to C and appear only derivatively on Infl heads (*v*), which lack them in the lexicon (*iv*). Following the ideal of the SMT, the inheritance mechanism responsible for spreading Infl-like features from the phase head C to an Infl head should be related to the legibility requirement of bringing semantic distinctions into the syntactic representation (*vi*).

3.3 Revising the Generalized Feature Inheritance Theory

In order to determine whether or not the C-Infl discontinuity is the result of inheritance mechanisms that take place from C to Infl, I shall revise first the status of T/Infl in the GFIT, and second, the status of C in the theory of phases. These are the propositions of the GFIT to be revised:

- (16) (i) C has a central role in the generation of syntactic objects
- (ii) Infl does not have a central role in the generation of syntactic objects
- (iii) Infl-like features are generated on C
- (iv) In the lexicon, Infl lacks Infl-like features

Propositions (16.iii) and (16.iv) are the core of the GFIT, whereas propositions (16.i) and (16.ii) are the core of the theory of phases. In section 3.3.1.1, I shall address proposition (16.iv). In section 3.3.1.2 I shall examine proposition (16.iii), whose plausibility depends on the accuracy of the theory of phases. Consequently, propositions (16.i) and (16.ii) will be investigated in section 3.3.2.

3.3.1 The status of the phase head C and of Infl heads

3.3.1.1 Infl

There are several open questions concerning the lexicon. For example, it is unclear whether the lexicon must contain redundant information which must be independently present in the outer systems (as selectional restrictions or categorial specification for substantive categories), or if, instead, it must be a unique or distributed component (see Bonet 1991, Hale & Marantz 1993 and Marantz 1997, among others). However, a very common view –which seems to be the minimal assumption–, is that the lexicon must contain at least what cannot be derived from principles of the syntactic component or interface conditions, i.e., idiosyncrasies, as argued by Chomsky (1995).

At first glance, this line of argumentation casts some doubt on the legitimacy of postulating a category like T, or more generally Infl, since it would not satisfy the minimal requirement of providing genuine instructions: the features Infl shows up with are not carried from the lexicon, but are inherited from the phase head C during syntactic derivations. The semantic instructions that appear on Infl are attributable to the existence of a device of the syntactic component, the inheritance mechanisms, and not to the alleged lexical item Infl. Thus, if the features that appear on Infl do not come from the lexicon but are due to the syntactic mechanisms of inheritance, it can be concluded that Infl is not a proper lexical item.

One might argue that Infl is a legitimate category in the lexicon in ‘a vacuous sense’: Infl is not originated with semantic instructions, but serves to inherit features and bring semantic distinctions into syntactic representations. Thus, it would be a content category. But all this is just an unnecessary complicated way of saying that inflectional categories are syntactic elements, not symbols of the lexicon.

Even if Infl heads are not originated with Infl-like features such as ϕ -features, tense, mood, modality or negation, a more careful inspection reveals that they provide a useful instruction for the C-I system. As has argued repeatedly in this study, C can be lexically defined as

[+clause typing] and bearing the [\pm embedded] distinction, whereas Infl can be lexically defined as [-clause typing] and lacking the [\pm embedded] instruction. Consequently, Infl heads lack Infl-like features in the lexicon (16.iv), but provide inherent semantic instructions different from those provided by C. Therefore, the GFIT should be slightly modified, reformulating proposition (16.iv) as (16'.iv).

(16') (iv) In the lexicon, Infl lacks Infl-like features, but it is defined as [-clause typing]

This modification has a certain propaedeutic value, for it helps to understand why inheritance mechanisms should apply not only to φ -features, but to Infl-like features in general. Recall that, in the case of φ -features, Chomsky (2005b) argues that feature inheritance is related to the semantic requirement of distinguishing between a position suitable for an operator to bind a variable (an A'-position) and a position suitable for a DP to instantiate φ -features (an A-position). In the general case of Infl-like features, an Infl-like feature should be inherited from C to Infl in order to associate the relevant Infl-like feature with a [-clause typing] position unspecified for the [\pm embedded] distinction. Keeping to the SMT, the mechanisms of feature inheritance are a solution to the C-I requirement by providing Infl-like features with a [-clause typing] position.

3.3.1.2 C

The empirical basis of the theory of phases is that two alleged properties of the syntactic component, cyclicity and successive cyclicity, are relativized to C and v, the so-called phase heads. Accordingly, C, and not Infl, will play a central role in the derivation of syntactic objects if the former, and not the latter, is responsible for defining the domains to be mapped from the syntactic component to the outer systems. In section 3.3.2 I shall discuss the general motivation for postulating syntactic cycles or phases in the theory of the syntactic component, and more particularly, on the motivation for claiming that C, and not Infl, defines syntactic cycles or phases, as the Phase-Impenetrability Condition states.

3.3.2 Revising the Phase-Impenetrability Condition

3.3.2.1 Cyclicity

The intuition under the property of cyclicity is that syntactic relations hold first between the most deeply nested element and its domain, then between the next least nested element and its domain, and so on. This derivational constraint is introduced into syntactic theory by Chomsky (1965) through the Transformational cycle:

(17) *Transformational Cycle*

In a derivation, for all syntactic domains α in a phrase marker, a linear sequence of transformations applies to a domain α_i before applying to α_j , where α_j contains α_i .

Accordingly, derivations should proceed in a bottom-up direction, from the most deeply nested domain to the least deeply nested one. Transformational rules cannot apply to a domain α_j unless they have applied to a domain α_i embedded under α_j in a previous derivational step.

Note that the Transformational Cycle does not predict if a domain or cycle can be tampered once it has been passed. Suppose that in derivation (18) probe γ has extracted a constituent ε_1 from its domain δ yielding the derivational stage Σ_1 , and that probe β has subsequently extracted ε_1 from its domain ζ mapping Σ_1 onto Σ_2 . When the derivation is at Σ_2 , the domain δ has already been passed. The question is, thus, whether the derivation can go back to Σ_1 so that the probe γ extracts the element ε_2 from δ and creates Σ_3 .

$$(18) \quad \begin{aligned} \Sigma_1 &= [\varepsilon_1 [\gamma, [\delta \ \varepsilon_1, \ \varepsilon_2]]] \\ \Sigma_2 &= [\varepsilon_1 [\beta [\zeta \ \varepsilon_1 [\gamma, [\delta \ \varepsilon_1, \ \varepsilon_2]]]]] \\ \Sigma_3 &= [\varepsilon_1 [\beta [\zeta \ \varepsilon_1, \ \varepsilon_2 [\gamma, [\delta \ \varepsilon_1, \ \varepsilon_2]]]]] \end{aligned}$$

Such a derivation is prohibited in Chomsky (1973: 243) under the Strict Cycle Condition, which sharpens the Transformational Cycle.

(19) *Strict Cycle Condition (SCC)*

No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node

The important empirical question was (and still is) to determine which nodes define a cycle. First, C (or S', in previous models) was responsible for defining the syntactic cycles, but for several reasons it was soon suggested that N (see, for instance, Akmajian 1975) or P (Riemsdijk 1978a, b) might also define syntactic cycles. According to the theory of phases, C* but not Infl, and v* but not V, are the cyclic nodes that define its domain as inaccessible once it has been passed. According to the theory of phases, this is not due to the Strict Cycle Condition, but rather to the existence of cyclic mappings from the syntactic component to the external systems, which eliminate a domain or render it invisible to further syntactic computations.

The intuition has been that cyclicity, under any of its formulations, is a rather plausible principle of efficiency that a computational device as the syntactic component should satisfy. Under the more precise version of the Strict Cycle Condition, both the number of possible derivations and the size of the structural span that must be active in the syntactic workspace are minimized, since a syntactic domain is forgotten once it has been passed. It is also worth noting that there are good reasons for thinking of cyclicity as an important principle for phonological computations, for which it was originally postulated. As Lasnik (in press) observes in his revision of the several conceptions of the cycle (see also Kean 1974 and Mascaró 1976 for early discussion on the phonological cycle):

«The cyclic principle, along with analysis into distinctive features, was one of two major syntactic innovations of the mid 1960s borrowed from phonology. The principle was first formulated in Chomsky *et al.* (1956), and applied in the phonological analysis of a variety of languages from the early 1960s on. Chomsky and Halle (1968), the monumental examination of the phonology of English, made crucial use of the cycle in virtually every analysis»

But as Freidin (1978) argued in a very important paper, the empirical support for the existence of syntactic cyclicity can be accounted for through other conditions independently motivated. This issue is more recently reconsidered by Freidin (1999) within a minimalist framework and by Lasnik (in press). In 3.3.1.1 I shall revise Freidin's (1978) observations about the classical empirical argument for strict cyclicity, and in 3.3.1.2 I shall focus on successive cyclicity, another particular facet of cyclicity.

3.3.2.1.1 The classical empirical argument for strict cyclicity

Ill-formed syntactic outputs that adhere to the following derivational scheme (adapted from Freidin 1978: 531) provide empirical support for the Strict Cycle Condition. X and Y are categories and α_1 and α_2 consecutive cycles.

$$(20) \quad \begin{aligned} \Sigma_1 &= [\alpha_1 \dots [X] \dots [Y] \dots] \\ \Sigma_2 &= [\alpha_2 [X] \dots [\alpha_1 \dots [] \dots [Y] \dots] \dots] \\ \Sigma_3 &= [\alpha_2 [X] \dots [\alpha_1 \dots [Y] \dots [] \dots] \dots] \end{aligned}$$

The derivational step mapping Σ_1 onto Σ_2 unproblematically displaces X from a position in α_1 to a position in α_2 . The derivational step that moves from Σ_2 to Σ_3 violates the SCC, since an operation affects only the syntactic subdomain α_1 of α_2 .

Let X and Y be two *wh*-phrases base-generated in the same embedded C*P, and consider the ill-formed sentence (21) and the partially detailed derivation (22), using for the moment traces and indexes instead of copies to keep track of Freidin's arguments.

(21) *What did she wonder where John put?

$$(22) \quad \begin{aligned} \Sigma_1 &= [{}_{v^*P} \text{John put what where}] \\ \Sigma_2 &= [{}_{C^*P} \text{what}_i [{}_{v^*P} \text{John put } [t_i] \text{ where}]] \\ \Sigma_3 &= [{}_{C^*P} \text{what}_i \text{ did she wonder } [{}_{C^*P} t_i \text{ John } [{}_{v^*P} t_i \text{ put where}]]] \\ \Sigma_4 &= [{}_{C^*P} \text{what}_i \text{ did she wonder } [{}_{C^*P} \text{where}_j \text{ John } [{}_{v^*P} t_i \text{ put } t_j]]] \end{aligned}$$

The derivation (22) violates the SCC, which arguably causes the ill-formedness of sentence (21). More concretely, the derivational step that maps Σ_3 onto Σ_4 is acyclic, since the operation of moving 'where' to the embedded SPEC-C* only affects a proper subdomain, the embedded C*P, dominated by the matrix C*P. According to this analysis, C* heads are cyclic nodes. Note that this class of ill-formed sentences cannot be banned by the weaker Transformational Cycle.

Freidin's principal observation is that the class of derivations that violate the SCC can be banned by other principles that have independent empirical motivation. Specifically, two opacity conditions for anaphora binding, the Tensed-S Condition and the Specified Subject Condition, serve this purpose.

(23) *Tensed-S Condition (TSC)*

In a structure

... X ... [α ... Y ...] ...

where α is a tensed sentence, X may not properly bind Y (Freidin 1978: 527)

(24) *Specified Subject Condition (SSC)*

In a structure

... X ... [α ... Y ...] ...

where α contains a subject distinct from Y and not controlled by X , X may not properly bind Y (Freidin 1978: 528)

The independent motivation of these conditions is, respectively, the inability of a subject DP to penetrate an embedded tensed sentence to bind a subject anaphor or an empty category, or to cross over a distinct and not controlled DP to bind an object anaphor or an empty category.

(25) TSC violations

- a. *John_i thought that himself_i was clever
- b. *They_i expected that each other_i would win
- c. *Jill_i was reported that e_i had won a prize

(26) SSC violations

- a. *John_i believed Mary to like himself_i
- b. *We_i expected John to consult each other_i
- c. *Jill_i was reported Jack to have insulted e_i

In the light of these two conditions, consider the contrast given in (27), from Freidin (1978: 529):

- (27) a. [_{C*P} who_i John said [_{C*P} t_i Fred saw t_i]]
 b. [_{C*P} who_j John knew [_{C*P} what_k [t_k frightened t_j]]]

Structure (27.a) is an acceptable syntactic output, whereas structure (27.b) is not. According to Freidin's reasoning, "the binding relation between 'who_j' and its trace [in the ill-formed (27.b)] violates the TSC and the SSC", whereas in the well-formed (27.a) there is a trace t_i in the embedded SPEC-C* that suffices to bind the trace in the base-position.

Freidin extends this line of reasoning to other ill-formed syntactic outputs which were attributable to the SCC and concludes that they all can be accounted for in terms of the TSC and the SSC, or rather, in terms of a more general Opacity Principle, defined as follows.

(28) *Opacity Principle*

In a structure

... [X ... [α ... Y ...] ...

where: a. $\alpha = S'$ (or NP), and

b. Y is not bound to any c-commanding category in α (linking convention)

if Y is in the domain of

i. a finite clause (TSC), or

ii. a subject not controlled by X (SSC)

then X may not properly bind Y (Freidin 1978: 530)

This principle treats S' (and N) not as a cyclic node that plays a crucial role in the definition of the syntactic material available to syntactic operations, but as an opacity inducer that defines its domain as non-transparent to a category c-commanding an anaphora. Such opacity impedes a DP from binding an anaphor or an empty category and a *wh*-phrase from binding its trace in the base-position across a S' .

In sum, “as the SCC has no independent empirical motivation, there is no reason to stipulate it as a part of the theory of grammar (Freidin 1978: 539)”. A theory of grammar that avoided the generation of sentences like (21) would be excessively precise since these syntactic objects can be banned by representational conditions or, in minimalist terms, by legibility conditions. This leads Freidin to argue that derivational restrictions such as the SCC are redundant and that conditions on representations are the decisive factor.

Reconsider the derivational step that maps Σ_3 onto Σ_4 in the derivation (22), repeated in (29).

- (29) $\Sigma_1 = [_{v^*P} \text{John put what where}]$
 $\Sigma_2 = [_{C^*P} \text{what}_i [_{v^*P} \text{John put } t_i \text{ where}]]$
 $\Sigma_3 = [_{C^*P} \text{what}_i \text{ did she wonder } [_{C^*P} t_i \text{ John } [_{v^*P} t_i \text{ put where}]]]$
 $\Sigma_4 = [_{C^*P} \text{what}_i \text{ did she wonder } [_{C^*P} \text{where}_j \text{ John } [_{v^*P} t_i \text{ put } t_j]]]$

It is crucial for Freidin’s analysis that in Σ_4 the *wh*-phrase ‘where’ is relocated in the position where the *wh*-phrase ‘what’ occurred in Σ_2 (embedded SPEC-C*), which results in the erasure of the bound trace t_i . It is precisely the erasure of such an intermediate trace that disallows the *wh*-phrase ‘what’ to bind its trace in the base-position in Σ_4 . Here the assumption is that C*, and any other head category, can have no more than one SPEC. This has been the general view for a long time, and the keystone for the analysis of *wh*-island effects, but indeed, as Chomsky has observed, there is no reason to think that a recursive operation like Merge should be restricted in such a way that it applies only twice to a sole target, which would create syntactic objects as (30.a), but successive applications of different items to a sole target being prohibited by stipulation (30.b).

- (30) a. $[_{HP} Y [_{HP} H X]]$
b. $[_{HP} Z [_{HP} \dots [_{HP} Y [_{HP} H X]]]]$

Nonetheless, since the internal or external applications of Merge are linked to properties of heads, it is expected that, in actual derivations, they become restricted. A further expectation is that languages vary in the number of SPEC positions, i.e., that different heads allow for different numbers of SPECs in different languages. Both expectations seem to be corroborated empirically: firstly, several syntactic phenomena suggest the possibility that multiple features are instantiated in different SPEC positions of a sole head in the relevant languages (Ura 2000), and secondly, several languages display overt fronting of multiple *wh*-phrases to multiple SPEC positions of C*, repeatedly instantiating a sole *wh*-feature (see Richards 1997, 1999, Boskovic 1999).

The second abovementioned case of multiple SPECs is particularly relevant to the study of *wh*-islands: a rather sensible prediction is that languages with multiple *wh*-fronting display weaker *wh*-islands effects or lack them altogether. This seems to be confirmed by some evidence from Bulgarian, a language with *wh*-fronting, provided by Richards (1997: 40).

- (31) a. koja kniga te popita učitelja kogo ubedi Ivan da publikuva?
 which book you asked teacher who convinced Ivan to publish
 ‘Which book did the teacher asked you who Ivan convinced to publish?’
- b. koj izdatel te popita učitelja kakvo ubedi Ivan da publikuva?
 which publisher you asked teacher what convinced Ivan to publish
 ‘Which publisher did the teacher ask you what Ivan convinced to publish?’

These well-formed sentences do not violate Freidin’s Opacity Principle, as the embedded C* in Bulgarian provides enough structural space for the two *wh*-phrases (one SPEC for each *wh*-phrase), which avoids the erasure of a bound trace. Therefore, the ability of a [*wh*] type features to trigger fronting of multiple *wh*-phrases is related to the obviation of the *wh*-island effects.

One might consider the possibility that, in languages like Bulgarian, multiple *wh*-fronting is a syntactic correlate of the semantic operation of absorption (Higginbotham & May 1981). Thus, the [*wh*] type feature in Bulgarian would be specified as [+absorbent] and the [*wh*] type feature of English as [-absorbent], with the caveat that the [\pm absorbent] distinction that refers to the [*wh*] type feature does not reflect the absence or presence of a semantic operation but rather of its syntactic correlate. Note that, apart from the adequacy of this suggestion, it would be a mistake to attribute the absence of *wh*-island effects particularly to the [+absorbent] specification of embedded *wh*-features, since the *wh*-phrase that moves to matrix SPEC-C* does not undergo absorption in the embedded C*P, as it is illocutionarily interpreted in the matrix C*P. This leads to the triggering problem of intermediate copies, which will be discussed in 3.3.2.1.2. For the moment, let us say that the C* head allows multiple SPECs in Bulgarian, but not in English, a conclusion we cannot easily avoid once we keep in mind that the former language, but not the latter, displays multiple *wh*-fronting constructions and lacks *wh*-island effects.

There is a further assumption in Freidin’s revision of the SCC, namely, that a *wh*-phrase₂ moves to the embedded SPEC-C* after a *wh*-phrase₁ has raised from embedded SPEC-C* to

matrix SPEC-C*, resulting in the erasure of an intermediate trace or in a counter-cyclic operation. But recall that the patterns of subextraction from a subject DP recently discovered by Chomsky lead to the conclusion that the π -features that trigger A'-movement to SPEC-C* work in parallel with the φ -features that trigger A-movement to SPEC-T; if these two operations did not take place in parallel, we could not account for the observation that a subject DP is an island for *wh*-subextraction only if the SPEC of a v*P is penetrated. Thus, there is reason to think that two probing relations of the same phase take place during the same derivational step, driven by the phase head, according to Chomsky's implementation. If A'-movement and A-movement take place in parallel at the same phase, it seems reasonable, or even obligatory, to assume that two A'-movement operations at the same phase take place also in parallel. Hence, *wh*-phrase₁ and *wh*-phrase₂ should move in parallel to two different SPEC positions of the embedded C*, just as a subject DP moves to SPEC-T and a π -phrase to SPEC-C* in parallel.

Accordingly, the ill-formedness of English sentences like ‘*what do you wonder where John put?’ cannot be attributed to the possibility that ‘where’ deletes the trace or copy of ‘what’ of the embedded SPEC-C*, but rather to the inability of (embedded) C* to attract more than one *wh*-phrase in English. Accordingly, in English, the syntactic properties of C* are the source of *wh*-islands, or in other words, Σ_3 is ill-formed in derivation (32) because in English C* cannot map Σ_1 onto Σ_2 . Bulgarian C*, arguably, displays the ability to trigger multiple parallel *wh*-movement.

- (32) $\Sigma_1 = [{}_{v^*P} \text{ John put what where}]$
 $\Sigma_2 = [{}_{C^*P} [\text{what}] [\text{where}] [{}_{v^*P} \text{ John put } [\text{what}] [\text{where}]]]$
 $\Sigma_3 = [{}_{C^*P} [\text{what}] \text{ did she wonder } [{}_{C^*P} [\text{what}] [\text{where}] [{}_{v^*P} \text{ John put } [\text{what}] [\text{where}]]]$

Recall that Freidin's observation was that the classical empirical evidence for the SCC, a constraint on derivations, could be accounted for in terms of an Opacity Principle relative to representations, which had independent motivation. Now, it seems that the ill-formed sentences that constitute the classical empirical evidence for the SCC do not involve any derivational step that violates the SCC, due to parallel probing at the phase level. The real source of *wh*-island effects is the inability of C* to provide a SPEC position for each of the two *wh*-phrases mapping Σ_1 onto Σ_2 .

From the preceding considerations we may conclude that the SCC, a particular formulation of the syntactic cycle, has no empirical foundation. The classical evidence for such a derivational principle is not valid because it does not adhere to the hypothetical derivational scheme (20), repeated in (33).

- (33) $\Sigma_1 = [{}_{a1} \dots [X] \dots [Y] \dots]$
 $\Sigma_2 = [{}_{a2} [X] \dots [{}_{a1} \dots [] \dots [Y] \dots] \dots]$
 $\Sigma_3 = [{}_{a2} [X] \dots [{}_{a1} \dots [Y] \dots [] \dots] \dots]$

As probe-goal relations take place in parallel at the phase level, *X* and *Y* would be attracted during the same derivational step.

For our concerns, it is important to stress that there seems to be an asymmetry in the role that C and Infl/T play in the derivation of syntactic objects: C, and not Infl, defines its domain as inaccessible to further probes. Firstly, we have evidence to conclude that the domain of T/Infl is not a cycle or a phase domain, since C* extracts its goals from SPEC-v, which is in the domain of Infl; and secondly, a matrix C* can probe a *wh*-phrase only if it is in a SPEC of the embedded C*P, but not if it is in the domain of the embedded C*P.

The relevant derivational condition that expresses the asymmetry between C and T/Infl is the Phase-Impenetrability Condition:

(34) *Phase-Impenetrability Condition (PIC)*

Consider a Phase $PH = [\alpha, [H, \beta]]$, H being the phase head
Call α and H the edge of PH , and β the domain of H

The domain of H is not accessible to syntactic operations beyond PH , only the edge, $\{\alpha, H\}$, since Transfer sends β to C-I and A-P once H has terminated its work (adapted from Chomsky 2001b: 5-6).

Therefore, the PIC defines an asymmetry between C* and T/Infl in the role they play in the generation of syntactic objects: C*, but not T/Infl, defines its domain as impenetrable, since C is a phase head.

3.3.2.1.2 Successive cyclicity

An important property of internal Merge is that a moving goal reaches its final landing site by means of shorter steps, successively passing through intermediate positions, which work as escape hatches, and leaving a copy in them. This raises the question of whether these intermediate positions for successive movement are provided by the phase head C and not by T/Infl. If that were the case, the categories that defined its domain as inaccessible to further probes would be the categories that provide escape hatches, a natural expectation, since escape hatches allow a category to move towards a further landing site.

The evidence for successive A'-movement is compelling, and will not be revised here. Both internal and external considerations strongly suggest that each C offers an intermediate SPEC position for a *wh*-phrase that undergoes apparently long movement. With respect to internal considerations, recall that a *wh*-phrase cannot move from a CP whose SPEC is filled by another *wh*-phrase in languages that do not allow for multiple *wh*-fronting. With respect to external considerations, intermediate copies are active at SEM, as binding effects reflect (Barss 2001), and they can have a morpho-phonological correlate in several languages (Felser 2004). It is also worth noting that intermediate A'-copies are spelled out in early child English, although the external evidence that children are exposed to does not overtly display intermediate copies (De Villiers et al. 1990, McDaniel et al. 1995). Interestingly, there is reason to think that successive cyclic movement of a *wh*-phrase must leave a copy not only at every CP phase but at every vP phase as well. These copies display binding effects (see Fox

2000 and Legate's 2003 subsequent extensions to non-agentive predicates) and morpho-phonological effects (Cole & Hermon 2000 and Saddy 1991).

There are three types of phenomena that argue that A-movement is comprised of shorter steps: quantifier stranding, subextraction from subjects and binding effects.

In Standard English, the quantifier 'all' can either be pied-piped to matrix SPEC-T (35.a) or stranded in different positions; in (35.a) and (35.b), 'all' is stranded in positions contiguous with embedded non-finite verbs, arguably in SPEC-T_{def}, and in (35.c), it is stranded in a postverbal position, inside the vP. This is proposed by Boeckx (2003), in his elaboration of McCloskey's work on quantifier float in successive A'-movement based on West Ulster English (McCloskey 2000).

- (35) a. All the boys seem to appear to like ice cream
b. The boys seem *all* to appear to like ice cream
c. The boys seem to appear *all* to like ice cream
d. The boys seem to appear to *all* like ice cream

In accordance with Sportiche's (1988) analysis of quantifier float, this pattern suggests that the complete quantificational phrase 'all the boys' moved from the base-position to its final landing site by passing through several SPEC-T_{def} positions, where the quantifier can be stranded.

Recall that, as observed in Chomsky (2005b), subextraction of a *wh*-phrase from an agentive subject is possible in raising constructions.

- (36) a. *it was the CAR (not the TRUCK) of which [the (driver, picture) [t caused a scandal]
b. *of which car did the driver cause a scandal
- (37) a. it is the CAR (not the TRUCK) of which [the (driver, picture) is likely [t to cause a scandal]]
b. of which car is [the (driver, picture) likely [t to [cause a scandal]]]

To account for the observation that the PP 'of which' can be extracted from the external argument 'the (driver, picture) of which' in (37) but not in (36) it is necessary for the probing C* in (37) to detect the PP when the external argument is not in the SPEC-v*. There are two possible positions: embedded SPEC-T_{def} and matrix SPEC-T. There is reason to think that matrix SPEC-T cannot be probed by matrix C*: if it were, the thematic characterization of the subject would not be a relevant factor in determining subextractability patterns where no SPEC-T_{def} is involved. Thus it seems that patterns (37) require that A-movement be successive cyclic.

Standard conceptions of binding principle A also suggest the necessity of successive cyclic A-movement. Consider the following paradigm, noted by Danny Fox, which plays a decisive role in the discussion of the successive cyclic status of A-movement.

- (38) a. John_i seems to Mary to appear to himself_i to be happy
 b. *Mary seems to John_i to appear to himself_i to be happy

An important factor in the analysis of (38) is that the experiencer NP embedded inside the PP has the ability to bind an anaphor not dominated by the PP (see Boeckx 1999 and references cited therein); note that, in (39.a), the NP ‘John’ cannot be correferential to the experiencer ‘him’, and that, in (39.b), the anaphor ‘himself’ is bound by the experiencer ‘John’, a further argument for the necessity of A-reconstruction.

- (39) a. They_i seem to him_j [to like John_{k/*j}]
 b. [Pictures of himself_i] seem to John_i [[pictures of himself_i] to be ugly]

By this reasoning, A-reconstruction in (38) is necessary in order to allow that the experiencer ‘Mary’ does not trigger intervention effects in (38.a) and to disallow that the experiencer ‘John’ in (38.b) binds the anaphor in the embedded sentence.

Therefore, both internal and external considerations suggest the necessity of successive cyclic A-movement. For internal considerations, intermediate positions of A-movement are necessary to account for the observed patterns of subextraction from agentive subjects in raising constructions (36-37), and for external considerations, a part of the intermediate copy is spelled out in the case of quantifier float (35) and the binding effects illustrated in (38-39) require the matrix subject to bind the anaphor from the SPEC-T_{def} position.

This leads us to the conclusion that the PIC is not a necessary condition for successive cyclicity: T_{def}, a category argued not to be a phase head, and whose domain should be accessible to further operations, forces a moving DP to pass through its SPEC.

One could claim that the successive cyclic nature of A'-movement and of A-movement derive from independent sources: the former is a consequence of the PIC, and every intermediate step is triggered at the level of every phase, whereas the latter is a consequence of a condition on chain formation that requires steps to be as local as possible, in a Takahashi/Boeckx-style (Takahashi 1994 and Boeckx 2003): a category *X* undergoes movement when its final landing site is present in the structure, by passing through all intermediate positions of type *X* to minimize the chain links.

But there is no reason to stipulate that successive cyclic A'- and A-movement are caused by two different economy conditions, a uniform account being the preferable one. Optimally, intermediate links are triggered either in a stepwise manner or in Takahashi/Boeckx-style for A'-movement and for A-movement alike. In this sense, note that the subextractability from an agentive subject in raising constructions is unexpected under the Takahashi/Boeckx-approach: if a PP remains inside the DP in the base-position and undergoes movement only when (matrix) C is merged, sentences (36) and sentences (37) should be equally degraded.

The best way to account for successive cyclicity arises from the following two observations (see Rizzi 2002, 2004 for very similar ideas):

- (40) (a) a moving category with a [+token] feature X passes through intermediate positions offered by categories with a [+type] feature X
- (b) the [+type] feature X that offers the intermediate position does not suffice to host the category with a [+token] feature X

Let's take a closer look at observation (a). A phrase with a marked value for π -features (a *wh*-phrase, for instance) passes through SPEC positions of C and v on its way to the SPEC-C position where it is illocutionarily interpreted. The fact that specifically C and v are the categories that offer an escape hatch for successive A'-movement is rather natural when we bear in mind that SPEC-C and extra-SPEC-v are associated with peripheral features. It is uncontroversial that C offers the domain for identifying peripheral features. Recently, Jayaseelan (2001) and Belletti (2002, 2001, 1999) argued that there is an InFLP-internal periphery above the vP, which offers the landing site for topical and focal phrases. This introduces an interesting parallelism between pre-sentential and pre-predicate positions, as both would provide suitable positions for π -phrases and intermediate positions for successive A'-movement. It is also clear that a DP with a token of φ -features uses SPEC-T specifically, the position where φ -features associated with nominative case are prototypically identified, as an escape hatch on its way to the final landing site. Thus, the choice of intermediate positions of successive A'-movement and successive A-movement is determined by the type of the token feature: π -phrases move through SPEC-C and SPEC-v, and nominative DPs through SPEC-T.

Although there is a type agreement between intermediate copies and the hosting head (observation (a)), the [+type] feature does not provide a suitable position for the [+token] category to be interpreted (observation (b)), and thus the category with the [+token] feature must still be remerged to another position. The anaphoric φ -features of T_{def} cannot offer an appropriate position for a DP with a complete set of φ -features; thus, in a raising construction or in an ECM construction, a DP moves from the SPEC of the head T_{def} , whose set of φ -features is anaphoric, to the SPEC of a higher occurrence of φ -features until it arrives at a SPEC- T_{compl} in raising constructions and at a SPEC- V_{compl} in ECM constructions. When the DP is in SPEC- T_{def} it is still active, as it cannot be properly interpreted there. Similarly, a moving *wh*-phrase cannot be properly interpreted in intermediate A'-positions. In (41), for instance, the type feature of force of the embedded C is [+assertive], whereas 'who' provides a [+interrogative] token feature; consequently, the embedded C cannot offer an appropriate position for the *wh*-phrase.

- (41) Who do you think [~~who~~ [that]_{C[+assert]} Peter saw]

Both *wh*-phrases and φ -phrases (or DPs) are 'active' when they are not in a semantically adequate position (SPEC- T_{def} in the case of DPs and SPEC- $C_{[+assert]}$ in the case of *wh*-phrases, respectively) and 'frozen in place' when they are in a semantically adequate position.

Recall that languages vary as to the number of SPECS provided by C, a factor that determines the presence or absence of *wh*-island effects. Bulgarian, a language that allows multiple *wh*-fronting, obviates *wh*-island effects. As noted above, the intermediate copy of a *wh*-phrase in

a SPEC of an embedded interrogative C cannot be analyzed in terms of absorption, since such a *wh*-phrase is not illocutionarily interpreted with respect to the embedded interrogative C but rather with respect to a higher interrogative C. This situation seems to suggest the necessity for the embedded interrogative Cs to contain a force feature specified as [+*wh*], which is instantiated by the *wh*-phrase that is spelled out and illocutionarily interpreted in embedded SPEC-C, as well as an extra-peripheral feature, which does not suffice to provide the semantically appropriate position for a *wh*-phrase illocutionarily related to a higher C. In Bulgarian, but not in English, an embedded C can have both an interrogative force feature and an extra-peripheral feature that triggers intermediate steps of A'-movement.

In sum, successive cyclicity is not an effect of the PIC, since both C/v and T provide intermediate positions for A'- and A-movement respectively, but rather a consequence of the presence of non-interrogative π -features in C and v that fail to provide an appropriate interpretive positions for a *wh*-phrase and of anaphoric φ -features of T that fail to provide an appropriate interpretive position for a DP with a complete set of φ -features. *Wh*-island effects emerge when a C specified as [+*wh*] fails to contain an extra π -feature, as in English embedded questions.

This line of reasoning forces us to revise the conclusion we arrived at in 3.3.2.1.1: there is an asymmetry between C and T in the role these categories play in the generation of syntactic objects, since C, but not T, defines its domain as inaccessible. Recall that the premise for such a conclusion was that the domain of T is accessible to C, whereas a superordinate C could attract a *wh*-phrase only if such a *wh*-phrase was at the edge of the immediately subordinate CP, that is, if A'-movement proceeded successive-cyclically. According to the arguments given above, successive cyclicity is attributed to the presence of defective features, and it cannot be attributed to the PIC: anaphoric φ -features in successive A-movement and non-interrogative π -features in successive A'-movement. This accounts for the existence of successive cyclicity, but does not account for the idea that a [*wh*] type feature fails to penetrate the domain of an embedded CP to probe a *wh*-phrase, as in English. However, the aforementioned asymmetry between the so-called phase heads C/v and Infl-heads is quite a natural one: if, for instance, a π -feature located on v penetrated the domain of an immediately lower C to probe a *wh*-phrase, it would penetrate the domain of a head containing π -features, but if a π -feature located on C penetrated the domain of an immediately lower Infl head to probe a *wh*-phrase at the edge of vP, then the penetrated domain would not be the domain of a head containing π -feature. In other words, the asymmetry between C/v and Infl-heads does not derive from a special status of the so-called phase heads in the generation of syntactic objects, but from a relativized opacity factor relative to probe-goal relation, which may be defined as follows:

Relativized Opacity Principle (ROP) (non-final definition)

In a syntactic object [α_1 [... [α_2 ... γ ...]]],

where: (i) α_1 and α_2 are two probes of the same type α ,

(ii) γ is in the domain of α_2 , and

(iii) γ is a goal with a token α ,

α_1 cannot readily probe γ

When *v* probes a *wh*-phrase that belongs to the domain of a lower *C*, it violates the ROP, but when *C* probes a *wh*-phrase into the domain of *Infl*, it does not.

It can thus be concluded that no deep asymmetry exists between *C* and *Infl*: successive cyclicity derives from a type-agreement relation between a type feature and a token feature that does not yield a satisfactory matching relation that provides an appropriate instance for the type feature and a semantically suitable position for the moving category, which remains accessible to further probes, and the observation that *C* can penetrate the domain of *Infl* but *C* cannot penetrate the domain of a lower *v/C* derives from the ROP. Consequently, neither successive cyclicity nor penetrability restrictions provide empirical motivation for the PIC.

It is clear that the ROP formulated above is not incompatible at all with the existence of cyclic mappings to the external systems, but it does not require them. Although the PIC or the existence of cyclic mappings to SEM and PHON seems unnecessary to account for impenetrability, it remains unresolved whether cyclic mappings to SEM and PHON are necessary for external requirements and whether they are defined by *C/v*, and not by *Infl/V*. Indeed, the possibility that cyclic mappings are relevant to minimize the computations that take place in the external systems, but not necessarily in narrow syntax, is not very far from Chomsky's position. Consider these two quotations from two recent papers:

«PIC sharply restricts search and memory for ϕ [the phonological component], and thus plausibly falls within the range of principled explanation [...]. It could be that PIC extends to NS [Narrow Syntax] as well, restricting search in computation to the next lower phase (Chomsky 2001b: 6)».

«Note that for narrow syntax, probe into an earlier phase will almost always be blocked by intervention effects. One illustration is agreement into a lower phase without intervention in experiencer constructions in which the subject is raised (voiding the intervention effect) and agreement holds with the nominative object of the lower phase (Icelandic). It may be, then, that PIC holds only for the mappings to the interface, with the effects for narrow syntax automatic (Chomsky 2005b: 9-10)».

There is something odd in this position. Firstly, the PIC, as defined, is a condition that reduces the search domain of a phase head: the edge of a lower phase is accessible, but its domain is not; prototypically, probe-goal relations take place at narrow syntax. Secondly, in the absence of an intervening category, a probe can detect a goal located in the domain of a lower phase, which argues that the relevant constraint on probe-goal relations is not the PIC but rather a minimality principle; this seems to be the case in experiencer constructions of several languages like Icelandic or Catalan, in which *V* agrees with a nominative object.

(42) Li agraden les pel·lícules
 3SG.DAT like.3PL the films
 'He likes films'

And thirdly, the PIC is argued to hold "only for the mappings to the interface", "restricting search and memory for ϕ ". What seems strange here is that the effects of the PIC, a condition

that restricts the probe-goal relations that take place at narrow syntax, are claimed to “hold only for the mappings to the interface”, restrictions on probe-goal relations being derived from other conditions as minimality.

In sum, the PIC is neither a necessary condition in order to account for successive cyclicity (there exists both A'- and A-movement are successive) nor a sufficient condition in order to account for impenetrability (when there is no intervener, the domain of a phase can be accessed by a higher probe).

3.4 Discontinuities

The plausibility argument we address in this section is as follows: if it can be motivated on independent grounds that “C and v [but not Infl] are the labels driving internal operations and relevant to external Merge, and are also the points of feature-valuation and transfer (Chomsky 2005a)”, then the C-Infl link might derive from inheritance mechanisms. This raises the possibility that the labels defining the domains accessible to syntactic computations offer the locus where functional features are base-generated, and that more elaborate structures revealed by the cartographic inquiries are based on linearization of features in these labels, and possibly labels closely linked to them (as in the C-T connection)”, as Chomsky (2005a: 18) speculates.

In order to decide if internal operations are driven by the phase head and if the C-Infl connection is due to inheritance mechanisms, it is crucial to find out whether there is an asymmetry between C and Infl with regard to the role they play in the generation of syntactic objects. This question has been examined in the context of derivational restrictions for extraction; since, firstly, there is no empirical reason to think that C defines strict cycles and, secondly, both C and Infl are relevant in the determination of successive cycles, it has been concluded that there is no such asymmetry. Consequently, the PIC (which assumes an asymmetry between C and Infl) has been replaced by the ROP.

It should not be forgotten that, in the theory of phases, the categories that define the fragments of structure to be transferred (and thus that become invisible) are not stipulated or defined on the basis of impenetrability effects, but rather they are independently determined on the basis of both internal and external considerations. The way phases are defined varies from the way barriers were defined.

The main internal consideration relevant to the idea that C* and v* are the phase heads is that Infl and V display a complete set of φ -features that can host a DP with a full set of φ -features if and only if C* and v* are present. The analysis of this observation states that φ -features are ‘inherent’ to C*/v* and ‘derivative’ to Infl/V. However, this is not a necessary conclusion; it might be the case, for instance, that C and Infl or v and V simply had to ‘agree’. The virtue of inheritance mechanisms is dependent on whether it can be independently argued that there is an asymmetry between C*/v* and Infl/V with respect to the role they play in the derivation of syntactic objects. For the same reason, the condition that probe-goal relations take place in parallel in the C-Infl system does not necessarily imply that C is the label driving both A'-movement and A-movement.

The external A-P considerations of which I am aware (see Chomsky 2000) do not provide a very strong argument, since “the points of feature-valuation” do not coincide with the fragments that display distributional freedom: whereas an embedded clause of a raising verb cannot be dislocated (43), a non-agentive vP can be (44). The observation that control structures can be clefted or dislocated is provided by Rizzi (1982a).

- (43) a. To stay at home I think he prefers
b. *To stay at home I think he seems
- (44) a. John wants to read the book, and read the book he will
b. John wants to stay at home, and stay at home he will

In fact external A-P considerations conflict with internal considerations, since it is not clear at all whether clauses embedded under control predicates display a complete set of φ -features: although they can be dislocated as finite clauses, their set of φ -features is anaphoric, like that of raising and ECM constructions (see above 2.1.1 and also 2.1.2). Thus, external A-P considerations are not compelling.

As to external C-I considerations, it might well be that C, and not Infl, is the locus of force features, and that v*, and not V, is the locus of agentivity; thus, in some sense, C and v* are responsible for closing the proposition and the predicate, respectively, and thus constitute objects with a relatively semantic autonomy. But again, this does not necessarily imply any significant asymmetry between C/v and Infl/V, as the semantic instructions provided by Infl-like features are as necessary as those provided by C/v in defining the properties of the proposition or the predicate.

These reasons lead me to work out an alternative to the GFIT in order to account for the C-Infl link, materialized in the following three properties:

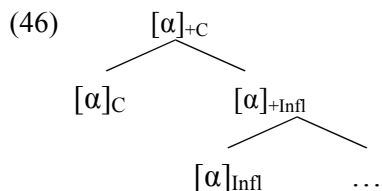
- (I) Complementizers can ‘replicate’ Infl-like features
- (II) There is a correlation between the characteristics of features surfacing on Infl and the choice of C
- (III) Infl-like features are involved in triggering V-to-C movement

Let us use the observation made in section 3.3.1.1 as a starting point. There we saw that the lexico-semantic feature with which C heads and Infl heads are originated is the same, [\pm clause typing], but with a different polarity: C is specified for a positive polarity ([+clause typing]) and Infl for a negative polarity ([-clause typing]). An issue that seems to me undecidable is whether such a distinction is inherited from the lexicon (in which case Infl and C would be lexical primitives) or whether it is a syntactic product that derives from splitting a hypothetical lexical unit [\pm clause typing] into two positions, one for each polarity. Be as it may, what is crucial for our concerns is that C and Infl turn out to be two different positions

with a different polarity of the same feature, and in that sense a discontinuous featural pattern.

(45) [+clause typing]_C ... [-clause typing]_{Infl}

Observe that Infl-like features introduce semantic instructions that are orthogonal to the [\pm clause typing] distinction; for instance the instruction of being 3rd person singular referred to an individual or of being negative referred to a proposition is independent of either the [-clause typing] instruction provided by Infl nodes or the [+clause typing] instruction provided by C nodes. This orthogonality with respect to both values of the [\pm clause typing] feature may force Infl-like features, when they are to be introduced into the discontinuous template ‘[+clause typing]_C ... [-clause typing]_{Infl}’, to be merged in both poles, and then to split into a [+clause typing] occurrence and a [-clause typing] occurrence. Thus, an Infl-like feature α is associated with two positions of a discontinuous syntactic template, and is therefore a discontinuous feature in syntax.



This would straightforwardly account for properties (I) and (II): the agreement relation observed between C and Infl is no more than a manifestation of inserting Infl-like features in both poles of the ‘[+clause typing] ... [-clause typing]’ discontinuity. Naturally, languages vary with respect to the morphological materialization given to such a deep link; as described in chapter 2, certain languages display a relatively rich inflectional paradigm of φ -features on C, while others only a [\pm fin] distinction, certain languages mark subjunctive distinctions on C, while others use verbal endings and others free particles, certain languages have negative markers encoding the [\pm embedded] distinction, while others do not, and so on. All this falls under the indisputable degree of morpholexical cross-linguistic variation.

The idea that the C-Infl relationship should be understood as a discontinuous unit is not new in the generative literature; as far as I know, at least Stowell (1981) as well as Rochette (1988) and Drubig (2001) have resorted precisely to this possibility at some point within their argumentations, Stowell to account for certain properties of tense and Rochette and Drubig to account for certain properties of modality. In Stowell’s words (1981: 241):

«we might adopt a suggestion of Y. Aoun (personal communication) to the effect that the complementizer and Infl form a discontinuous element. The matching between complementizers and Infl would then follow from the fact that the two actually form a single unit at some level, so that selection for one implies selection for the other».

My proposal is that this intuition, thus far not seriously considered nor systematically researched, may turn out to be the appropriate analysis of the connection between complementizers and inflectional heads.

Property (III), or the observation that Infl-like features can trigger V-movement both to C and to Infl, can be argued to be just a side-effect of properties (I) and (II). If Infl-like features are inserted both in C and in Infl, it is expected that they can work as probes from one position or the other. This raises the very complex question of which factors determine which of the two occurrences of Infl-like features is active in which language and in which construction. This issue will be partially explored in the next chapter.

First we shall tentatively explore the possibility of the existence of other discontinuous syntactic patterns beside C-Infl. In this case, the possibility for a feature to be associated during syntactic computations with two positions is not an idiosyncrasy of Infl-like features but a more general property of the syntax of functional features. I shall focus on two possible cases: the connection between prepositions and the case endings of nominals and the connection between the functional head *v* and the V. Please mind the tentative character of the following inquiries.

It is also a traditional observation that P ‘assigns’ case to a nominal, or, in other words, that the value of K (case) correlates with the choice of P as the value of several Infl-like features correlates with the choice of C. As a matter of fact, the problem of precisely describing the correlation between the mood endings and the choice of complementizer in a language like classic Latin is analogous to that of precisely describing the correlation between the case endings of nominals and the choice of preposition. This close relationship suggests that P and K constitute a sole unit split into two positions, just as C and Infl do.

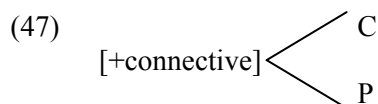
Again languages vary in how they materialize the P-K link with morpholexical material, resorting only to prepositions, only to nominal endings (which are often grammaticalized postpositions in the case of oblique cases), to both or to none of them.

The functional similarity between C and P is quite clear. As Rizzi (1997: 283) observes, “We can think of the complementizer system as the interface between a propositional content (expressed by the IP) and the superordinate structure (a higher clause, or possibly, the articulation of discourse, if we consider a root clause)”; analogously, we might add that the prepositional system can be thought of as an interface between an argumental content (expressed by the DP), and an embedding structure, a higher VP or DP, or perhaps discourse if we consider nominal fragments.

As Emonds (1985) argued in a work with the very explicit title ‘S as P and COMP as P’, it may be that P and C are not two distinct categories but rather instantiations of the same archicategory appearing in different syntactic frames. Emonds’ particular position was that complementizers are a subcategory of prepositions appearing in the frame —S (where S is equivalent to Infl), and thus that a CP is a subtype of PP (in Emonds’ terms, S would be a subtype of P). Although I sympathize with Emonds’ methodological intuition that it is desirable to reduce the number of primitive grammatical categories of UG, “if empirically supported”, I must adopt a different technical implementation. The principal reason for this is that, in the model we are developing –but not in Emonds’–, there is no such a thing as a frame —S that defines P as C, for Infl is neither an independent element of C nor an element

pre-existing C, as C and Infl constitute a sole unit spread into two positions. In other words, claiming that P becomes C in the frame —Infl would be a circular argument as the frame itself constitutes a unit with C.

Since the subtypes C and P cannot emerge contextually as in the course developed by Emonds, I propose that both C and P are instances of a more abstract category that cannot be directly identified with C or P. If we recall that both C and P connect an internal structure (IP, DP) to an external one, it seems reasonable to define such an archicategory with a [+connective] feature.



It is as desirable for the theory of UG to find an accurate characterization for the common substrate of C and P as to define the source of divergence between them. The simplest way to achieve this second objective, keeping to the idea that the C-Infl discontinuity is defined by the [\pm clause typing] feature, is to define a distinctive feature for the P-K discontinuity.

The semantic instruction provided by prepositions seems to be a topological relation (direction, origin, location, etc.) between arguments. The choice of P correlates with the value of K in a relatively systematic way; for example, in Latin, where both P and K can be overt, a preposition like ‘ad’, which conveys a relation of direction, correlates with the accusative ending [-em] in ‘ad civitatem’ (‘to/toward the city’), and a preposition like ‘e’/‘ex’, which conveys a relation of origin, correlates with ablative ending [-e] in ‘ex civitate’. Therefore C can be characterized as a [+clause typing] connector and P as a [+topological typing] connector, and Infl and K as [-clause typing] and [-topological typing], respectively.

- C: [+clause typing]
- Infl: [-clause typing]
- P: [+topological typing]
- K: [-topological typing]

As observed by Kratzer (1996), there must be a connection between the thematic role of the external argument and the Aktionsart properties of V, which are argued to originate “from selectional restrictions for the event argument” (“actions, states, events proper and so on”). Kratzer attributes the link between the external argument and the event argument to the functional head voice (or v), which is responsible for conjoining the external argument and the VP. Thus, strictly speaking, the external argument is not selected by V but chained to it by virtue of the functional head voice. Consider the following informal sketch from Kratzer (1996: 123):

«Suppose there are two kinds of voice heads in English: active and non-active. Active voice heads add external arguments and assign (check) accusative Case. Non-active voice heads do not add external argument and do not assign (check) accusative case. [...]. Suppose furthermore that the repertoire of (basic) active voice heads is very limited. Maybe there are

just two of them, one adding an agent argument to an active verb, and the other one adding the holder to the stative verb».

Like C and P, *v* connects an internal structure (a VP) to an external one (an external argument) ensuring that their semantic properties are compatible. This raises the question of whether the node *v*, which selects the external argument (\emptyset , agent or holder, following Kratzer's approximation), and the node V (non-active, active agentive, active stative) constitute a discontinuous template, which would straightforwardly account for the one-to-one relation that exists between the choice of *v* and the properties of V: a particular definition of the properties of the external argument automatically imply a particular definition of the Aktionsart properties of V simply because, at some level, they are the same unit.

The head *v* may still connect the internal properties of V to another kind of external element, manner adverbs, an observation that goes back to Chomsky (1957, 1965) and Lees (1960): "The Verbs that do not take Manner Adverbials freely Lees has called 'middle verbs' (Lees 1960: 8), and he has also observed that these are, characteristically, the Verbs with following NP's that do not undergo the passive transformation (Chomsky 1965: 103)". Therefore, if $v-V = [+middle]$, then *v* can neither become passive nor conjoin a manner adverb with V.

- (48) a. *John married Mary well
a'. *Mary was married by John
b. *John resembled Mary well
b'. *Mary was resembled by John
c. *John had a car well
c'. *A car was had by John

Cinque (1999a: 101-103) resorts to this observation and some cross-linguistic evidence to argue that manner adverbs appear precisely in the SPEC position of a head he terms voice.

Following Kratzer's spirit, the two nodes of the *v*-V discontinuous syntactic pattern can be characterized as two polarities of the same feature: *v* is [+voice typing] and V is [-voice typing].

In sum, after revising some aspects of the theory of phases, it has been concluded that there is no asymmetry between C and Infl in the generation of syntactic objects. Consequently, the GFIT, a sub-theory of the theory of phases, has been discarded: the C-Infl connection is not due to feature inheritance mechanisms that spread features inherent to C to Infl, but to the discontinuous status of C and Infl. Similarly, the condition that A'-movement and A-movement are triggered in parallel does not imply that C is the label that triggers internal merge operations. It has been pointed out that, beside the C-Infl connection, the P-K connection and the *v*-V connection might also be analyzed as discontinuous syntactic patterns.

3.5 Subextraction: (II): relativized opacity for probe-goal relations

Before closing this chapter, I must turn to issue related to the theory of phases that has not been satisfactorily analyzed in our previous discussion.

A very important observation from Chomsky (2005b) is that a π -probe can subextract a π -phrase from the SPEC of a lower head if such a head is T_{def} but not v or C . Putting some potential empirical complications aside (see Uriagereka 2004, Gallego 2006 and section 5.3.2.1 of this study), Chomsky's refinement of Huang's subject islands seems to be solid. However, it still remains to be explained why subextraction from SPEC-C or SPEC- v is problematic; in Chomsky's words (2005b: 14):

«It remains to explain why the probe for *wh*-movement cannot readily access the *wh*-phrase within the external argument of α [α = the v^*P phase]. That could reduce to a locality condition: *which* in α is embedded in the lower phase, which has already been passed in the derivation. We know that the external argument itself can be accessed in the next higher phase, but there is a cost to extracting something embedded in it».

This possibility is unclear: to be precise, the lower phase has not been transferred, but only its domain; consequently, according to the PIC, we would expectt its edge and the constituents of its edge to be accessible to a further phase head.

On this point I shall attempt to argue that the asymmetry attested between subextracting a π -phrase from SPEC- v^*/C^* and from SPEC- $T_{(\text{def})}$ can be derived from a redefinition of the Relativized Opacity Principle. Consider, as a point of departure, the first approach of the ROP given above and repeated here:

Relativized Opacity Principle (ROP) (non-final definition)

In a syntactic object [α_1 [... [α_2 ... γ ...]]],

- where: (i) α_1 and α_2 are two probes of the same type α ,
- (ii) γ is in the domain of α_2 , and
- (iii) γ is a goal with a token α ,

α_1 cannot readily probe γ

As argued, such a definition accounts for the observation that a π -phrase cannot be extracted from the domain of a lower C (*wh*-islands effects), whereas it can be extracted from the domain of a lower T (i.e., from a SPEC- v), since C , but not Infl , contains π -features, and thus induces relativized opacity effects for A' -movement. In other words, a π -probe cannot penetrate into the domain of a lower π -probe, but it can detect its SPEC.

However Chomsky's observations reveal that, although a higher probe can detect and extract a lower phrase at a SPEC of a π -projection, it cannot subextract a constituent from a SPEC of a π -projection: a C head can subextract a π -phrase from SPEC- $T_{(\text{def})}$, but not from SPEC- C/v . This leads to the conclusion that, for a higher probe, there is no distinction between penetrating into a complement or into a SPEC: a *wh*-probe can neither penetrate into the

complement of a lower π -head (*wh*-island effects) nor into the SPEC of a lower π -head (subextraction patterns), but it can penetrate into the complement of T (extraction of a π -phrase from SPEC-*v* to SPEC-C of the same clause) or into SPEC-T_{def} (subextraction of a π -phrase from an agentive subject of a predicate embedded into a raising or ECM construction).

Recall that φ -features and π -features must probe in parallel if they belong to the same discontinuous pattern (or to the same phase, in Chomsky's terms); accordingly, the factor that prevents C from subextracting a π -phrase from SPEC-T when C and T constitute a discontinuous pattern is not the ROP (T is not a π -probe), but rather a condition that requires operations to take place in parallel at the same discontinuous pattern (or phase). In the case of raising/ECM constructions, it is a superordinate π -head that penetrates embedded SPEC-T_{def}, and thus C and T_{def} do not constitute a discontinuous domain and do not probe in parallel. Note that it is crucial that the embedded clause lacks a CP projection (as Chomsky assumes) or that its C head is defective: it provides the [+embedded] semantic instruction (as proposed above), but lacks π -features. If it had π -features, we would expect subextraction from an agentive subject in raising/ECM constructions to be impossible. It is also important to note that a contradiction occurs when two criteria for defining phases are applied to raising constructions: on the one hand, T_{def}P should not be a phase because the φ -features and tense features of T_{def} are anaphoric, but on the other hand, T_{def}P should be a phase because T_{def} must probe its goal before matrix C, and not in parallel. This suggests that the syntactic unit where probe-goal relations take place in parallel is a syntactic discontinuity, which also includes embedded clauses of raising constructions, C_{def}-T_{def}.

According to the above reasoning, penetrating into a SPEC of a π -projection is as troublesome as penetrating into the complement of a π -projection for a higher π -probe. To keep the parallelism as close as possible, extracting the whole complement of a π -projection should be as easy as extracting the whole SPEC. This means that moving an InflP and a VP, with C and *v* stranding, should be as easy as extracting a *wh*-phrase from SPEC-C/*v*. Although this seems to be false, it does not necessarily challenge the proposed analysis, since the static nature of InflP and VP may be due to other factors; for instance, it could be that InflP and VP did not manifest, contrary to arguments or CPs, a suitable token to instantiate a type π -feature (i.e., they would not be born with a *wh*-feature or a topic/focus-feature), or that syntactic operations had problems in affecting only one subelement of a discontinuous element. I shall continue to assume that the whole InflP and VP are visible to a further probe, as SPEC-C and SPEC-*v* are, its apparent unextractability being related to other factors.

Plainly, the ROP, as defined above, cannot account for subextraction patterns, but a minimal modification would automatically account not only for *wh*-islands effects and for the local nature of movement, but also for the asymmetry between subextracting from SPEC-C/*v* and from SPEC-T_{def}. This slight modification is to consider that the opacity inducer is not the head α but the label of the α -projection.

Relativized Opacity Principle (ROP)

(final definition)

In a syntactic object $[_{\alpha_1 P} [_{\alpha_1} \dots [_{\alpha_2 P} \Delta [_{\alpha_2} \Gamma]]]]$,

- where: (i) α_1 and α_2 are two probes of the same type α each projecting an αP
(ii) Δ is SPEC- α_2 and Γ is α_2 -COMPL, and
(iii) $\exists \delta$: δ is a constituent of Δ and $\exists \gamma$: γ is a constituent of Γ ,

α_1 can probe Δ or Γ if they provide a suitable token for α_1 , but it cannot readily probe γ or δ

According to this definition, the factor that decides whether a feature α_1 of type α can readily probe a goal is not being inside or outside the search domain of α_2 , but their relative depth in the α_2 -projection whose label is of the same type as the searching probe α_1 . $\alpha_2 P$ does not render the complete Δ and Γ opaque to α_1 , but search in them becomes difficult. In other words, α_1 manages to detect Δ and Γ but not δ and γ because Δ and Γ are located on a more superficial layer than δ and γ .

As desired, the definition of the ROP given above allows a π -probe to search into a SPEC and a complement of a non- π -projection. Crucially, v can penetrate its complement VP to probe an object *wh*-phrase or an object DP and it can also penetrate such an object DP to subextract a *wh*-phrase; C can also penetrate into its complement TP to extract a *wh*-phrase which is at SPEC- v , but it cannot penetrate the v -COMPL or the SPEC- v , since vP is a π -projection; and finally, a superordinate C can penetrate the subordinate SPEC- T_{def} without penetrating any π -projection (because there is no embedded CP between the matrix C^* and the embedded SPEC- T_{def} or because there is a defective CP lacking non-interrogative π -features).

It is worth noting that, if the conclusions on locality that this chapter has drawn are on the right track, the principle governing locality is no more than a variation of the first locality principle formulated in syntax, the A-over-A principle (Chomsky 1964a, b). See Ross' (1967) classical revision of the A-over-A principle, which gave raise to a great deal of empirical generalizations and ideas that still keep on feeding syntactic theory.

A-over-A principle

If a phrase X of category A is embedded within a larger phrase ZXW which is also of category A , then no rule applying to the category A applies to X but only to ZXW (Chomsky 1964a: 931).

The question whether the ROP is the best definition of the metrics for internal Merge requires further research.

Chapter 4. Analytic Syntactic Patterns

This chapter considers the possibility that the Conceptual-Intentional system generally favors analytic representations with a one-to-one relation between positions and features. It is argued that several distributional restrictions observed by Cinque (1999) derive from the Prohibition Against Tangled Structures and the Prohibition Against Vacuous Quantification, two instances of a general legibility condition, the Full Interpretation Condition. Accordingly, the ordering of functional categories observed by the cartographic project does not lead to a primitive syntactic component, but rather to the study of the C-I system. Finally, it is observed that discontinuous syntactic patterns are no more than a subtype of analytic syntactic patterns.

4.1 On cartographies

The general observation of the cartographic project is that languages do not seem to differ as to the ordering relations between functional categories they allow. If this is correct, the combinatorial possibilities (roughly, literal precedence and scope) of the multiple functional categories are very limited; indeed, it is argued that, for any three different categories *A*, *B* and *C*, it is always the case that they follow a unique strict linear order. Since allegedly there is no flexibility as to the linearization of functional categories, a very precise map of the functional structure where every category occupies its position can be drawn. Belletti & Rizzi (2002: 123) express the basis of the Cartographic Project very clearly in an interview with Noam Chomsky:

«If it is true that a constitutive characteristic feature of natural languages is to privilege representations with many dedicated positions, each with simple interpretive properties, it becomes important to draw a map as precise and fine-grained as possible of this complex-positional system. This is the rationale behind the so-called cartographic studies, which are pursued intensely in some research centers in Italy and elsewhere».

Hence, the main issue of the cartographic project concerns the relation between functional categories and positions, and the conclusion the cartographic studies yield is that UG favors a strict linear order of functional categories or analytic syntactic representations.

As argued in the preceding chapter, there is a very general distinction between [+clause typing] categories and [-clause typing] categories, and it is the case that the former are always in a higher (or less nested) position than the latter. If we recall that *C* links a proposition (say, *Infl*) to a superordinate structure (see above 3.4), then it becomes entirely natural that *C* is higher than *Infl*. But if the cartographic observations are correct, the [=clause typing] distinction is just the starting point, and both *C* and *Infl* must split into many semantically specialized positions, the ‘*C*-area’ and the ‘*Infl*-area’.

According to Rizzi (1997), the map of the *C*-area is composed of two sub-areas, the Force-Finiteness sub-area and the Topic-Focus sub-area. Rizzi asserts that the Force-Finiteness is “the essential part of the *C* system” that connects the superordinate structure and the propositional content, the *Infl*-area. Complementizers express the clause type or the force of a proposition, whether it be a question, an assertion or an exclamation. This instruction is

There is a crucial issue that must be noted here: if adverbs appear in SPEC positions of several silent heads, then a verbal head may appear between two adverbs (see Cinque's chapter 2, 'A Case for Adverb Phrases in Spec'). Cinque illustrates this phenomenon with sentences like (3), where 'an active past participle can be found preceding, or following, each AdvP in the sequence [...] (p. 45)':

- (3)
- a. Da allora non hanno *rimesso* di solito mica più sempre completamente tutto bene in ordine
 - b. Da allora non hanno di solito *rimesso* mica più sempre completamente tutto bene in ordine
 - c. Da allora non hanno di solito mica *rimesso* più sempre completamente tutto bene in ordine
 - d. Da allora non hanno di solito mica più *rimesso* sempre completamente tutto bene in ordine
 - e. Da allora non hanno di solito mica più sempre *rimesso* completamente tutto bene in ordine
 - f. Da allora non hanno di solito mica più sempre completamente *rimesso* tutto bene in ordine
- 'Since then, they haven't usually not any longer always put everything well in order'

In trying to support such a plausibility argument, Cinque's work on the structure of the Infl-area may have brought one of the most careful, subtle and exhaustive syntactic descriptions ever attempted. Such a descriptive milestone raises some important theoretical questions. For instance, in his chapter 4 ('Some Implications and Residual Remarks'), Cinque poses the following two questions:

- (4)
- (a) does the C-I system require full-fledged structures containing specialized features with a default value when there is no overt category with a marked value?
 - (b) can combinatorial restrictions –and hence, cartographies– be reduced to C-I conditions?

I shall not attempt to make any direct contribution as to how or whether morphologically absent categories are coded (4.a); as far as I can see, such a question cannot be empirically substantiated, and one is forced to give purely conceptual arguments, which too often can be indistinctively constructed either for one choice or the other. Hence I shall not attempt to argue for any of the three conceivable possibilities outlined below.

- (5)
- (a) A functional feature with a [-marked] value is always present in the syntactic representation and projects its own projection (Cinque's view)
 - (b) A functional feature with a [-marked] value is present in the syntactic representation structure, but syncretized with a [+marked] feature in a sole position (Giorgio & Pianessi's 1996, 1997 view)

- (c) Functional features are only present in the structure when they are [+marked], in which case they host morpholexical material (Rizzi's 1997 view on the particular case of the topic/focus field)

But I shall directly face question (4.b). Although this question may be premature, as much work is still needed in order to understand the nature of cartographies, it is clear that cartographies lead us to inquire into the way syntactic representations provide useful instructions to the C-I system, an extremely difficult enterprise.

Let us thus consider whether cartographies belong to the set of unexplained properties of the Faculty of Language, or in other words, to the set of properties specific to the Faculty of Language.

4.2 The nature of ordering restrictions

The general skeleton of the clause is one of the most solid syntactic properties. To the best of my knowledge, the observation that tense is higher than aspect or voice in a mono-clausal structure is undisputed; similarly, it is also unquestioned that aspectual markers are taken to be systematically lower than epistemic markers and mood markers. It is generally accepted that deontic modals are lower than epistemic modals, though some cases where deontic modals scope over epistemic modals have been given (Kratzer 1976 and Cormack & Smith 2002). However, there is still no satisfactory account for why the general skeleton of the clause has the form it seems to have.

If Cinque is correct, each of these categories must split into several specialized projections, again, rigidly ordered; accordingly, there would be a universal projection for proximative/retrospective aspect, which would be universally lower than perfective aspect and higher than durative aspect, and a universal evaluative mood, which would be universally lower than speech act mood and higher than evidential mood, and so forth. Languages would vary only as to instantiating these type features of a mono-clausal Infl-area as free heads, verbal periphrasis, affixes or adverbials occurring in a SPEC position of the relevant head.

Cinque's empirical results are preliminary, and therefore may be (partially) inaccurate; for instance, as Cinque acknowledges, "It is not entirely clear whether 'retrospective' and 'proximative' are two values of one and the same aspect. As seen with 'terminative' and 'continuative', their appearing opposite, and complementary, dimensions might be a consequence of their semantics (and of their contiguity)". The same reason may also apply in epistemic necessity and epistemic possibility and virtually in other cases. It is also true that the above cartography may have to be enlarged to include several specialized projections for aspect and root modality, if Romance functional verbs can be analyzed as heads of a mono-clausal structure, as Cinque (1999b, 2000) argues (but see Solà 2002, where the mono-clausal/bi-clausal debate is revised and arguments for analyzing clitic climbing as restructuring in a reduced bi-clausal structure –without a CP layer in the embedded clause– are given).

As will become clear, it is not a minor task checking whether an apparent ordering restriction is real or not. For this reason, the empirical scope of this endeavor will be relatively narrow: I shall focus on the general skeleton of the Infl-area and some of the more fine-grained ordering restrictions, but I shall ignore the cartographic studies concerning other categorial areas (see for instance Koopman's 1993 and Riemsdijk 1978a, b, 1990, 1996 works on prepositional phrases, Cinque's 1994, 1996, 2005 works on determiner phrases and Corver 1997a,b and Zamparelli's work on adjective phrases). Similarly I shall neglect to discuss several studies concerning the clausal structure (as Zanuttini's 1991, 1997 work on negation, Beghelli & Stowell's 1997 work on nominal quantifiers and Cormack & Smith's 2002 work on modals).

4.2.1 Some allegedly non-primitive order restrictions

Cinque's position on the relation between ordering restrictions and semantic constrains is that, "while the relative order of some such notions may indeed reflect intrinsic logical relations among them, the hierarchy is only indirectly related to such semantic, or logical, properties. For instance, certain possibilities which in terms of logical relative scope would be expected are not found, or are downright impossible, thus suggesting that the hierarchy is a construct of the computational system of language, not completely reducible to other components (Cinque 1999a: 135)". Let's begin by considering those restrictions that are allegedly related to C-I constraints (4.2.1.1 and 4.2.1.2).

4.2.1.1 Epistemic modality and tense

Cinque asserts that "the fact seen in chapters 3 and 4 that epistemic modality is higher than (takes scope over) Past, or Future, Tense (see also Bybee 1985, 119ff) appears to reflect the intrinsic relative scope of the two types of operators. Epistemic modality, as noted, expresses the degree of the speaker's commitment to the truth of the proposition expressed by the sentence. But the truth of the sentence can only be evaluated if the proposition is located in a precise moment of time. Consequently, epistemic modality presupposes an operand which is already tensed (and thus its scope is external) (Cinque 1999a: 135)".

The claim that epistemic modality requires a temporal indication in its operand does not ban a temporal operator from taking scope over a proposition whose epistemic modal force has already been calculated, which would create a recursive structure [TP₂... [MODP... [TP₁...]]]. Consider the statement given in (6), where the plausibility of the proposition "Great Britain avoid- war with Germany" is evaluated with respect to a set of premises that some presupposed epistemic agents shared in 1937, a past time. Such a proposition was then evaluated as compatible with the relevant premises, although the evidence was not conclusive, and thus an existential epistemic modal auxiliary is selected instead of a universal one. Note that here we are concerned only with the epistemic interpretation of 'could', and not with the deontic interpretation.

- (6) The general opinion in 1937 was that Great Britain could/might avoid war with Germany, but in 1939 nobody believed that

Thus, (6) presents two tenses, one in the matrix, which scopes over the epistemic modal operator ('could'/'might'), and one in the embedded clause, within the scope of the epistemic modal operator. The former locates the time where the epistemic agents evaluate the proposition, a past tense specified by a temporal PP ('in 1937'), and the latter locates the proposition that is evaluated in a time prior to the speech time but posterior to the matrix tense. Note that the time of evaluation can be shifted, for instance, to 1939, where the epistemic agents have access to a new set of premises with which the proposition under evaluation is incompatible.

The formula (Tense₂ (Epist-Mod (Tense₁ (φ)))) does not display any conceptual deviance: a proposition that takes place in a time₁ is evaluated with respect to a time₂ (in (6), time₂ < time₁ < speech time), and not with respect to the utterance time. The relevant issue to decide if the alleged hierarchy of inflectional functional features is hardwired in such a way that tense cannot recursively scope over epistemic modals is whether the scope relation observed in the bi-clausal structure (6) can be expressed in a mono-clausal structure such as (7).

(7) In 1937 Great Britain could/might avoid war with Germany, but not in 1939

Meaning A:

Nowadays it is thought that it was possible that Great Britain avoided war in 1937; but (nowadays) it is thought that it was impossible that Great Britain avoided war in 1939

Meaning B:

In 1937 it was thought that it was possible that Great Britain avoided war in a time t such as $1937 \leq t < \text{utterance time}$; but in 1939 it was thought that it was impossible that Great Britain avoided war in a time t such as $1939 \leq t < \text{utterance time}$

Sentence (7), but obviously not sentence (6), can express the meaning A, where the PP 'in 1937' does not define an outer past tense where the proposition "Great Britain avoid- war with Germany" is epistemically evaluated, but instead, it is associated with the inner tense under the scope of the epistemic operator. In the interpretation A, the evaluation time is, by defect, the utterance time (or rather a present time that includes the utterance time – 'nowadays'–). Meaning A is notably more prominent than meaning B, but nonetheless it is not impossible for meaning B to emerge from the mono-clausal structure (7). Consider the following passage, in which such an interpretation is contextually salient:

«From 1937 to 1939, there were no objections to German expansion: British Prime Minister seemed to be able to do everything to warrant peace and stability and, although France feared German hegemony, it was safer not to undergo any active protest, provided that Maginot line was a secure defense. For these reasons, according to most analysts, {in 1937 Great Britain could avoid war with Germany}/in 1937 it was possible that Great Britain avoided war with Germany}, but peace expectations dimmed in April of 1939, when Britain engaged to keep Poland's security signing an assistance pact with Poland».

The modal ‘could’ of the underlined sentence can be epistemically interpreted under the scope of the outer past tense defined by the PP. Thus, one of the interpretations of the mono-clausal structure (7) can be paraphrased as ‘that Great Britain avoided war with Germany was compatible with what most analysts knew in 1937’, an interpretation equivalent to the bi-clausal structure (6). Note that it is not discussed whether Great Britain had the ability to avoid war in 1937 but not in 1939, but rather whether the possibility that Great Britain avoided war was compatible with what some epistemic agents knew about the political context. Mono-clausal structure (7) is therefore ambiguous, although one of the two possible readings is less prominent unless the context makes it salient, as in the above passage.

The scope relation between nominal quantifiers and epistemic auxiliaries constitutes a similar case: although nominal quantifiers appear more naturally under the scope of epistemic modals, it is not impossible for them to scope over epistemic modals. At least in Catalan, the following reasoning does not introduce any logical incongruence and the nominal quantifier must scope over the epistemic modal (but see Iatridou and von Stechow 2003, who assert that nominal quantifiers cannot scope over epistemic auxiliaries in English).

- (8) (i) Alguns llibres són a can Joan
(ii) Altres a la biblioteca
(iii) No sabem on és cada llibre
(iv) Per tant, pel que sabem, tots els llibres poden ser a can Joan, tot i que no és possible que tots hi siguin

‘(i) Some books are at John’s house; (ii) others are at the library. (iii) We do not know where each book is. (iv) Thus, as far as we know, all the books may be at Peter’s house, although it is not possible that they are all there’

In sum, it remains unclear whether the Infl-area is rigidly defined to avoid tense from scoping over an epistemic auxiliary, a scope relation that seems to be possible although it is not generally prominent. In fact, if tense could not scope over an epistemic modal, Cinque’s semantic argument (“epistemic modality presupposes an operand which is already tensed – and thus its scope is external–”) would be insufficient: it could not avoid an outer tense to scope over a modal operator if such a modal operator scoped over an inner tense.

4.2.1.2 Tense and aspect

The second, and final, ordering restriction that may reflect semantic constraints taken under consideration by Cinque is that tense must appear higher in the hierarchy than grammatical aspect. The semantic reason given for this is that “aspect is more closely related to the predicate (it expresses different ways of viewing the event expressed by the predicate) than tense (which locates the time of the event –whatever its aspectual make up– with respect to the speech time (p. 135)”, a broadly extended standpoint that Cinque attributes to Foley & Van Valin (1984: 209ff). It could be that, informally speaking, the internal structure of an event must be defined before the event is located in time, and hence that an aspectual marker

merged higher than tense could not be associated with any event, and therefore the created object would be semantically vacuous.

But quantificational aspectual adverbs such as ‘already’, ‘still’, ‘always’, ‘usually’ and ‘often’ are expected, on purely conceptual grounds, to be able to scope over epistemic adverbs or auxiliaries, as the bi-clausal structures in (9) suggest.

- (9) a. Ja és el cas que probablement és a casa seva
‘It is already the case that he is probably at home’
- b. Ja és el cas que pot ser a casa seva
‘It is already the case that he can be at home’

The interesting issue here is whether a quantificational aspectual marker can scope over an epistemic marker when they both are at the same Infl-area, a scopal relation that is conceptually acceptable. Although this is not the most prominent reading and sounds quite unnatural to me, I think that it cannot be regarded as impossible. Thus, the Catalan sentence ‘ja hi pot haver vida a Mart’ (“there already can be life on Mars”), an odd way of saying ‘ja és possible que hi hagi vida a Mart’ (“it is already possible that there is life on Mars”), can be a felicitous conclusion to the following reasoning:

- (10) Fins i tot fa només un any, cap científic assenyat hauria pensat que hi podia haver vida a Mart; però ara, gràcies al descobriment tan important que acabem d’esmentar, hem de concloure que {ja hi pot haver vida a Mart/ja és possible que hi hagi vida a Mart}

‘Even only a year ago, no sensible scientist would have thought that there could be life on Mars; but now, thanks to the very important discovery aforementioned, it must be concluded that {there already can be life on Mars/it is already possible that there is life on Mars}

This context eliminates the deontic reading of the modal: it is not discussed whether scientists permit life to exist on Mars, but the compatibility of the proposition ‘There is life in Mars’ with what is known about the conditions of the planet and what is known about the conditions where life may emerge. It also eliminates the interpretation whereby ‘ja’ is under the scope of the epistemic modal: life was already existing (or had already existed) before the discovery. Thus, the reasoning in (10) forces the underlined mono-clausal structure to be interpreted as the bi-clausal structure ‘It is already the case that now –and not a year ago– life may exist on Mars’.

In sum, the two hypothetical ordering restrictions that were attributed to C-I considerations have been revised, and we have concluded that:

- (a) tense can be coerced to scope over epistemic modality, a relation that is not prominent (or, according to Cinque, impossible),

- (b) aspect appears lower than tense because it is more closely related to the event; however, quantificational aspectual adverbs cannot readily scope over tense, contrary to what would be expected on purely conceptual grounds; again, it seems that quantificational aspect can be coerced to scope over tense,
- (c) the situation in which tense could not scope over epistemic modality and quantificational aspectual adverbs could not scope over tense does not seem to reflect purely conceptual restrictions.

4.2.2 Some allegedly primitive order restrictions

Cinque (1999a: 135-136) provides three cases of ordering restrictions that allegedly do not derive from conceptual restrictions, which will be briefly illustrated in this section and considered in detail in section 2.3. Cinque's method for proving that an ordering restriction between two projections of a mono-clausal structure cannot be derived from conceptual constraints involves constructing a bi-clausal structure where the ordering relation forbidden in the mono-clausal structure becomes acceptable and does not introduce any logical incongruity.

- A. Evidential mood, instantiated by affixes, particles or adverbials, is higher than epistemic mood
 - (12) a. Allegedly John will probably give up
b. *Probably John will allegedly give up
 - (13) a. (?)Evidentemente Gianni ha probabilmente lasciato l'albergo
'Evidently G. has probably left the hotel'
b. *Probabilmente Gianni ha evidentemente lasciato l'albergo
'Probably G. has evidently left the hotel'
 - (14) E' probabile che sia evidente que lui e' il colpevole
'It is probable that it is evident that he is the guilty one'
- B. Retrospective/proximative aspect ('soon') are higher than prospective aspect ('almost'/'imminently')
 - (15) a. He will soon almost be there
b. *He will almost soon be here
 - (16) a. He is about to soon be admitted to hospital
- C. Habitual aspect ('di solito') is higher than terminative aspect ('più')
 - (17) a. Dopo le 10, Gianni non beve di solito più niente
'After 10 o'clock, G. drinks usually no longer anything'

b. *Dopo le 10, Gianni non beve più di solito niente

- (18) Gianni ha smesso di andare di solito a trovare suo padre la sera
 ‘G. stopped usually going to visit his father in the evening’

It remains unclear whether the so-called terminative aspect must occur lower than the habitual aspect. As Svenonius (2002: 213-214) has observed, in English, both precedence relations and both scope relations are possible. The static character of the terminative aspect in Italian, and its relatively low position, may be a peculiarity of ‘più’.

- (19) a. After 10, John usually no longer drinks anything
 ‘It is usually the case that John no longer drinks anything’
 b. After 10, John no longer usually drinks anything
 ‘It is usually the case that John no longer drinks anything’

Although Cinque’s arguments for the inherently grammatical status of the Infl-hierarchy are scarce and too weak to arrive at a well-grounded conclusion, Cinque may be correct in claiming that the set of combinations in a mono-clausal structure allowed by UG is smaller than the set of combinations that are legitimate on merely conceptual grounds, a possibility that merits some attention.

4.3 Toward a principled account for some order restrictions

Interestingly, auxiliaries must be combined in a rigid order in a language like English in such a way that a modal (‘should’) or a temporal auxiliary (‘will’) must precede aspectual auxiliaries. In fact, if a perfective periphrasis (‘have + past participle’) and a continuative auxiliary (‘being’) appear in the same sentence, the perfective periphrasis must precede the continuative auxiliary, and they both must precede the lexical verb with a voice mark [-d] (‘produce-d’), which appears in the final position of the clauses in (20). If a complementizer (‘that’) is present, it must precede all aforementioned categories.

- (20) a. [C That (this model) [Mod should [Asp-Perf have been [Asp-Cont being [voice produced]]]]]
 b. [C That (this model) [Fut-T will [Asp-Perf have been [Asp-Cont being [voice produced]]]]]
 [C [Mod [T [Asp-Perf [Asp-Cont [v-V [V]]]]]]]]

These facts do not seem to be an accident of English grammar, since ordering restrictions among grammatical categories (verbal auxiliaries, affixes, free particles or adverbials) do not vary cross-linguistically. As important and pressing as this question is, I shall not attempt to construct a good answer for it. However, I hope to show that there is no reason to think that hierarchies are hardwired in a clumsy fashion.

4.3.1 Voice and tense

Let's begin with the ordering between *v* and tense. If, as argued in section 3.4 of chapter 3, *v* forms a discontinuous element with *V* and Tense (an inflectional head for the Infl-like feature of tense) forms a discontinuous element with *C*, then it is quite natural that *v* cannot occur higher in the structure than Tense; if it did, the two discontinuous syntactic templates would be tangled, and thus the *C*-Infl connection and the *v*-*V* connection would be broken off. Therefore, some functional features would be required to be ordered in such a way that two discontinuous units do not get tangled, a requirement that may be expressed in terms of the following Prohibition Against Tangled Structures.

Prohibition Against Tangled Structures (PATs)

Two discontinuous syntactic templates $[A_1 \dots A_2]$, $[B_1 \dots B_2]$ cannot be broken off by yielding the tangled object $[A_1 \dots [B_1 \dots A_1] \dots B_2]$

It would be more illuminating to define the semantic effects of tangled structures. On the particular case of the ordering between tense and voice, Kratzer (1996) observes that if the external argument is not strictly an argument of *V*, then nothing prevents voice from carrying out temporal or aspectual information, or in fact, from occurring higher than temporal or aspectual projections. Kratzer explores a way of locating the voice projection lower than tense based on the order in which two semantic operations, Event Identification and Existential Quantification, apply. Kratzer proposes that, whereas the operation of Existential Quantification is responsible for providing a truth-value (p. 122), "Event Identification is one of several admissible conjunction operations" responsible for chaining "together various conditions for the event described by a sentence"; more specifically, the addition of the external argument might "proceed via the operation of Event Identification", yielding "the connection between the Aktionsart of a verb and the thematic role of the external argument". Kratzer's operation of Event Identification is a post-syntactic algorithm that checks the compatibility of the properties of two predicates, the external argument and the verb.

Kratzer asserts that voice can appear anywhere in the hierarchy of a verb's inflectional heads "as long as the event argument is not existentially quantified (p. 125)", since the operation of Event Identification cannot apply once a truth value has been obtained. "If we have more than one inflectional head in addition to voice (possibly all of Tense, Agr, Mood and Aspect), then we have to find out which head does the existential quantification (p. 126)". "For the time being, let us tentatively assume that Voice is located directly above the VP, but stay open to the possibility that it may turn out that there are intervening functional heads after all".

Thus, the semantic algorithms of Event Identification and Existential Quantification over events cannot apply in the right order in the tangled structure $[C [v T] V]$: the external argument and *V* can be properly linked at the right stage before the event variable is bound.

An alternative way of grasping Kratzer's intuition could be in terms of a condition on representations that avoided vacuous quantification. Consider the following definition. See

Kratzer (1995) and Chomsky (1982, 1991/1995) on the prohibition against vacuous quantification in natural language.

Prohibition Against Vacuous Quantification (PAVQ)

For every occurrence of a variable x there must be a quantifier Q binding x , and for every quantifier Q there must be a variable x such that Q binds x

If voice and its argument (the so-called external argument) occur higher than tense ($\exists e$), we obtain formula (21.a), whereas if they occur lower than tense, we obtain (21.b).

- (21) Peter bought the house
 a. $\exists e(e: \text{past}) [\text{Agent}(\text{Peter})(e) \ \& \ \text{bought}(\text{the house})(e)]$
 b. $\text{Agent}(\text{Peter})(e) \ \& \ [\exists e (e: \text{past}) [\text{bought}(\text{the house})(e)]]$

In (21.a) the external argument and the verb are unproblematically chained before tense has provided a truth-value and before all occurrences of the eventive variable are bound by the existential quantifier; but in (21.b) Event Identification fails to link the properties of the event denoted by the external argument and the verb (because a truth-value has been obtained via Existential Quantification), and the occurrence of the event variable of the external argument ($\text{Agent}(\text{Peter})(e)$) is free. Thus, the PAVQ is violated if the voice projection (the voice head and its argument) occurs higher than the tense projection.

The PAVQ may be useful in understanding why some combinations that do not seem to involve tangled structures and that seem to be legitimate on pure conceptual grounds are not allowed. The general expectation is that a conceptually possible scope relation is banned because it involves a free variable or a quantifier with no variable to bind. Let's examine, in the light of the PAVQ, the two ordering restrictions illustrated in 4.2.2 that seemed to be irreducible to purely conceptual restrictions.

4.3.2 Quantificational aspectual adverbs and completive/prospective aspectual adverbs

Aspectual adverbs such as 'usually', 'often', 'always' or 'soon' can be regarded as quantifiers over events, whereas others, such as 'completely'/'partially', can be regarded as predicates that must be conjoined to the verb. Accordingly, if predicate adverbs were to occur higher than quantificational adverbs, the resulting expression would be equivalent to that of voice occurring higher than tense: the eventive variable associated with the predicate adverb would be free, violating the PAVQ.

- (22) a. Peter will have soon completely eaten the sandwiches
 soon(e): [(completely)(e) & Agent(Peter)(e) & eat(sandwiches)(e)]
 a'. *Peter will have completely soon eaten the sandwiches
 (completely)(e) & [soon(e): [Agent(Peter)(e) & eat(sandwiches)(e)]]

- b. Peter usually completely eats the sandwiches
usually(e): [(completely)(e) & Agent(Peter)(e) & eat(sandwiches)(e)]
- b'. *Peter completely usually eats the sandwiches
(completely)(e) & [usually(e): [Agent(Peter)(e) & eat(sandwiches)(e)]]

Note that a prospective adverb like ‘almost’ cannot readily be combined with a predicate that does not involve a gradation: #‘he is almost tall’ vs. ‘he is almost 2 meters tall’; #‘he is almost sitting down’ vs. ‘he is almost at home’; in a closed interval $[0, 1]$ of a property p where 0 is the minimal degree of p and 1 the maximal degree of p , and for all x such as $0 \leq x \leq 1$, ‘almost’(p) denotes an x prominently close to 1. This suggests that there must be a semantic conjunction operation (Kratzer’s Event Identification) that checks whether the properties of ‘almost’ are compatible with those of the verb. If ‘almost’ is merged higher than the existential quantifier, then the eventive variable of the adverb cannot be bound, which yields a violation of the PAVQ condition.

- (23) a. Peter will soon almost be at home
soon(e): [(Almost)(e) & be (Peter, at home)(e)]
- a'. *Peter will almost soon be here
(Almost)(e) & [soon(e): [be (Peter, at home)(e)]]
- b. Peter often is almost on time
often(e): [(Almost)(e) & be (Peter, on time)(e)]
- b'. *He almost is often on time
(Almost)(e) & [often(e): [be (Peter, on time)(e)]]

Thus we account for Cinque’s observation that prospective adverbs cannot occur higher than proximative adverbs (23), a scope relation that cannot be banned by purely conceptual restrictions, along with the more general observation (22) that predicate aspectual adverbs (‘completely’, ‘partially’) cannot occur higher than quantificational aspectual adverbs (‘always’, ‘often’, ‘soon’, ‘usually’).

4.3.3 Evidential mood and epistemic modality

There is a clear distinction between evidential categories such as the adverb ‘evidently’ or the adjective ‘evident’ and epistemic categories such as the adverb ‘possibly’ or the adjective ‘possible’: the former are veridical operators, while the latter are non-veridical operators.

- (24) Given an epistemic model M , an operator Op is
- a. veridical iff $[[Op(x)]] = 1 \models [[x]] = 1$ in M
- b. non-veridical iff $[[Op(x)]] = 1 \not\models [[x]] = 1$ in M
- c. anti-veridical iff $[[Op(x)]] = 1 \models [[x]] = 0$ in M

Thus, when the sentence ‘it is evident that he has left the hotel’ is uttered, the proposition under the scope of ‘evident’, ‘he has left the hotel’, must be true in the relevant *M*. When the sentence ‘it is possible that he has left the hotel’ is uttered, the proposition ‘he has left the hotel’ is neither true nor false in *M*.

An epistemic marker scopes over a proposition that does not yet have a truth-value specified in *M* because its variable over worlds has not been bound, whereas an evidential marker takes a proposition with a truth-value in *M* because its variable over worlds has already been bound by an overt or covert modal quantifier. If epistemic modals are merged lower than evidential markers, they provide an existentially quantified proposition for the evidential marker. However, when epistemic modals are introduced higher than an evidential marker, problems arise, suggesting that there is no variable over worlds to be bound by the epistemic modal.

More precisely, the source of the violation of the PAVQ in (25.a) lies in the conjunction of (i) an opacity factor and (ii) a poverty factor: (i) the evidential mood projection prevents the higher epistemic modal from binding a lower variable over worlds and (ii) the evidential mood projection fails to introduce a suitable variable over worlds for the higher epistemic modal. If the epistemic modal scopes over a full CP containing an evidential projection, the scope relation is legitimate (25.b): the matrix clause where ‘probabile’ is merged contains a variable over worlds to be bound with no intervening category.

- (25) a. *Probabilmente Gianni ha evidentemente lasciato l'albergo
 ‘Probably G. has evidently left the hotel’
- b. E' probabile che sia evidente que lui e' il colpevole
 ‘It is probable that it is evident that he is the guilty one’

This reasoning can be applied to other familiar cases. If Cinque (1999a) and Iatridou & von Stechow (2003) were correct in claiming, respectively, that quantificational aspectual adverbs and tense cannot scope over epistemic modality and that nominal quantifiers cannot scope over epistemic modals, the source of such restrictions could be (i) that tense, quantificational aspectual adverbs and nominal quantifiers fail to bind its variable, which is arguably in the VP, across epistemic modality (this possibility is in fact explored by Iatridou & von Stechow 2003 in the case of epistemic modality and nominal quantifiers), and (ii) that the projection of epistemic modality fails to introduce a suitable variable. Nonetheless, discourse may favor these non-prominent readings and coerce the epistemic modality projection into introducing a new eventive variable.

4.4 The Full Interpretation Principle

The main idea I have tried to pursue in this chapter is that the Infl-hierarchy is not an arbitrary solution, but satisfies conditions imposed by the C-I system; concretely, it must satisfy the PATS and the PAVQ. A syntactic object that violates the PAVQ contains elements that are not useful to the Conceptual-Intensional system, such as a free variable or a quantifier that does not bind any variable. Similarly, a discontinuous template that is broken off automatically becomes a useless element: C and *v* would fail to connect Infl and V to the

superordinate structure/discourse and to the external argument/manner adverb. Thus, the PATS and the PAVQ are two instances of the more general Full Interpretation Condition.

Full Interpretation Condition

An object generated by a grammatical component must be constituted only of useful features for a particular level of interpretation when it attains such a level of interpretation

Consequently, the empirical results of the cartographic project do not lead us to a primitive element of the syntactic component, but rather to the study of the C-I system. Although we are far from a good understanding of the restrictions the C-I system imposes on the combination of functional categories, there is no compelling reason to think that cartographies belong to the set of primitive elements of the faculty of language, as there is no compelling reason to think that a deeper understanding of a phenomena cannot be attained by reducing an apparent property of it to higher order principles.

Devices such as the PATD and the PAVQ are clearly relative to the levels of interpretation of the C-I system, and hence one may be skeptical about coding them in the theory of grammar in the form of derivational devices or in the form of universal hierarchies. As Chomsky (1991/1995: 151) observes: “when he deals with the prohibition of vacuous quantification, an aspect where “language differs from typical formal systems that permit vacuous quantification freely”, “[I]f some theory of grammar stipulates specific devices and rules to bar such constructions and interpretations [(28) and (29)], we conclude that it is the wrong theory: it is generating expressions and structures too accurately and is therefore incorrect. There is nothing paradoxical about this conclusion. The unwanted constructions are excluded on general grounds, in terms of the overarching condition FI [Full Interpretation]; there is no reason to suppose that the mechanisms of language include superfluous devices and rules to achieve, redundantly, the same result in special cases. Similarly, the phonological component contains no rules to express special cases of general properties of universal phonetics or of phonetic representations”.

- (28) a. who John saw Bill
- b. who did John see Bill
- c. every some person

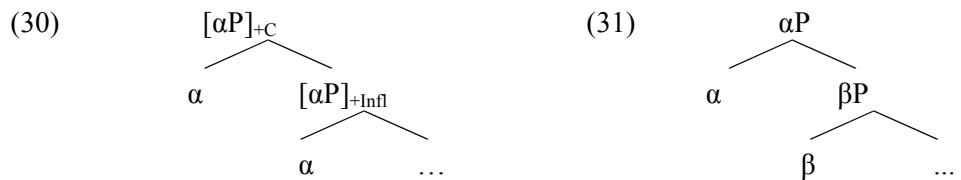
- (29) a. who did Mary see him
- b. the man that Mary saw him

The theory of the syntactic component of the faculty of language need not –and therefore must not– encode devices that translate the kind of external requirements to be satisfied: the syntactic component does not have to be defined to avoid the generation of (28)-(29); similarly, if cartographies derive from the Full Interpretation condition, cartographies do not reflect the knowledge of grammar. To be accurate, there is no overgeneration: there are no generated expressions that cannot be used at the performance systems, but a multitude of effects that expressions may have; if the devices of the levels of interpretation of the C-I system that are independently in place are provided with the information to discriminate the

multiple type of deviance effects –an empirical requirement that the theory must satisfy–, overgeneration disappears. The syntactic component is thus a computational mechanism that arranges and re-arranges features or instructions, with a wide range of external effects, and with no real distinction between grammatical and ungrammatical expressions (see Chomsky 2005a and references cited therein, especially Borer 2004.a, b).

4.5 Cartographic effects

It is important to note that discontinuous syntactic objects (30) can coexist with analytic syntactic objects (31).



Discontinuous syntactic objects do not contradict the idea that the C-I system requires articulated syntactic representations with semantically devoted positions, but rather they adhere to it: according to the line of argumentation of this study, the simultaneous insertion of Infl-like features into C and Infl is the way to provide Infl-like features with a [+clause typing] occurrence and a [-clause typing] occurrence, the semantic instruction provided by Infl-like features being orthogonal to the [\pm clause typing] distinction. Thus, when an Infl-like feature such as tense is introduced into the C-Infl discontinuous syntactic template, it is simultaneously associated with a [+clause typing] position and with a [-clause typing] position. Both occurrences of tense are minimally distinguished, and therefore, to be accurate, they are not the same element in two positions but two covariations. Thus, the creation of [\pm clause typing] variations (or discontinuous syntactic objects) is at the service of the creation of syntactic representations with semantically devoted positions (or analytic syntactic objects). Discontinuous syntactic objects are a subtype of analytic syntactic objects.

Chapter 5. Syncretic Syntactic Patterns

This chapter argues that syncretic syntactic patterns, where more than one feature is matched in one projection, are favored, under special syntactic circumstances, by a principle of structural minimization, the Maximize Matching Effects Principle. I begin by considering the nature of V-movement and verbal inflection, concluding that (i) movement of both maximal and minimal categories takes place in narrow syntax, (ii) functional features attached on V convert V into a multicategorial word, which does neither mean that they are uninterpretable on V nor that they must be deleted, (iii) the position of V is determined by the presence (and not the richness) of inflectional features attached on V, and (iv) the so-called Null Subject Parameter derives from the richness (and not the presence) of verbal inflection. I propose three instances of syntactic syncretic patterns: (i) in the TP of Null Subject Languages, φ -features are matched by a finite V in T and π -features such as topicality and referentiality are matched in SPEC-T by a preverbal subject or a CLLD element (Solà 1992); (ii) in the TP of V2 languages, φ -features are matched by A-movement to SPEC-T and tense is matched in T (and not in C) only when the first constituent is a subject (Zwart 1993), which avoids initiating the CP to match only tense, a feature that can be matched either in T or in C, due to its discontinuous character; (iii) in English, nominative interrogatives match φ -features and wh-features in SPEC-T, which becomes a mixed A'-A-position (similar patterns are obtained in Dutch with topical/focal 2SG subjects). Finally, I raise the possibility that C and T may contract when both φ -features and π -features are matched in T in certain languages, which would yield that-trace effects.

5.1 On structural minimization

All syntactic models implicitly or explicitly assume a constraint that bans superfluous structure. Consider, as a starting point, the following very rough formulation of this constraint in terms of an Elemental Economy Principle.

Elemental Economy Principle

A syntactic object cannot contain superfluous elements

Since the quality of superfluity referred to a syntactic object is not inherent to the object but relative to the requirements imposed precisely by a level of interpretation, what bans syntactic objects with superfluous elements is not an economy or minimization constraint that favors simplicity in the creation of structure, such as the Elemental Economy Principle, but rather a general convergence condition such as the Full Interpretation Condition.

Full Interpretation Condition

An object generated by a grammatical component must be constituted of only useful elements for a particular level of interpretation when it attains such a level of interpretation

Thus, a syntactic object containing elements that do not provide useful instructions for a particular level is not an excessively sophisticated object but an object crashing at that particular level of interpretation; such an object is to be filtered by a convergence condition,

such as the Full Interpretation Condition, which applies at the relevant level of interpretation, not by an economy principle. Convergence conditions and principles of structural minimization are not to be confused.

The conception of the Full Interpretation Condition in Chomsky & Lasnik (1993) and in the foundational texts of Minimalism (Chomsky 1991, 1993, 1995) has an unclear status; for instance, Chomsky & Lasnik (1993: 27) claim that “There are also certain general ideas that appear to have wide acceptability, among them, principles of economy stating that there can be no superfluous steps in derivations (Chomsky 1986b, chapters 2-4 of [Chomsky 1995])” and Chomsky (1991: 130) argues that “Some of these guidelines have a kind of ‘least effort’ flavor to them, in the sense that they legislate against ‘superfluous elements’ in representations and derivations. Thus the notion of ‘Full Interpretation’ (FI) requires that representations be minimal in a certain sense”. But to be accurate, the Full Interpretation Condition does not require representations to be minimal but rather to provide objects that converge with the relevant legibility requirements.

In chapter 3 I revised two plausible principles of derivational minimization, the Strict Cycle Condition, a particular version of the Transformational Cycle, and the Phase-Impenetrability Condition, and I concluded that it is unclear whether they play any role in narrowing the set of conceivable syntactic derivations: there is no compelling argument for strict cyclicity, especially once parallel probing at the phase level (or equivalently, at the level of discontinuous units) is taken into account, and successive cyclicity is not relative to the so-called phase heads. Note that the Relativized Opacity Principle, proposed as an alternative to account for *wh*-islands effects and subextraction patterns, can be indistinctly viewed either as a constraint reducing the search domain of a probe and the number of possible derivations, or as a sign that the searching power of a probe is not unbounded. Thus, it is unclear whether the Relativized Opacity Principle is an efficiency factor or an effect of the feebleness of probes.

This chapter explores whether genuine and unambiguous factors of structural minimization matter for the recursive generative procedure of the syntactic component. The candidate that will be examined is the Maximize Matching Effects Principle, understood as follows:

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

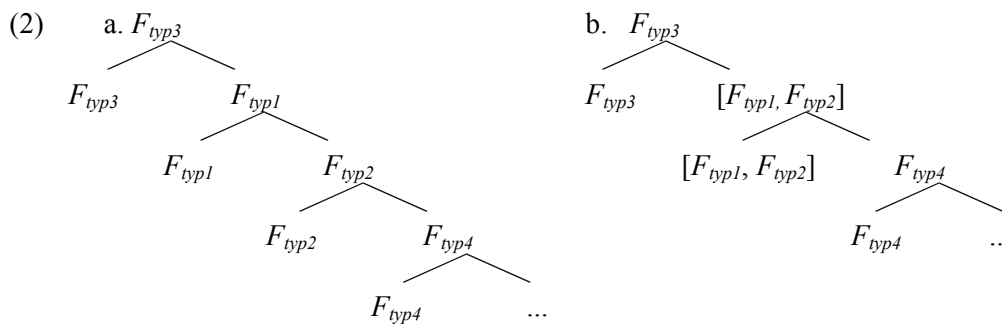
The most obvious difficulty in exploring the MMEP is how to measure the size of a span of structure where a set of type features are instantiated by external or internal merge of token features. Take a syntactic derivation that must match two different type features F_{typ1} and F_{typ2} with two different token features F_{tok1} and F_{tok2} . In assigning each type feature a particular position in the syntactic hierarchy of functional type features, viewed as a strict order relation O , the syntactic component may connect F_{typ1} and F_{typ2} to the rest of elements and to each other (1.a) or it may not connect them to each other although it connects them to the rest of type features (1.b).

- (1) a. $O_1 = \{ \langle F_{typ3}, F_{typ1} \rangle, \langle F_{typ1}, F_{typ2} \rangle, \langle F_{typ2}, F_{typ4} \rangle, \langle F_{typ3}, F_{typ2} \rangle, \langle F_{typ3}, F_{typ4} \rangle, \langle F_{typ1}, F_{typ4} \rangle \}$

$$b. \quad O_2 = \{ \langle F_{typ3}, F_{typ1} \rangle, \langle F_{typ3}, F_{typ2} \rangle, \langle F_{typ2}, F_{typ4} \rangle, \langle F_{typ3}, F_{typ4} \rangle, \langle F_{typ1}, F_{typ4} \rangle \}$$

In O_2 , F_{typ1} and F_{typ2} are not ordered with respect to each other ($\langle F_{typ1}, F_{typ2} \rangle \notin O_2$ and $\langle F_{typ2}, F_{typ1} \rangle \notin O_2$), and they appear in the same minimal context (their immediate precedent is F_{typ3} and their immediate consequent is F_{typ4}).

In standard arborian representations, it could be claimed that two different type features occur in the same minimal context when they are contained in the same head of the same projection. Thus, the syntactic hierarchies of functional features given in (1) can be expressed as in (2).



(2.a) is an analytic syntactic representation where all type features are strictly ordered, whereas in (2.b) there are two type features that are not ordered with respect to each other, F_{typ1} and F_{typ2} . The tree diagram (2.b) contains what can be called a syncretic syntactic pattern, since two different type features appear in the same minimal context, occurring in the same head ($[F_{typ1}, F_{typ2}]$). If it can be argued that two different type features can be instantiated in the same head, then there is reason to think that the MMEP is operative in the syntactic component, as the span of structure used to instantiate two type features in the syncretic syntactic pattern (2.b) is smaller than the one used in the analytic syntactic object (2.a): in the former case, a sole projection is used, while in the latter two different projections are used.

Whereas the preceding chapter argues that cartographic effects (analytic syntactic patterns and discontinuous syntactic patterns) derive from the C-I requirement of bringing semantic distinctions into the syntactic representation, this chapter states that a principle of structural minimization, the MMEP, is the source of anti-cartographic effects (syncretic syntactic patterns). Note that the MMEP is not concerned with banning useless elements or crashing objects but with condensing the structural span where useful features are coded and matched; in this sense, it is a genuine principle of structural minimization.

Accordingly, the way to explore whether the MMEP plays any role in the computations of the syntactic component is to determine whether syncretic syntactic patterns exist and define the special syntactic circumstance that favors a principle of structural minimization to the detriment of the C-I legibility condition that requires semantically devoted positions or a strict linear order of type features. Several controversial issues such as the role of verb

movement in the computations of the syntactic component, the nature of inflectional verbal morphology, the inexistence of vacuous movement, the existence of mixed A-/A'-positions and the possibility that structure contracts will be studied carefully leading to an array of delicate conclusions on the mechanics of creating structure.

5.2 Why and where V moves

The following four propositions may be taken to reflect the –perhaps idealized– predominant current view on verb movement in the minimalist approach (henceforth, V-movement). Consider them as a starting point.

- (3)
- I. V-movement does not take place at narrow syntax but at the PF-branch
 - II. Infl-morphology on V is uninterpretable
 - III. V moves from the vP if and only if it bears rich Infl-morphology
 - IV. The Null-Subject Parameter derives from V-to-T movement, and hence from the existence of a rich Infl-morphological paradigm in a particular language

Whereas propositions I and II are characteristic of the main stream of minimalism, and they are not present, as far as I know, in the pre-minimalist period of the Principles and Parameters era, propositions III and IV reflect two important generalizations formulated in the 80s that allowed for certain progress in deriving the degree of cross-linguistic variation from morphological properties: if V shows up ‘rich’ Infl-morphology in a language *L*, *L* displays movement of V outside the vP (see, among others, Pollock 1989) and allows for postverbal subjects and null subjects (see, specially, Taraldsen 1980, Chomsky 1981 and Rizzi 1982b, 1986). Propositions III and IV have not been discredited by minimalist inquiries, although there has not been much progress in understanding them. I find that the belief that SPEC-T must be universally filled (the Extended Projection Principle translated in current terminology as the EPP feature of T) and the minimalist tenets conveyed in propositions I and II have not been a good guide.

Proposition I. V-movement does not take place at narrow syntax but at the PF-branch

It may be worth considering two arguments given for the idea that V-movement does not take place at narrow syntax (see Acedo & Fortuny 2006 for a more general discussion of operations that do not seem to have semantic effects, such as scrambling and clitic climbing).

Firstly, Chomsky (1995) observes that, whereas movement of maximal categories follows the Extension Condition as it always targets the root, head movement never does, since it is a case of syntactic infixation. For several reasons given throughout this study, Chomsky’s observation is no longer empirically valid: at the phase level, the Extension Condition is not operative, since movement of maximal categories takes place in parallel; crucially, because movement to SPEC-T and movement to SPEC-C occur during the same derivational step,

only the operation that creates SPEC-C extends the root, any operation of A-movement being a case of syntactic infixation, triggered by the phase head C in Chomsky's (2005b) analysis. Indeed, if one were to cast doubt on any of the premises of the argument at hand, this is the Extension Condition, an even stricter derivational constraint than the Strict Cycle Condition: if operations take place in parallel at the phase level, the Extension Condition is neglected at the phase level and the classical empirical evidence given for the Strict Cycle Condition (that is, for an extension condition at an inter-phase level) is not relevant, as argued in chapter 3.

Secondly, whereas movement of maximal categories has clear semantic effects (it provides peripheral positions, scope positions for quantifiers and binding positions for nominative and object DPs), head movement is usually argued to be semantically vacuous (Chomsky 2000; but see Acedo & Fortuny 2006, and references cited therein, for cases in which clitic climbing reflects semantic changes in the eventive structure). On the one hand, the observation that V-movement does not have the type of semantic effects that are attested in movement of maximal projections cannot be taken as a valid premise to conclude that it should be confined to the PF branch: V cannot have the same semantic contribution as a quantifier because it is not a quantifier and it cannot be interpreted correferentially to an anaphor because it is not an argument. On the other hand, it should not be ignored that, as discussed below, V-movement interacts with movement of maximal categories in such a way that a DP moves to SPEC-T only when V remains inside the vP, an observation that can be more easily accounted for if both V-movement and DP-movement take place at the same computational level.

Thus far there is no reason to assume that movement of maximal categories and movement of minimal categories take place at different computational levels, say, the former at narrow syntax and the latter at the PF branch. They might both take place at narrow syntax or at the PF branch. In fact, the only requirement at the C-I system is for token features and type features to be matched; whether or not probe-goal relations become spatially material could be a matter of parameterization linked to the morpho-phonological component.

However, some well-known patterns in English discussed by Laka (1990) suggest that V-movement takes place at narrow syntax, since it is relevant for operations that hold at the C-I system. As illustrated in (4.a), a negative polarity item occurring at SPEC-T cannot be licensed by a lower negative marker in an assertive matrix sentence; if V raises to C carrying the attached negative marker as a free-rider, then the negative polarity item can be licensed (4.b).

- (4) a. *Anybody didn't come
b. Didn't anybody come?

For these reasons, I shall keep to the idea that both movements of minimal and maximal categories take place at narrow syntax, a position that will be further supported as the argumentation proceeds (see Zwart 2001 for a broader discussion of the nature of V-movement).

Another interesting issue observed by Laka (1990) is that an embedded C selected by a negative matrix verb has the ability to license a negative polarity item in SPEC-T. In (5.a), ‘anybody’ is unlicensed, because it is not in the domain of the embedded negative marker, whereas in (5.b) it is licensed. The contrast between (5.b) and (5.c) shows that it is not the negative characterization of matrix V that directly licenses the embedded negative polarity item, but the embedded C: if the embedded CP layer is absent, the negative polarity item is unlicensed (Progovac 1988).

- (5) a. *I say that anybody didn’t leave
b. I deny that anybody left
c. I deny *anybody/everything

This suggests, as argued by Laka (1990), that a matrix negative verb selects an embedded C specified as [+neg], just as a matrix interrogative verb selects an embedded C specified as [+wh] (‘I wonder what he saw’), a further piece of evidence for the idea that negation can be operative on C.

Proposition II. Infl-morphology on V is uninterpretable

In section 1.3 of chapter 1, I argued on the basis of the problem of generality, the problem of determinacy and the problem of consistency that the source of movement was not the existence of uninterpretable functional features located in the syntactic hierarchy of features just to be deleted, a device called suicidal greed. φ -features of T, like the rest of Infl-like features located in the hierarchy (tense, mood, modality and negation), were claimed to provide useful instructions for the C-I system since they are responsible for referential displacement. Since the property of movement dissolves into the basic operation of merge, the trigger of internal merge has to be the same as that of external merge: a type feature present in the syntactic hierarchy must be instantiated by a suitable token feature and a token feature needs to be hosted in a suitable position in order to be properly interpreted; when a type feature is instantiated by a token feature present in its search domain, a probe-goal relation takes place, which may then become spatially material.

A partially different question is whether the tokens of inflectional features attached to V are uninterpretable, as stated in proposition II, and consequently, whether the source of V-movement lies in deleting the uninterpretable features on V before the derivation arrives at the C-I system. If that were the case, uninterpretable features of V would instantiate a type Infl-like feature and be deleted from the syntactic object when moving toward the C-I system.

It could be argued that φ -features are uninterpretable on V because V denotes properties of an event and person-number distinctions are relative to arguments of that event, to the nominative DP in the case of subject agreement, to the accusative DP in the case of object agreement or to the dative DP in the case of dative agreement. This reasoning can be applied to other functional features that appear attached on V, such as tense and mood, as they do not reflect internal properties denoted by V such as telicity or duration, but rather an ordering relation between the time where the event takes place and a reference time in the case of

tense and the [\pm realis] characterization of the proposition or the source of evidence for a statement in the case of mood.

Let us assume that the reasoning outlined above is correct, and that therefore φ -features, tense and mood are uninterpretable on V. As will be argued below, type φ -features can be alternatively instantiated through movement of a V inflected for φ -features or by a DP; similarly, it is a strong cross-linguistic tendency for languages to instantiate tense or mood type features (and generally, functional features) through V-movement or by inserting a maximal category (an adverbial) or a minimal category (a free particle). This means that matching Infl-like type features could either provide a suitable position for the relevant token to be interpreted or a suitable position for a token to be deleted in the sole case that such a token is attached to V. All this seems to be an unnecessary complication: it remains unclear why the morpho-phonological property of being attached or free should be relevant in determining whether the matching type feature offers a position for the token to be deleted or to be interpreted. In fact, to say that φ -features, tense and mood are uninterpretable on V seems to be a misleading way of expressing that V contains (in some languages) not only inherent features (say, thematic-features and Aktionsart features), but also non-inherent features that are acquired through the morphological process of attachment.

There is still a further problem for the assumption that attached features are uninterpretable and must be deleted before the object attains the C-I system. V undergoes (overt) movement only when it displays rich inflectional paradigms; otherwise, it remains inside the vP (proposition III). If the need for uninterpretable attached features to be deleted were the source of V-movement, rich inflection would never be overt: when inflectional paradigms were pre-syntactically rich, V would undergo movement and its attached features would be deleted during syntactic computations. Accordingly, moved verbs would be expected to be bare verbal roots and in situ verbs, poorly inflected verbs or bare verbal roots. However, we can see that displaced Vs do not get their inflection deleted from the object to which morpho-phonological form is assigned.

Stipulating that the deletion of attached features takes place during the C-I branch after the syntactic object has been sent to the A-P system is futile: if V has undergone movement, a probe-goal relation must have taken place; in this case, a type feature is instantiated and a token feature is hosted or deleted (in the sole case that it is attached). On the one hand, it is unclear why instantiation of a type and deletion of a token should have different timings, the former holding before the syntactic object is sent to the A-P system, and the latter during the C-I branch after the syntactic object is sent to the A-P system; optimally, both instantiation and deletion should be two immediate effects of the same matching operation, and therefore should take place during the same derivational step. On the other hand, if both typing and deletion took place at the C-I branch, after the syntactic object was sent to the A-P system, then there would be no reason for V to overtly move. The most straightforward account is that there is no deletion operation and no uninterpretable features to be deleted, all token features being hosted by a suitable type feature at the same step at which the type feature is automatically instantiated.

These conceptual considerations solely serve to show, along with the problem of generality, the problem of determinacy and the problem of consistency exposed in section 1.3, that the interpretable/uninterpretable dichotomy can neither be a powerful analytical tool nor provide a deeper understanding of the relation between displacement patterns and inflectional paradigms. Here I shall keep to the idea that any device especially constructed for displacement is suspicious. There are thus four valid alternative strategies for a token feature to instantiate a type feature, with no interpretable/uninterpretable dichotomy: (a) external merge of a minimal category, (b) internal merge of a minimal category, (c) external merge of a maximal category and (d) internal merge of a maximal category.

Proposition III. V moves outside the vP if and only if it bears rich Infl-morphology

As argued by Pollock in a very important paper (Pollock 1989), which in fact set the stage for cartographic studies, there is a systematic difference as to the position at which a lexical verb surfaces in French and in English: in French, but not in English, a lexical verb can move to an Infl-position, preceding the negative marker *pas* (5), a frequentative aspectual adverb (6), or a stranded quantifier (7) argued to be, à la Kayne (1975), in an adverbial position, or to a C position, which gives rise to auxiliary-subject inversion (8).

- (5) a. * $[\text{InflP John } [\text{InflP likes } [\text{NegP not } [\text{NegP likes } [\text{vP likes Mary}]]]]]$
 b. $[\text{InflP Jean } [\text{InflP (n')aime } [\text{NegP pas } [\text{NegP (n')aime } [\text{vP aime Marie}]]]]]]]$
- (6) a. * $[\text{InflP John } [\text{InflP kisses } [\text{AdvFreqP often } [\text{vP kisses Marie}]]]]]$
 b. $[\text{InflP Jean } [\text{InflP embrasse } [\text{AdvFreqP souvent } [\text{vP kisses Marie}]]]]]$
 c. $[\text{InflP John } [\text{InflP } \emptyset [\text{AdvFreqP often } [\text{vP kisses Marie}]]]]]$
 d. * $[\text{InflP Jean } [\text{InflP } \emptyset [\text{AdvFreqP souvent } [\text{vP embrasse Marie}]]]]]$
- (7) a. * $[\text{InflP My friends } [\text{InflP love } [\text{AdvP all } [\text{love Marie}]]]]]$
 b. $[\text{InflP Mes amies } [\text{InflP aiment } [\text{AdvP tous } [\text{vP aiment Marie}]]]]]$
 c. $[\text{InflP My friends } [\text{InflP } \emptyset [\text{AdvP all } [\text{vP love Marie}]]]]]$
 d. * $[\text{InflP Mes amies } [\text{InflP } \emptyset [\text{AdvP tous } [\text{vP aiment Marie}]]]]]$
- (8) a. * $[\text{CP Likes } [\text{InflP he } [\text{vP likes Marie}]]]?$
 b. $[\text{CP Aime-t-il } [\text{InflP } [\text{vP aime-t-il Marie}]]]?$

(Adapted from Pollock 1989: 367)

In (6.b), Pollock assumed that the optional preclitic negative marker *n'*- (<'ne') had moved pied-piped by V from the head of NegP, where 'pas' is spelled out as a SPEC, to the head Infl via successive head movement, thereby satisfying Emonds' Head Movement Constraint (Emonds 1978). Another possibility is that 'ne' and 'pas' occur in different negative projections (see Zanuttini 1997).

It is tempting to connect, as Pollock did, these syntactic minimal pairs, which reflect that lexical verbs must move outside the vP in French and cannot move outside the vP in English, to the different degree of robustness or richness of inflectional paradigms in the two

languages. Clearly, French inflectional paradigms are richer than English ones: whereas French inflectional paradigms display distinctions of tense (present, past and future), mood (indicative and subjunctive) and person-number features (1st [\pm singular], 2nd [\pm singular] and 3rd [\pm singular]), English inflectional paradigms do not display any productive subjunctive/indicative distinction and future is not expressed by inflectional marks but by means of verbal periphrasis ('will kiss', 'is going to kiss', etc.); apparently, through inflectional means English expresses only a temporal present/past distinction and a 3SG mark for present.

(9)

	<i>Present</i>	<i>Past</i>
1SG.	kiss	kiss-ed
2SG.	kiss	kiss-ed
3SG.	kiss-(e)s	kiss-ed
1PL.	kiss	kiss-ed
2PL.	kiss	kiss-ed
3PL.	kiss	kiss-ed

Thus, a richly inflected V leaves the vP in French and a poorly inflected V does not leave the vP in English.

According to the reasoning above, the richness of inflectional paradigms is assumed to be determined by the number of overt morpho-phonological distinctions, an observable property, although the morphological analysis of inflectional forms is rarely transparent, especially in syncretic morphological patterns. Since languages vary as to the number of overt morpho-phonological distinctions, languages vary as to the relative richness of their inflectional paradigms. Consequently, one should ideally be able to predict whether in a language L V leaves the vP simply on the basis of the number of overt morpho-phonological distinctions of the inflectional verbal paradigm in L . Or equivalently, one can test whether proposition III is accurate simply by correlating the position of V in different languages and the number of morpho-phonological distinctions in the inflectional verbal paradigms in each language; if, all things being equal, the verbal inflectional paradigms of a language L_1 display a smaller or equal number of morpho-phonological distinctions than the verbal inflectional paradigms of a language L_2 and V leaves the vP in L_1 but not in L_2 , then proposition III is contradicted. See Platzack & Holmberg (1989), Roberts (1993) and Vikner (1994) for two attempts to decide the critical number of substantial distinctions for V to move to Infl.

However, there are both empirical and conceptual reasons to suspect that inflectional richness is not the real trigger of V-movement.

For instance, Swedish, like Norwegian and Danish, manifests very poor inflectional paradigms. In Swedish, mood distinctions are conveyed exclusively through verbal periphrasis ('*må* + infinitive), as future tense ('*skulle*' + 'infinitive'). Inflectional verbal paradigms only distinguish past from present tense, with no person-number opposition.

(10)

	<i>Present</i>	<i>Past</i>
1SG.	kyss-er	kyss-te
2SG.	kyss-er	kyss-te
3SG.	kyss-er	kyss-te
1PL.	kyss-er	kyss-te
2PL.	kyss-er	kyss-te
3PL.	kyss-er	kyss-te

Therefore, inflectional verbal morphology is slightly poorer in Swedish than in English: apparently, in both languages V displays a past-present distinction, but only in English V displays a 3SG marking. Since a finite lexical V leaves the vP in Swedish but not in English, our starting supposition that morphological richness determines whether V leaves the vP should be brought into question.

Proposition III reflects a conceptual oddity that deserves some attention. It would be quite natural if Infl type features did not require a lexical V to move to T if inflectional features attached on V did not display a relatively high number of distinctions: type features would not bother to host an inflected V unless its inflection was robust enough to satisfactorily instantiate them. However, verbal inflectional features need to undergo movement to be assigned a suitable position to be interpreted; therefore, unless type Infl features did not unselfishly host a poorly inflected V, the derivation would crash, because it would contain tokens that could not be appropriately interpreted. Thus, for convergence reasons, any V should move to the C-T discontinuity if it bears attached inflectional features, as relatively poor or relatively rich as they may be, and a V can remain inside the vP only if it completely lacks attached inflectional features. This is clear for French and Mainland Scandinavian lexical Vs, which display Infl-like features, but it seems problematic for the standard analysis of English lexical verbs, according to which they manifest tense features and a person mark for present 3SG.

The idea that the trigger of V-movement is not the relative richness of inflectional paradigms or the strength of features (in the sense of Chomsky 1993, 1995) but the mere presence of inflectional features attached on V has already been developed in some detail by Solà (1996). In this study, Solà addresses the problem posed by English lexical verbs, and he concludes that they can be viewed as non-finite forms, concretely, as participial forms. The core argument goes as follows (see Solà 1996 for a detailed argumentation). A very solid observation is that the past form of regular verbs in modern English has become homophonous with the participial form ('work-ed'). This means that, at some historical stage of English, the past form may have been reanalysed as a past participle. If this is so, the lexical verb 'worked' appears in the same position in both 'John has worked' and 'John worked'. Such a participle is claimed to bear an aspectual feature: [-progressive]. Moreover, "a present morpheme –Solà (1996: 233-234) argues– cannot exist if it does not stand in opposition to a past morpheme (just like singular cannot exist without a plural). It should then also have been reanalyzed as a participle". However, 'works' and 'worked' do not contrast in aspect, as they both are [-progressive]. Solà solves this potential problem by arguing that they

are selected by null tense morphemes: “when [-Past], it would select ‘arrives’, when [+Past], it would select ‘arrived’. In other words, temporal features are not glued in V but covertly present in the syntactic hierarchy of functional features as null morphemes. English lexical verbs would not directly convey temporal features, but aspectual features, and the morpho-phonological contrast between ‘arrives’ and ‘arrived’ derives from the value of the selecting tense; hence, the commonly called inflectional morphology of ‘present’ -Ø/-s and the commonly called inflectional morphology of ‘past’ are suppletive forms of a [-progressive] morpheme. Finally, the morpheme -s is analyzed as just number agreement (see also Kayne 1989), and the observation that number agreement is confined to the 3rd person is related to the participial status of the verbal form.

Solà’s reasoning implies that English lexical verbs are not spelled out in their base position, since their aspectual [-progressive] features must be hosted by an appropriate aspectual type feature. It is known that the position where V is spelled out must be lower than negation and quantificational adverbs (5-8) but, as argued above, it must be higher than the A-position where an accusative DP occurs, which has a VP-adjunct in its domain, as it is able to bind an anaphor and to license an NPI into such a VP-adjunct. The following constructions, adapted from Lasnik (2001), are repeated from section 3.2.2.

Condition A satisfaction

- (11) a. The DA proved [two men to have been at the scene of the crime] during each others’s trials (ECM)
 b. ?*The DA proved [that two men were at the scene of the crime] during each other’s trial (finite embedded sentence)
 c. The DA accused two men during each other’s trials

NPI licensing

- (12) a. The DA proved [noone to have been at the scene] during any of the trials
 b. ?*The DA proved [that noone was guilty] during any of the trials
 c. The DA cross-examined none of the witnesses during any of the trials

These surprising patterns were accounted for following Chomsky’s (2005b) idea that an accusative DP is attracted by v* to SPEC-V, just as a nominative DP is attracted by C to SPEC-T, and V moves to a higher functional head, v*, or perhaps an aspectual head.

I thus adhere to Solà’s main idea that there is good reason to keep to the simplest possibility, namely that V-movement is triggered by the presence of inflectional features, and not by the relative richness of inflectional paradigms. The counterexample provided by English verbs may be just apparent, as a finer morphological analysis of English lexical verbs, along the course developed by Solà, should make clear. This opens a very interesting and subtle path for cross-linguistic research: the cross-linguistically variant distribution of V must be determined by the cross-linguistically variant number and type of non-inherent functional features (voice, aspect, tense/mood/ ϕ -features, force) glued on V.

Proposition IV. The Null-Subject Parameter derives from V-to-T movement, and hence from the existence of a rich Infl-morphological paradigm in a particular language

Null Subject Languages (NSLs) are defined by their ability to allow not only preverbal subjects (13) but also postverbal subjects with no overt preverbal expletive (14) and omitted subjects (15). Accordingly, Catalan is a NLS whereas English is a non-NSL, or Catalan is positively specified for the Null-Subject Parameter whereas English is negatively specified for it.

(13) *Preverbal subjects*
a. En Pere va treballar
b. Peter worked

(14) *Postverbal subjects*
a. Va treballar en Pere
b. *Worked Peter

(15) *Omitted subjects*
a. Va treballar
b. *Worked

In the wake of studies by Taraldsen (1980), Chomsky (1981) and Rizzi (1982b, 1986), it has been generally understood that the Null-Subject Parameter is an epiphenomenon derived from morphological properties, more particularly, from the richness of verbal inflectional paradigms. This is indeed a very solid empirical generalization that offers an interesting framework for the study of parameter setting and morpho-syntactic diachronic process.

Note that, if both the Null-Subject Parameter and the observation that V leaves the vP were to be derived from verbal inflectional richness, the rate of inflectional richness triggering V-movement would be different from the rate of inflectional richness that yields the Null-Subject Parameter, as modern French, a language displaying V-to-T/C movement (16.a) is not parametrized as a NSL (16.b-c).

(16) a. Marie travaille souvent
b. *Travaille souvent Marie
c. *Travaille souvent

A more straightforward account can be attained if, as argued above, the trigger of V-to-T/C movement is not the inflectional richness, but rather the presence of Infl-features on V. The T-head in a language like modern French would have to unselfishly trigger V-to-T movement in order to warrant convergence at the external systems (i.e., to warrant that inflectional features attached on V were interpreted at the appropriate position); however, such a token would not be rich enough to instantiate type ϕ -features in a language like modern French, and thus a DP must be merged to SPEC-T to avoid the derivation to crash (i.e., to avoid a type feature to arrive at the C-I system without being satisfactorily instantiated). Therefore, EPP-effects appear when V is [-fin], as in English, but also when V has undergone V-to-T

movement and is poorly inflected, as in French. They do not appear when V is richly inflected, in which case a V inflected for Infl-features is hosted by an Infl type feature in an appropriate position, and this Infl type feature is satisfactorily instantiated by the richly inflected V.

The idea that it is precisely the inflectional richness, and not the presence of φ -features on V, that defines a language as a NSL is strongly supported by cross-linguistic and diachronic considerations. Old French, a language with verbal inflectional paradigms richer than Modern French, was a NSL (see Adams 1987 and Roberts 1993). Similarly, several dialects of Modern Brazilian Portuguese manifest simpler inflectional verbal paradigms than do traditional Brazilian Portuguese or modern European Portuguese: the 2SG form ‘tu falas’ and 2PL form ‘vós falais’ have been replaced by the old address form ‘você fala’ and ‘vocês falam’. In addition, as Kato observes (Kato 1999: 4), “the first person plural ‘nós’ is being replaced by the nominal ‘a gente’ (the folks = we folks), which triggers 3rd person singular agreement, thus reducing the paradigm to three distinctive forms”.

(17) Present indicative of ‘falar’ (“speak”)

	Traditional Brazilian Portuguese Modern European Portuguese		Modern Brazilian Portuguese	
1SG	eu	fal-o	fal-o	fal-o
2SG direct	tu	fal-as	---	---
2SG indirect	você	fal-a	fal-a	fal-a
3SG	ele/-a	fal-a	fal-a	fal-a
1PL	nós	fal-amos	fal-amos	(a gente) fal-a
2PL direct	vós	fal-ais	---	---
2PL indirect	vocês	fal-am	fal-am	fal-am
3PL	eles/-as	fal-am	fal-am	fal-am
		(6/8)	(4/8)	(3/8)

Whereas Traditional Brazilian Portuguese and Modern European Portuguese are NSLs (and the inflectional richness of present indicative paradigms is the same as two other NSLs, Standard Spanish *–yo habl-o, tu habl-as, usted habl-a, él habl-a, nosotros habl-amos, vuestros habl-ais, ustedes habl-an, ellos hablan* (6/8)– and Standard Catalan *–jo parl-o, tu parl-es, vostè parl-a, ell parl-a, vosaltres parl-eu, nosaltres parl-em, vostès parl-en, ells parl-en* (6/8)–), the dialects of Modern Brazilian Portuguese we are concerned with do not omit referential subjects (21), although they (still) manifest null expletives (18), omission of arbitrary subjects (19), null bound pronouns (20.a) and anaphoric null subjects (20.b). See Duarte (1993) and Kato (1999) for further discussion. The following illustrations are taken from Kato (1999: 5).

- (18) a. Tá chovendo
‘It is raining’
b. Tem novidade
‘There is news’

- c. Parece que vai chover
'It seems that it is going to rain'
- (19) a. Aqui pode fumar
'You/one can smoke'
b. Aqui conserta sapatos
'One repairs shoes'
- (20) a. Ninguém acha que é estúpido
'Nobody_i thinks that he_i is stupid'
- b. O João disse que comprou um carro
'John_i said that he_i has bought a car'
- (21) a. *(Eu) como pizza
'I eat pizza'
b. *(Você) come pizza
'You eat pizza'
c. *(Ele) come pizza
'He eats pizza'
d. *(A gente) come pizza
'The folks (we folks) eat pizza'

As argued by Sigurðsson (1993), Old Icelandic allowed not only null agreeing subjects (22) but also null non-agreeing verbal objects (23) and null non-agreeing prepositional objects (24).

- (22) ok kom hann_i þangat, ok var Hoskuldr uti, er —_i reið i tún
and came he there and was H. outdoors when rode into field
'And *he* came there, and Hoskuldr was outdoors when (*he*) rode into the field'
- (23) dvergrinn mál_i lti, at sa baug_i skyldi vera hverjum hofuþsbani, er atti —_i
the dwarf said that *that ring* should be to-anybody a headbane that possessed
'The dwarf said that *that ring* should bring death to anybody possessed (*it*)'
- (23) þá tla ek, at þú nýtir eigi boga minni þottu
believe I that you (can-)use not bow my even-if-you
spyrnir fótum i —_i
push with-feet in
'I believe that you cannot use *my bow* even if you push with your feet in (*it*)'

(Sigurðsson 1993: 248)

Thus, the grammatical strategies that allow null subjects in languages like old Icelandic and Catalan must be different: in the former, null subjects are part of more general coindexing algorithms that do not depend on verbal inflection and hence extend to non-agreeing objects,

whereas in the latter, the existence of null subjects derives from inflectional richness, which is not involved in licensing null objects (25). In simpler terms, old Icelandic was not a Null-Subject Language but a Null Argument Language, and Catalan is a Null-Subject Language (NSL) but not a Null Argument Language (NAL).

- (25) [Ø]_{subject} va donar el llibre a tots els estudiants que *(el) van demanar
 ‘(He/she) gave the book to all the students that asked for (it)’

According to Sigurðsson’s analysis, null subjects and objects in old Icelandic were “identified under free coindexing with an NP in preceding discourse (much like Old Icelandic null-topics, which could, however, also be identified under free coindexing with a construed discourse topic)”. The claim that it is not rich verbal inflection that identifies type φ -features but coindexing with an NP present in preceding discourse is strongly supported by the observation that in old Icelandic prose null arguments never initiate discourse, since a preceding linguistic context is required. This is a clear difference with respect to null subjects in NLSs: the subject is omitted in (26.b) although it cannot be coindexed with any NP present in preceding discourse.

- (26) a. Què fa en Joan?
 what do.3SG ART Joan?
 ‘What is Joan doing?’
 b. (#jo) no ho sé
 I no it-ACC know.1SG
 ‘I don’t know it’

Interestingly, modern Icelandic has lost both null referential subjects and null referential objects “without any concomitant or previous reduction of rich verbal inflectional”, as illustrated below with the present indicative of *leita* ‘search’, *segja* ‘say’, and *sjá* ‘see’, borrowed from Sigurðsson (1993: 249). Non-referential subjects can still be null in modern Icelandic, as in the dialects of modern Brazilian Portuguese considered above, which suggests that abstract φ -features do not need to be instantiated by rich inflectional morphology.

(27)	Old Icelandic			Modern Icelandic		
1SG.	leita	segi	sé	leita	segi	sé
2SG.	leitar	segir	sér	leitar	segir	sér λ
3SG.	leitar	segir	sér	leitar	segir	sér
1PL.	leitum	segjum	sjáum	leitum	segjum	sjáum
2PL.	leitit	segit	sjáit	leitit λ	segit λ	sjáit λ
3PL.	leita	segja	sjá	leita	segja	sjá

Sigurðsson’s results are very important for the understanding of discourse indexing mechanisms, leading to the conclusion that, even if in old Icelandic the inflected V obligatorily moves to T, it does not instantiate type features on T, which are instantiated by coindexation to an NP in preceding discourse. Note that Sigurðsson does not reject the idea that rich verbal inflection is the decisive factor in defining a language as a NSL, but he convincingly shows that verbal inflection plays no role in defining a language as a Null

Argument Language, in which free discourse indexing (and not verbal inflection) plays a crucial role in instantiating Infl-features.

From this revision I conclude that the three propositions in (3) must be replaced or reformulated in the following terms:

- I. Both movement of maximal and minimal categories take place during syntactic computations; thus both V-movement to T/C and DP-movement to T take place during syntactic computations.
- II. Inflected verbs contain not only inherent features (thematic-features and Aktionsart features), but also non-inherent functional features (tense, φ -features, mood, aspect, etc.). Functional features attached on V convert V in a multicategorical word, which means neither that they are uninterpretable on V nor that they must be deleted.
- III. The position of V is determined by the non-inherent functional features attached on V; particularly, V leaves the vP moving up to C or to T if it is finite.
- IV. A language *L* is a NSL only if verbal inflection is rich in *L*; *L* is a NAL if and only if *L* displays free discourse indexing mechanisms, independently of verbal inflection.

5.3 A side-effect of V-to-T movement

In this section I shall introduce the first possible case of syntactic syncretism, and the least controversial one, for the analysis I shall pursue is somehow present in the literature on NSLs: since richly inflected Vs instantiate φ -features on T, SPEC-T is available to match other features.

I shall begin by providing in 5.3.1 some evidence for the idea that SPEC-T is not a position where φ -features are matched in NSL (see Solà 1992, among others, for a broader discussion): SPEC-T is both too restrictive (only a subset of subjects can occur there) and too permissive (non-subject Clitic Left Dislocated (CLLD) elements do also occur there). On this basis, I shall argue that π -features, namely, topic and referential features, are matched in SPEC-T, but not φ -features.

In 5.3.2 I shall discuss the source of π -features matched in SPEC-T; I shall first focus on a very recent proposal by Gallego (Gallego 2006): in NSLs, the phase head *v* moves along with V to T, a syntactic operation that has the effect of transforming T into a phase head (5.3.2.1); thus π -features on T in NSLs would be provided by *v*. After showing some technical and empirical problems for Gallego's 'phase sliding', I shall propose that π -features surfacing on T are inherent to the C-T discontinuity, not inherited from the lower *v*-V discontinuity. Accordingly, [*v*-V]-to-T movement does not provide T with π -features, but creates a syntactic circumstance that favors structural minimization to the detriment of the C-I legibility requirement of creating semantically devoted projections: as an inflected V is responsible for matching φ -features, the MMEP favors SPEC-T to be used to match π -

features, and not to redundantly match φ -features. If this is so, in a single projection, both φ -features and π -features are matched, thereby minimizing the span of structure used for syntactic computations.

In 5.3.3 I shall close the discussion of NSLs by noting the problematic status of null expletives (Rosselló 2000, among others) and the inadequacy of postulating covert movement of postverbal subjects to SPEC-T.

5.3.1 SPEC-T becomes an A'-position when V instantiates φ -features on Infl

Consider the following contrasts between English and Catalan, which illustrate a very well known generalization: any kind of subject occurs preverbally in a non-NSL like English, whereas, in a NSL like Catalan, only a subset of subjects does.

5.3.1.1 Definite subjects in NSLs

The constructions in (28) show that a definite subject that bears new information cannot occur preverbally in Catalan (28.a), unless it bears focal stress and undergoes focus fronting, in which case it arguably occurs at Rizzi's FocP of the C-area (28.b); but subjects that bear new information can occur postverbally (28.c), and the verb, which conveys old information present in discourse, is optionally omitted (28.d).

(28) *Definite subjects bringing new information*

Qui va venir?

- a. #En Pere va venir
- b. EN PERE va venir
- c. Va venir en Pere
- d. En Pere

The constructions in (29) illustrate that, if a definite subject expresses old information present in discourse, it can occur preverbally (29.a), and optionally, an intonational pause can appear between the topical preverbal subject and the verb (29.b), in which case the preverbal 'subject' may be left-dislocated to a TopP of the C-area. Definite subjects that bring old information cannot occur postverbally (29.c) and can optionally be omitted (29.d).

(29) *Definite subjects bringing old information*

Què fa en Pere?

- a. En Pere llegeix
- b. En Pere(,) llegeix
- c. #Llegeix en Pere
- d. Llegeix

Crucially, English referential preverbal subjects are not informationally restricted: referential subjects can (indeed, must) appear preverbally, independently of whether they bring new (30) or old (31) information.

(30) *Definite subjects bringing new information*

Who came?

- a. Peter came
- b. *Came Peter

(31) *Definite subjects bringing old information*

Where did Peter study?

- a. Peter studied in Barcelona
- b. *Studied Peter in Barcelona

5.3.1.2 Non-definite subjects in NSLs

Non-definite preverbal subjects must satisfy further constraints; as argued by Solà (1992), bare indefinite subject NPs cannot occur preverbally in Catalan and Italian, while they must occur preverbally in English (32). Negative subject quantifiers and counting subject quantifiers too cannot occur preverbally in NSLs (34-35), but must occur preverbally in non-NSLs (33). The distribution of these three kinds of non-definite quantifiers in NSLs and non-NSLs is thus complementary. All these contrasts are adapted from Solà (1992, chapter 5), where the very intricate restrictions that govern preverbal subjects are delineated in great detail. My sole intention here is to express in a clear way Solà's general results, which will be useful for research on syntactic syncretism.

- | | | |
|------|--|---------|
| (32) | a. Students have arrived
a'. *Have arrived students | English |
| | b. *Estudiants han arriuat
b'. Han arriuat estudiants | Catalan |
| | c. *Studenti sono arrivati
c'. Sono arrivati studenti | Italian |
| (33) | a. Nothing has happened
a'. *Has happened nothing
b. Few students have come
b'. *Have come few students | |
| (34) | a. *Res ha passat
a.' No ha passat res
b. *Pocs estudiants han vingut
b'. Han vingut pocs estudinats | |

- (35) a. *Niente è successo
 a'. Non è successo niente
 b. *Pocchi studenti sono arrivati
 b'. Sono arrivati pochi studenti

Indefinite subjects can appear preverbally or postverbally in Catalan; if they appear preverbally they are unambiguously interpreted as referential expressions, whereas if they appear postverbally they can be either referential or non-referential (37). In English, these subjects obviously must occur preverbally, and are thus ambiguous (36).

- (36) *Indefinite subjects in English: 'a stone'*
 a. A stone has fallen down the hill
 ≡ 'One of the stones has fallen down the hill' (referential reading)
 ≡ 'An indefinite stone has fallen down' (pure existential reading)

- (37) *Indefinite subjects in Catalan: 'un roc' ("a stone")*
 a. Un roc ha caigut de la muntanya (referential interpretation of *un roc* 'a stone')
 ≡ 'One of the stones has fallen down the hill' (referential reading)
 ≠ 'An indefinite stone has fallen down' (pure existential reading)
 a'. Ha caigut un roc de la muntanya
 ≡ 'An indefinite stone has fallen down' (pure existential reading)
 ≡ 'One of the stones has fallen down the hill' (referential reading)

Therefore, in NSLs, definite subjects can occur preverbally only if they are topical; bare NPs, negative quantifiers and counting quantifiers cannot occur preverbally at all, and indefinite quantifiers can occur preverbally only if they are referentially interpreted. If the preverbal position is so restricted for subjects in NSLs and so unrestricted for subjects in non-NSLs, it is reasonable to conclude that preverbal subjects in non-NSLs match only φ -features; as to preverbal subjects in NSLs, there are two possibilities: they match (a) both φ -features and topicality/referentiality features or (b) only topicality/referentiality features, which are indeed those that constrain the set of legitimate preverbal subjects in NSLs. As will be illustrated below following Solà's arguments, there is reason to think that option (b) is the correct one.

5.3.1.3 Clitic-Left Dislocated elements in NSLs

Non-definite quantifiers can be resumed by a clitic only if they can be interpreted referentially; thus, bare NPs, negative quantifiers such as 'nothing' and counting quantifiers such as 'few' cannot be CLLD elements (38.a-c), whereas universal quantifiers and existential quantifiers can be (38.d-f), in which case they are unambiguously referential. English translations do not reflect resumption, for obvious reasons.

- (38) a. *Estudiants_i els_i he vist (Catalan)
 students them.have.1SG seen

'I have seen the students'

- b. *Res ho he vist
nothing it.have.1SG seen
'He/she has done nothing'
- c. *Pocs estudiants_i els_i he convidat
few students them.have.1SG invited
'He/she has invited few students'
- d. A tots aquests alumnes_i els_i veuré
to all these pupils them.will.see.1SG
'I will see all the pupils'
- e. A un meu amic_i el_i van arrestar
to one my friend him.arrested.3PL
'They arrested one of my friends'; or rather, 'one of my friends was arrested'
- f. A dos amics meus_i els_i han arrestat
to two friends mine them.have.3PL arrested;
'They have arrested two friends of mine'; or rather 'two friends of mine have been arrested'

Definite DPs can be CLLD elements only if they are topical (39a' vs. 39.b'), in which case they involve a pause only optionally (39.a').

- (39) a. Has vists els estudiants?
Have.2SG seen.PL the students
'Have you seen the students?'
- a'. Els estudiants_i(,) no els_i he vist
the students not them.have.3SG seen
- b. A qui has vist?
to whom have.2SG seen?
'Who have you seen?'
- b'. #Els estudiants_i els_i he vist
the students them.have.1SG seen
'I haven't seen the students'

Therefore, the restrictions imposed on preverbal subjects and resumed elements are the same (if they are definite, they must be topical (28, 39), if they are non-definite, they must be referential (37, 38)). A further question is whether it can be decided whether the position where both preverbal subjects and CLLD elements occur is the same, as would be expected: if they satisfy the same semantic requirements, they should instantiate the same type features.

5.3.1.4 Both preverbal subjects and CLLD elements match π -features in SPEC-T in NSLs

It seems that “the more a quantifier is likely to be interpreted as D-linked or referential, the more it is acceptable as dislocated (Solà 1992)”. Firstly, a [-human] negative quantifier cannot appear preverbally and with a resumptive clitic, as illustrated in (38.b), whereas a [+human] negative quantifier (*ningú* ‘nobody’) or a universal quantifier (*tothom* ‘everybody’) can appear preverbally and with a resumptive clitic (in the case they are non-subjects). Although ‘ningú’ and ‘tothom’ are referential enough to be resumed (contrary to *res* ‘nothing’), they cannot be dislocated; but other quantifiers that have a more prominent referential reading, such as *dos amics* ‘two friends’ and *alguns amics* ‘some friends’ can be resumed by a clitic (if they are non-subjects) and moreover dislocated. Examples are again from Catalan.

(44) *Dislocation*

- | | | | | |
|-----|---|--------------------------|----------|-----------------------|
| a. | A ningú _i (^{*/#} ,) | l _i 'han | acceptat | (resumed non-subject) |
| | to nobody | him.have.3PL | accepted | |
| a'. | Ningú(^{*/#}) | ha aprovat l'examen | | (subject) |
| | nobody | has passed the exam | | |
| b. | A tothom _i (^{*/#} ,) | l _i 'han | acceptat | (resumed non-subject) |
| | to everybody | him.have.3PL | accepted | |
| b'. | Tothom(^{*/#}) | ha aprovat l'examen | | (subject) |
| | everybody | has passed the exam | | |
| c. | A dos amics meus _i (,) | els _i han | arrestat | (resumed non-subject) |
| | to two friends of mine | them.have.3PL | arrested | |
| c'. | Dos amics meus(,) | han tingut un accident | | (subject) |
| | two friends of mine | have.3PL had an accident | | |
| d. | Alguns amics _i (,) | els _i han | arrestat | (resumed non-subject) |
| | some friends | them-have.3PL | arrested | |
| d'. | Alguns amics(,) | no hi vindran | | (subjects) |
| | some friends | not there.come.3PL | | |

Crucially, those quantifiers that cannot be dislocated must occur in a position adjacent to the inflected verb both if they are subjects and non-subjects, which strongly argues that both subjects and resumed non-subjects appear in the same position, SPEC-T. Those quantifiers that can be dislocated, because they are more prominently referential, may have moved

further, to SPEC-Top of the C-area, independently of whether they are subjects or non-subjects.

(45) *Adjacency effects*

- a. A ningú_i (*aquesta vegada) l_i'han acceptat (resumed non-subject)
 to nobody (this time) him.have.3PL accepted
- a'. Ningú (*^{/?}aquesta vegada) ha aprovat l'examen (subject)
 nobody (this time) has passed the exam
- b. A tothom_i (*^{/?}aquesta vegada) l_i'han acceptat (resumed non-subject)
 to everybody (this time) him.have.3PL accepted
- b'. Tothom (*^{/?}aquesta vegada) ha aprovat l'examen (subject)
 everybody (this time) has passed the exam
- c. A dos amics meus_i (avui) els_i han arrestat (resumed non-subject)
 to two friends of mine (today) them.have.3PL arrested
- c'. Dos amics meus (avui) han tingut un accident (subject)
 two friends of mine (today) have.3PL had an accident
- d. Alguns amics_i (avui) els_i han arrestat (resumed non-subject)
 some friends (today) them-have.3PL arrested
- d'. Alguns amics (al casament) no hi vindran (subjects)
 some friends (to the wedding) not there.come.3PL

It is thus fair to conclude that preverbal subjects in Catalan, a NSL, do not match φ -features but only π -features such as topicality and referentiality; since (non-subject) resumed elements instantiate the same π -features and, as preverbal subjects, appear adjacent to V unless they can be dislocated to a TopP of the C-area, both CLLD elements and preverbal subjects occur in SPEC-T, a position used for matching topic and referential features, instead of φ -features, unlike in English. The ability of both preverbal subjects and resumed elements to be dislocated to SPEC-Top depends on their relative degree of referentiality.

5.3.2 The source of π -features associated with SPEC-T when V instantiates φ -features on T

5.3.2.1 Subextraction (III): phase sliding

Gallego (2006) considers the possibility that, although C and v are universal inherent phase heads, the fact that v moves along with V to T (instantiating Infl-features) in NSLs endows the TP with the phase properties characteristics of v and frees the vP from taking over its

phase responsibilities. A crucial point for Gallego's analysis is that of subject islands in NSLs: if preverbal subjects are in SPEC-T and postverbal subjects are in SPEC-v, and subextraction from preverbal agentive subjects is more problematic than subextraction from postverbal agentive subjects, then there is good reason to think that, in NSLs, TP behaves, derivatively, as a phase, while v*P has lost its phasehood; in Gallego's terms, a 'phase sliding' would have taken place. This obviously raises the possibility, as Gallego argues, that the A'-status of SPEC-T in NSLs (its topicality/referentiality features) is introduced by the v head by means of [V-v]-to-T movement.

The cases of subextraction that Gallego considers can be classified in two types. Though he does not mention this distinction, it is rather instrumental in understanding subextraction patterns in NSLs:

- (a) The subextracted PP and the DP from where the PP has been subextracted are spelled out in the same clause: $[_{CP1} [_{PP} \dots]_i C_1 \dots [_{DP} \dots t_i]]$
- (b) The subextracted PP and the DP from where the PP has been subextracted are spelled out in different clauses, the former in a superordinate one and the latter in a subordinated one: $[_{CP1} [_{PP} \dots]_i C_1 \dots [_{CP2} C_2 \dots [_{DP} \dots t_i]]]$

In (a), neither subextraction from what Gallego considers an agentive preverbal subject nor from what he considers a postverbal agentive subject leads to an entirely acceptable output (46.a/b). Note that subextraction from a postverbal non-agentive subject is unproblematic (46.c), as expected. These Spanish examples are borrowed from Gallego (2006: 84); example (46.c) is from Uriagereka (2004: 9). Direct questions in written Spanish are always preceded by a question mark '¿' ('¿Qué dices?' "what are you saying?"), which is not to be confused with the superindex '??', which indicates acceptability judgements.

- (46) a. ?? $[_{CP} De\ qué\ artistas_i\ han\ herido_z [_{TP} t_z [_{v*P} tu\ sensibilidad_j [_{v*P} [las\ obras\ t_i] t_z t_j]]]]?$
of what artists have.3PL hurt your sensitivity the works
'Which artists have the works by hurt your sensitivity?'
- b. * $[_{CP} De\ qué\ artistas_i\ han\ herido_z [_{TP} [las\ obras\ t_i]_j t_z [_{v*P} t_j t_z tu\ sensibilidad]]]]?$
of what artists have.3PL hurt the works your sensitivity
'Which artists have the works by hurt your sensitivity?'
- c. $[_{CP} De\ qué\ artistas_i\ han\ llegado_z [_{TP} ya\ t_z [_{vP} t_z [las\ obras\ t_i]]]]?$
of what artists have.3PL arrived already the works
'Which artists have the works of arrived?'

Thus, both subextraction from a preverbal and a postverbal agentive subject seem to yield degraded outputs, which means that the meager contrasts provided by type (a) does not provide a solid argument for phase sliding.

Observe in the above representations that Gallego assumes that V has moved from T to C in questions across the adverb 'ya' ("already"); thus the position of the subject is determined

with respect to the object ('tu sensibilidad'): when it precedes the object, it is considered to appear in SPEC-T, following V in C, and when it follows the object, it is considered to be in its base-position, SPEC-v*, with the object shifted. However, it is unclear that V moves from T to C in (46): the position of V relative to aspectual adverbs such as 'ya' does not indicate whether V is in T or in C, for V can follow a subject and precede such an adverb also in an assertive clause in Spanish (47.b). In either of these two constructions, no pause intonation is needed. Note also that the verb can follow 'ya' in questions (47.c).

- (47) a. Las obras ya han llegado
 the works already have.3PL arrived
 'The works have already arrived'
- b. Las obras han llegado ya
 the works have.3PL arrived already
 'The works have already arrived'
- c. De qué artistas ya han llegado las obras?
 of what artists already have.3PL arrived the works
 'Which artists have the works of arrived?'

Whether V ever raises to C in Catalan and Spanish is a topic of debate; see Suñer (1994) and Ordóñez (1998) on the claim that V-to-C never occurs in Spanish and Catalan, and Gallego (2004) on the possibility that it does. However the case may be, the adverb 'ya' is not a good positional reference.

One might also be skeptical about Gallego's test to decide whether a subject is in SPEC-T or in SPEC-v* in questions: if it follows V and precedes the object it is in SPEC-T, and if follows the shifted object it is in SPEC-v*; again, in an assertive clause, the subject can follow V and at the same time precede the object (48.a). Since claiming that V raises to C in Spanish assertive clauses would be entirely unmotivated, the subject 'las obras de Pedro' may be in a position lower than T even if it precedes the object 'tu sensibilidad' (48.a).

- (48) a. Han herido las obras de Pedro tu sensibilidad
 have hurt the works of Pedro your sensibility
 'The works by Pedro have hurt your sensibility'
- b. Han herido tu sensibilidad las obras de Pedro
 have hurt your sensibility the works of Pedro
 'The works by Pedro have hurt your sensibility'

It would be more reasonable to conclude that the subject is in SPEC-v* in both (48.a) and (48.b); if this is the case, what would be variable is the position of the object: it remains in its base-position in (48.a) and it is shifted to a higher position in (48.b).

- (48') a. [_{TP} [Han herido]_i-T [_{v*P} [las obras de Pedro] t_i [_{VP} tu sensibilidad]]]
 b. [_{TP} [Han herido]_i-T [_{v*P} [tu sensibilidad]_j [_{v*P} [las obras de Pedro] t_i [_{VP} t_j]]]]

Consider the following contrast, also from Gallego (2006: 86), which seems to be stronger than the one provided in (46). The analysis in brackets of these constructions is mine. For abovementioned reasons, I am not following Gallego’s positional analysis.

- (49) a. $??\zeta_{CP} [De\ que\ universidad]_j [_{TP}\ causaron_i [_{VP^*} [los\ estudiantes\ [t_j]] [_{VP}\ t_i\ problemas]]]]?$
of what university caused.3PL.PAST the students problems
‘Of what university did the students cause problems?’
- b. $\zeta_{CP} [De\ que\ universidad]_j [_{TP}\ causaron [_{v^*p}\ problemas_i [_{v^*p}\ [los\ estudiantes\ [t_j]] [_{VP}\ t_i]]]]?$
of what university cause.3PL.PAST problems the students
‘Of what university did the students cause problems?’

If the PP ‘de qué universidad’ is replaced by the PP ‘de qué coche’ (“of which car”), and the DP ‘los estudiantes’ by the DP ‘el conductor’ (“the driver”), then there is no strong acceptability contrast: both (50.a) and (50.a’) are degraded outputs, as in (46.a/b), repeated in (50.b/b’). A further factor to consider is the presence of an article: if ‘problemas’ is replaced by ‘un problema’, subextraction yields an even less acceptable output (50.a’’).

- (50) a. $??\zeta_{CP} [De\ que\ coche\ causó\ problemas\ el\ conductor]?$
of what car cause.3SG.PAST problems the driver
‘Of which car did the driver cause problems?’
- a’. $*\zeta_{CP} [De\ que\ coche\ causó\ el\ conductor\ problemas]?$
of what car cause.3SG.PAST the driver problems
‘Of which car did the driver cause problems?’
- a’’. $???\zeta_{CP} [De\ que\ coche\ causó\ un\ problema\ el\ conductor]?$
of what car cause.3SG.PAST a problem the driver
‘Of which car did the driver cause a problem?’
- b. $??\zeta_{CP} [De\ que\ artistas_i\ han\ herido_z [_{TP}t_z [_{v^*p}\ tu\ sensibilidad_j [_{v^*p}\ [las\ obras\ t_i] t_z t_j]]]]?$
of what artists have.3PL hurt your sensitivity the works
‘Which artists have the works by hurt your sensitivity?’
- b’. $*\zeta_{CP} [De\ que\ artistas_i\ han\ herido_z [_{TP} [las\ obras\ t_i] t_z [_{v^*p}\ t_j t_z\ tu\ sensibilidad]]]]?$
of what artists have.3PL hurt the works your sensitivity
‘Which artists have the works by hurt your sensitivity?’

There are indeed several intricate issues involved in subextraction. For instance, Chomsky (2005b, note 38) observes that examples with the expression ‘the picture-PP’ introduce “extraneous issues because of the ambiguity of the phrase, which can be understood with the PP interpreted not as a complement of ‘picture’ but rather as a reduced relative clause (roughly, ‘I have a picture which is of Boston’ contrary to ‘*I saw a driver who is of the car’,

‘I saw an author who is of the book)’’. The same caution that led Chomsky to avoid constructions such as ‘Mary, of whom a picture hit me on the head’ should lead us to distrust expressions such as ‘the students of which university’, which can be interpreted as ‘I saw the students which are of University of Barcelona’, contrary to expressions such as ‘the driver of that car’ (‘the driver who is of that car’).

Thus, the empirical generalization is that subextraction of a PP from a DP located in SPEC-v* yields a degraded output when the PP is generated as a complement of the DP ([_{DP} [_{PP}]]), as in (50); when the DP and the PP can be interpreted as forming a reduced relative clause [DP PP] located in SPEC-v, extraction of PP is unproblematic (49.b). This is not at all surprising, as the PP is more embedded in the nested structure [_{DP} [_{PP}]] than in the symmetric structure [DP PP], and indeed, it could be accounted for by the Relativized Opacity Principle, reproduced below.

Relativized Opacity Principle (ROP)

In a syntactic object [_{α IP} [_{α 1} ... [_{α 2P} Δ [_{α 2} Γ]]]],

- where: (i) α_1 and α_2 are two probes of the same type α each projecting an α P
- (ii) Δ is SPEC- α_2 and Γ is α_2 -COMPL, and
- (iii) $\exists \delta$: δ is a constituent of Δ and $\exists \gamma$: γ is a constituent of Γ ,

α_1 can probe Δ or Γ if they provide a suitable token for α_1 , but it cannot readily probe γ or δ

When the nested structure [_{DP} [_{PP}]] is agentive, it is base-generated in the SPEC position of a π -projection, the v*P. Thus, the immediately higher C, a π -probe, can probe Δ , i.e., the whole [_{DP} [_{PP}]], but not a constituent δ of Δ , i.e., the PP. If reduced clauses are symmetric structures, with no projecting head, no complement and no specifier, then there is no projection Δ , and C can detect the whole [DP PP] but also the PP, which is not too deeply nested.

It is not entirely clear to me why extraction of PP from [DP PP] should be less degraded when the object moves to an extra-SPEC-v* preceding the agentive DP located in SPEC-v* (49.b) than when it follows the agentive DP (49.a). As a matter of fact, a similar contrast is attested in (49’), where the *wh*-phrase is not extracted from a subject.

- (49’) a. ?¿Cuándo causaron los estudiantes problemas?
 when cause.PAST the students problems
- b. ¿Cuándo causaron problemas los estudiantes??
 when cause.PAST problems the students
 ‘When did the students cause problems?’

If so, the contrast in (49) may not be related to subextraction, but rather to ordering restrictions between objects and subjects in Spanish questions. This reasoning may apply to

(46) as well: both (46.a) and (46.b) are degraded, but the least degraded one is that in which the object moves over the subject.

In sum, on the one hand, the contrast illustrated in (46.a/b) is too weak to be considered as proof of phase sliding, and on the other hand, it is doubtful that such a slight contrast indicates that subextraction from SPEC-v* is less problematic than subextraction from SPEC-T in NSLs, since there is no reason to think that the subject is spelled out in SPEC-T in the slightly more degraded (46.a): the subject can precede the object and follow the verb not only in questions (46.a) but also in assertions (48.a), which indicates that neither in (46.a) nor in (48.a) is V in C nor the subject in SPEC-T. A PP can be more easily extracted from a reduced relative clause [DP PP] located in SPEC-v* (49) than from a nested structure [DP [PP]], because in the former the PP is less embedded than in the latter.

Examples of type (b) given by Gallego seem to provide, initially, a stronger contrast between subextraction from a preverbal agentive subject and a postverbal agentive subject. Example (51) is from Uriagereka (2004:10).

- (51) a. ^{??*}¿De qué conferenciantes_i te parece [CP que [TP [las propuestas t_i]_j me van a impresionar [v*P t_j]]]?
Of what speakers CL-to-you seem-3SG that the proposals me go to impress
'Which speakers does it seem to you that the proposals by will impress me?'
- b. ¿De qué conferenciantes_i te parece [CP que [TP me van a impresionar [v*P [las propuestas t_i]]]]]?
Of what speakers CL-to-you seems-3SG that me go to impress the proposals
'Which speakers does it seem to you that the proposals by will impress me?'
- (52) a. ^{??}¿De qué escritor_i dice María [CP que [TP [muchas novelas t_i]_j la han impresionado [v*P t_j]]]]]?
of what writer say-3SG María that many novels CL-her have-3PL impressed
'Of what writer does María say that many novels have impressed her?'
- b. ¿De qué escritor_i dice María [CP que [TP la han impresionado [v*P [muchas novelas t_i]]]]]?
of what writer say-3SG María that CL-her have-3PL impressed many novels
'Of what writer does María say that many novels have impressed her?'
- (53) a. ^{??}¿De qué universidad_i dice la prensa [CP que [TP [muchos estudiantes t_i]_j han protestado [v*P t_j]]]]]?
of what university says the press that many students have protested
'Of what university do the media say that many students have protested?'
- b. ¿De qué universidad_i dice la prensa [CP que [TP han protestado [v*P [muchos estudiantes t_i]]]]]?
of what university say-3SG the press that have-3PL protested many students

‘Of what university do the media say that many students have protested?’

There are at least two misleading factors that may easily go unnoticed in examples (51-53).

Firstly, the examples of subextraction from agentive subject involve psych-verbs like ‘impress’ (51, 52) and unergative verbs like ‘protest’ (53); as Gallego warns (note 113), “I will avoid introducing verbs with explicit direct objects as much as I can, since those may add additional problems”. Clearly, once verbs with explicit objects are used, subextraction from a postverbal agentive subject is degraded, and thus providing the same results as in English.

- (54) a. ??¿De qué coche causó problemas el conductor?
of what car caused problems the driver
- b. *¿De qué coche causó el conductor problemas?
of what car caused the driver problems
‘Of which car did the driver cause problems?’

Gallego does not specify what kind of problems may be added by explicit objects, but they may be not additional. A conceivable additional problem could be that an explicit object yielded minimality effects; but, significantly, what we observe are anti-minimality effects: in (49.a), the explicit object precedes the subject, in (49.b), the object follows the subject, and nonetheless, subextraction yields a degraded output in (49.b) and not in (49.a).

It could be simply that different θ -grids correspond to different vP-structures, and that the acceptability of subextraction from an argument depended on the exact position where such an argument is base-generated. This is envisioned by Chomsky, as Gallego acknowledges (2006: note 113):

“Choice of v* may have an effect. Perhaps, ‘of which books did the author receive the prize’ is more acceptable than ‘of which car did the driver cause a scandal’. If so, difference among theta roles might be relevant, perhaps requiring a deeper analysis of base structures” (Chomsky 2005b: note 39)

In other words, the contrast of acceptability between (55.a/a’) and (55.b) may not reflect that objects add additional problems, but that transitive predicates and psych-predicates have different argument structures.

- (55) a. ??¿De qué coche causó problemas el conductor?
of what car caused problems the driver
- a’. ??¿De qué coche causó el conductor problemas?
of what car caused the driver problems
‘Of which car did the driver cause problems?’
- b. ¿De qué conferenciantes me van a impresionar las propuestas?
Of what speakers I-DAT GO.PL to impress the proposals

‘Of which speakers are the proposals going to impress me?’

In (55.a/a’) the base position of the agentive DP ‘el conductor de qué coche’ (“the driver of which car”) is outside the search domain of v* (it is an external argument of a transitive verb), whereas the DP ‘las propuestas de qué conferenciantes’, which is not an external argument or an agent of the psych-verb, is base-generated in the search domain of v. Accordingly, the vP structure of the predicates in (55) would be like these.

- (55’)
- a. [_{v*P} [el conductor de qué coche] [_{v*P} v*_[+φ, +ag] [_{VP} V problemas]]] (‘cause’)
 - b. [_{VP} v_[-φ, -ag] [_{VP} [las propuestas de qué conferenciantes] [_{VP} V GOAL]]] (‘impress’)

And secondly, in (53), if the PP ‘de qué universidad’ is replaced by the PP ‘de qué coche’ and the DP ‘los estudiantes’ by the DP ‘el conductor’, which cannot be interpreted as constructing a small clause [DP PP], then subextraction of PP is degraded both when the subject follows the object and when it precedes the object, as noted above with respect to (49-50), repeated in (56.b/b’). Accordingly, subextraction from the subject of unergative verbs and subextraction from the subject of a transitive verb, both external arguments, yield the same result.

- (56)
- a. *¿De qué coche dice la prensa que el conductor ha protestado?
of what car says the press that has protested the driver
‘Of which car does the media say that the driver has protested?’
 - a’. ??¿De qué coche dice la prensa que ha protestado el conductor?
of what cars says the press that have protested many drivers
‘Of which cars does the media say that many drivers have protested?’
 - b. ??¿De qué coche causó problemas el conductor?
of what car cause problems the driver
‘Of which car did the driver cause problems?’
 - b’. *¿De qué coche causó el conductor problemas?
of what car cause the driver problems
‘Of which car did the driver cause problems?’

If the tests are defined by avoiding misleading factors such as psych-verbs (which lack an agent and hence a thematic SPEC-v*) or small relative clauses [DP PP] appearing in SPEC-v*), we can conclude that in Catalan and Spanish, two NSLs, C cannot subextract a PP from SPEC-v*.

Firstly, both subextraction from the external argument of a transitive verb (57) and an unergative verb (58) yield degraded outputs; (57) illustrates that subextraction from the agentive subject is problematic either when the subject raises to an extra SPEC-v* preceding the object (57.a) as well as when it remains in situ (57.b). Secondly, subextraction from non-agentive subjects is unproblematic (59, 60, 61), as is subextraction from verbal objects (62): the v head probes the PP when it is in V-COMPL and relocates it in SPEC-v, a position in

which C can unproblematically access the PP. The examples of subextraction from non-agentive subjects involve passive subjects (59), subjects of impersonal constructions (60), and unaccusative subjects (61), all base-generated in V-COMPL. This paradigm can be extended to expressions such as “the architect of which building”, the writer of which book”, etc.

(57) *Subextraction from an agentive subject of a transitive verb*

Catalan

a. *[De quina pel·lícula]_i va provocar [el director t_i] un escàndol?
of which film PAST.SG cause the director a scandal

a'. *[De quina pel·lícula]_i va provocar un escàndol [el director t_i]?
of which film PAST.SG cause a scandal the director

‘Of which film did the director cause a scandal?’

Spanish

b. *¿[De qué película]_i provocó [el director t_i] un escándalo?

b'. *¿[De qué película]_i provocó un escándalo [el director t_i]?

Relativization test

c. *The director who is of that film

(58) *Subextraction from an agentive subject of an unergative verb*

Catalan

a. *[De quina pel·lícula]_i va treballar [el director t_i]?
of which film PAST.SG work the director

‘Of which film did the director work?’

Spanish

b. *¿[De qué película]_i trabajó [el director t_i]?

(59) *Subextraction from a non-agentive subject of a passive verb*

Catalan

a. [De quina pel·lícula]_i va ser aplaudit [el director t_i]?
of which film PAST.SG be applauded the director

‘Of which film was the director applauded?’

Spanish

b. ¿[De qué película]_i fue aplaudido [el director t_i]?

(60) *Subextraction from a non-agentive subject of impersonal 'se'-constructions*

Catalan

- a. [De quines pel·lícules]_i s'han aplaudit [els directors t_i]
of which films SE-have.3PL applauded the directors
'Of which films were the directors applauded?'

Spanish

- b. ¿[De qué películas]_i se aplaudiron [los directores t_i]?

(61) *Subextraction from an unaccusative subject*

Catalan

- a. [De quina pel·lícula]_i ha arribat [el director t_i]?
of which film have.3PL arrived the director
'Of which film has the writer arrived?'

Spanish

- b. ¿De qué película ha llegado el director?

(62) *Subextraction from a verbal object*

Catalan

- a. [De quina pel·lícula]_i va aplaudir [el director t_i] en Pere?
of which film PAST applaud the director ART Pere
'Of which film did Pere applaud the director?'

Spanish

- b. [De qué película]_i aplaudió Pedro [al director t_i]?

Although a PP cannot readily be subextracted from SPEC-v* in NSLs, as in English, there remains an important question to consider. The principal difference between type (a) and type (b) of Gallego's examples of subextraction is that in type (b) we can observe when the DP from which the PP is subextracted appears in a preverbal or in a postverbal position. Recall examples (51), (52) and (53).

- (63) a. ^{??/*}¿De qué conferenciantes_i te parece [_{CP} que [_{TP}[las propuestas t_i]_j me van a impresionar [_{v*P}t_j]]]?

Of what speakers CL-to-you seems-3SG that the proposals me go to impress
'Which speakers does it seem to you that the proposals by will impress me?'

- b. ¿De qué conferenciantes_i te parece [_{CP} que [_{TP} me van a impresionar [_{v*P} [las propuestas t_i]]]]]?

Of what speakers CL-to-you seems-3SG that me go to impress the proposals
'Which speakers does it seem to you that the proposals by will impress me?'

- (64) a. ^{??}¿De qué escritor_i dice María [_{CP} que [_{TP} [muchas novelas t_i]_j] la han impresionado [_{v*P} t_j]]]]?
 of what writer says-3SG María that many novels CL-her have-3PL impressed
 ‘Of what writer does María say that many novels have impressed her?’
- b. ¿De qué escritor_i dice María [_{CP} que [_{TP} la han impresionado [_{v*P} [muchas novelas t_i]]]]?
 of what writer says-3SG María that CL-her have-3PL impressed many novels
 ‘Of what writer does María say that many novels have impressed her?’
- (65) a. ^{??}¿De qué universidad_i dice la prensa [_{CP} que [_{TP} [muchos estudiantes t_i]_j] han protestado [_{v*P} t_j]]]]?
 of what university says-3SG the press that many students have-3PL protested
 ‘Of what university does the media say that many students have protested?’
- b. ¿De qué universidad_i dice la prensa [_{CP} que [_{TP} han protestado [_{v*P} [muchos estudiantes t_i]]]]?
 of what university says-3SG the press that have-3PL protested many students
 ‘Of what university does the media say that many students have protested?’

In these cases, when the DP is postverbal, subextraction of a PP is possible because, in (63) and (64), the DP is not an external argument but a theme or source, and in (65), the DP and the PP form a small clause [DP PP]. But why is subextraction more degraded when the DP is spelled out in SPEC-T of the embedded clause? Spanish and Catalan SPEC-v* and English SPEC-v* do not differ, but why do Spanish (and Catalan) SPEC-T differs from English SPEC-T, as illustrated in (63-65)?

A possible answer could be as follows. Gallego may be correct in claiming that [V-v]-to-T movement gives SPEC-T the same status in NSLs as SPEC-C and SPEC-v*, since it provides π -features to TP. Thus, although v*P is still a phase in NSLs, V-movement creates a new phase by inheritance; like in the frame of barriers, there would be inherent phases and derived phases. Accordingly, there would not be, strictly, phase sliding, but rather the creation of new phases as part of syntactic computations. Indeed, the idea that movement creates new effects is uncontroversial and, but the hypothesis that it eliminates previous effects from syntactic representations is problematic. Why should [V-v]-to-T movement free the v*P from its phasehood? Or why should [V-v*]-to-T movement prevent V from assigning a θ -role to its complement or to its SPEC? If movement could eliminate previous effects from syntactic representations, it would destroy semantic relations, and there would be no way of accounting for the observation that movement brings semantic richness by associating a category to different positions (predicate positions –v-area–, infl-positions –Infl-area– and illocutionary positions –C-area–).

The notion of “phase” disappears at the moment in which phases are claimed to be created derivationally. The notion of “phase” is not equivalent to the notion of “barrier”. Barriers

were defined on the basis of syntactic opacity phenomena, but phases are defined on the basis of completeness at the external systems: a syntactic object is a phase if it can be argued to be complete at the external systems C-I and A-P. Syntactic opacity is an effect of multiple transference operations that cyclically map those autonomous units from the syntactic workspace to the external systems, and thereby rendering them invisible at syntax. Since v-V and C-T constitute interpretive units (roughly, C-T conveys propositional features and v-V predicate features) or discontinuous elements both in NSLs and in English, there is no reason to think that [V-v]-to-T movement redefines TP as a ‘phase’, although it may still be accurate to consider that [V-v]-to-T movement redefines TP as a syntactic ‘barrier’ or as an opacity inducer for syntax.

Let us finally consider the source of π -features that surface in TP when an inflected V has instantiated φ -features on T. Reconsider Gallego’s intuition: [V-v]-to-T movement provides T with π -features. We know from Chomsky’s work on English subject islands that φ -features of T and π -features of C work in parallel; as far as I see, this is a conclusion one cannot easily discard. Thus, in English, φ -features trigger movement of a DP to SPEC-T and π -features trigger movement of a *wh*-phrase to SPEC-C. In Catalan, as argued above, φ -features are instantiated by a finite V in T, not by movement of a DP to SPEC-T. Consider now, for instance, the sentence given in (63.a). π -features of embedded C would probe a *wh*-phrase at the same time that φ -features probe the finite V (Σ_1 in (66)). In a further step (Σ_2), once T had been provided with topic-features by V-movement, the topical DP would move to SPEC-T. And this is the derivational paradox: SPEC-C would be created by remerging ‘de qué conferenciantes’ before SPEC-T was created by remerging ‘las propuestas’, and nonetheless, it is the presence of ‘las propuestas’ in SPEC-T that complicates the way C subextracts the PP ‘de qué conferenciantes’.

(66)

Σ_1 : [_{CP} [de qué conferenciantes]_i C [_{TP} impresionarán_j-T [_{v*P} [las propuestas t_i] t_j]]]
 (Parallel movement of V to T and of PP to SPEC-C)

Σ_2 : [_{CP} [de qué conferenciantes]_i C [_{TP} [las propuestas t_i]_k impresionarán_j-T [_{v*P} t_k [t_j]]]]]
 (Probing of DP to SPEC-T movement by the just inherited π -features)

However, this is not the derivation proposed by Gallego. In Gallego’s derivation, π -features of C fail to extract the PP from SPEC-T, which means that π -features and φ -features triggering V-to-T movement would not work in parallel. In other words [v-V]-to-T redefines TP as a phase only if φ -features and π -features of the C-T discontinuity do not probe in parallel; accordingly, phase sliding, a device postulated to account for the Spanish contrasts illustrated in (63-65) is at odds with parallel probing, which is necessary to account for English patterns of subject islands.

The most straightforward solution to this tension is that π -features surfacing on T in NSLs are a further instance of the C-T connection, and that therefore they are not provided by [V-v]-to-T movement. π -features and φ -features probe in parallel: φ -features trigger V-to-T movement

and π -features prototypically trigger movement to a projection of the C-area. Particularly, a type of π -features, topic-features, can be matched in SPEC-T in NSLs, as was illustrated above with preverbal subjects and CLLD elements. Thus, during the same derivational step, φ -features trigger V-to-T movement and π -features trigger movement of the whole DP ‘las propuestas de qué conferenciantes’ to SPEC-T.

(67)

Σ_1 : [_{CP} C [_{TP} [las propuestas de que conferenciantes]_i impresionarán_i-T [_{v*P} t_i t_j]]]
 (Parallel movement of V to T and of DP to SPEC-T)

Once SPEC-T has been created, the label of TP contains π -features that create opacity effects, as the Relativized Opacity Principle predicts. Thus, a further π -probe, say matrix v , cannot subextract a PP from SPEC-T.

In sum, in NSLs, a PP cannot be subextracted from a [DP [PP]] if this asymmetric object is located in SPEC- v^* ; some apparent counterexamples provided by Gallego have been taken into account, concluding that they involve psych-verbs or reduced relative clauses [DP PP], where the PP is located in a less nested layer than in the asymmetric object. Accordingly, v^*P creates opacity effects in non-NSLs and in NSLs, as expected under the intuition that movement does not eliminate previous effects or semantic relations. It has been noted that the hypothesis that [V- v] endows TP with topic-features introduces a timing problem: the PP cannot be extracted when the DP is spelled out in SPEC-T, but T would trigger movement of the DP to SPEC-T once SPEC-C had been created by merge of the extracted PP. For this reason, it has been proposed that π -features of T are not endowed from v , but offer a further instance that C and T constitute a derivational unit: φ -features trigger V-to-T movement and π -features trigger movement of a topic to SPEC-T in parallel; a further π -probe cannot subextract the PP from SPEC-T, since TP is a π -projection in NSLs. The next question is why SPEC-T is used to match π -features in NSLs, both by subjects and non-subjects.

5.3.2.2 Maximize Matching Effects Principle

Recall the above formulated Maximize Matching Effects Principle:

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

As argued in (5.1), one way to explore whether the MMEP plays any role in the syntactic computations of the C_{HL} is to find out whether two different features are matched in the same projection, creating a syncretic syntactic object. It is also important to determine, for any possible case of syncretism, which syntactic circumstance favors the application of a genuine principle of structural minimization to the detriment of the C-I legibility requirement of creating analytic syntactic objects and discontinuous syntactic objects, where different features are conveyed in different projections.

All the necessary premises have already been independently motivated to analyze the TP of NSLs as a case of syncretic syntactic pattern: in the head T, type φ -features are matched by the finite verb, and at SPEC-T, π -features such as topicality and referentiality are matched by a preverbal subject or a CLLD element. Two matching relations thus take place in a sole projection, opportunistically using SPEC-T in NSLs as an A'-position.

In this particular case of syntactic syncretism, the key is that the finite V is responsible for instantiating φ -features of T, which renders A-movement redundant. Thus, [V-v]-to-T movement does not directly provide T with π -features, but creates the syntactic circumstance that favor syntactic syncretism (two different features are matched in a sole projection) to the detriment of the general legibility condition that requires a strict order of type features by assigning to each feature a different projection.

5.3.3 When is SPEC-T created in NSLs?

It has often been claimed that SPEC-T must be universally filled, either by an overt DP (68.a), an overt expletive (68.b), or a covert or null expletive (68.c).

- (68)
- a. *(Three students) remained in the class
 - b. *(There) remained three students in the class
 - c. [\emptyset]_{expl} van quedar tres estudianta a la classe

If this were correct, then SPEC-T would be universally filled at overt syntax, contrary to the analysis developed above, according to which SPEC-T in NSLs is only created when π -features need to be matched. However, the existence of null expletives in NSLs (68.c) lacks empirical support, and seems to be no more than a prejudice derived from English patterns (68.a/b); it is not motivated by 'theory-internal reasons', but simply stipulated as a formal requirement that must be universally met, traditionally known as the Extended Projection Principle, currently viewed as a feature of T (the EPP feature) or a generalized feature of heads requiring an occurrence in its SPEC (the OCC feature).

The EPP-OCC feature is entirely dispensable: it is neither necessary to account for local A-movement or A'-movement, nor to account for intermediate steps. In any of these instances of movement of maximal categories, type features with a semantic content can be defined (φ -features for A-movement, π -features for A'-movement and defective φ/π -features for intermediate steps), avoiding the postulation of a feature that cannot be externally motivated by detecting its effects at the C-I system. Although the theory of grammar could quite easily dispense with merely positional features such as the EPP, the thesis that SPEC-T in NSLs is created only when it is necessary to match features other than φ -features would be further supported if null expletives can be independently rejected.

In this respect, a compelling observation noted by Rosselló (2000) is that the postulation of a null expletive to account for postverbal subjects or null subjects is, in Rosselló's words, "abusive", since "it does not predict the standard properties of the existing overt expletives: expletives in NSLs would by far exceed the well-known restrictions on lexical expletives,

which, aside from the case in which they are the correlate of a sentence (*it/it* in constructions like *it seems that.../il semble que...*), are only compatible with nominals in the 3rd person (the *zero* person, strictly speaking), whether they agree with their associate, as in the case of English *there* or German *es* (*there/es*, $V_{3sg/pl}$, $associate_{3sg/pl}$), or not (*it/il*, V_{3sg} , $associate_{3sg/pl}$)” (Rosselló 2000: 108). Thus, the alleged null expletives would violate the characteristic restrictions on overt expletives expressed in (69). This is adapted from Rosselló (2000: 109). For similar ideas and further arguments, see Rosselló (1986), Solà (1992), Alexiadou & Anagnostopoulou (1998) and Picallo (1998).

(69) *Null expletives:*

I.A Are not restricted to indefinite associates:

a Vindran els professors
 come.FUT the professors
 a'. *There would come the professors

I.B Do not license postverbal indefinite subjects with transitive and unergative predicates

a. *Ho saben candidats
 it.ACC-know candidates
 a'. *There know it candidates

b. *Badallaven vells
 yawned old people
 b'. *There know it old people

II. Are not restricted to unaccusative constructions

a. Ho saben els candidats
 it.ACC-know the candidates
 a'. *There know it the candidates

b. Badallaven els vells
 yawned the old people
 b'. *There know it the old people

III. Are compatible with lexical subjects in the 1st and 2nd person

a. Vindrem nosaltres
 come we
 a'. *There come we
 b. Vindreu voaltres
 come you(plur)
 b. *There come you

One could argue that, although SPEC-T is not created in overt syntax, it is created by covert movement, to check the structural case of the DP, in rather traditional terms. However, there is some reason to think that postverbal subjects do not covertly move to SPEC-T. Consider, for instance, the following English minimal pair, where a counting quantifier (CQP) ‘more than’ scopes over negation when it is nominative and under negation when it is accusative. This observation led Beghelli & Stowell (1997) to argue that Counting Quantifier Phrases (CQPs) take scope from ‘case’ positions, i.e., from A-positions (namely, SPEC-T and SPEC-V).

- (70) a. Peter doesn’t know more than two linguists
 The number of linguists Peter does not know is 2 or 1
 b. More than two linguists don’t know Peter
 The number of linguists that do not know Peter is 2, 3, ...

Nominative CQPs can occur preverbally and postverbally in Catalan, and significantly, when they are postverbal, they cannot scope over negation. Thus, postverbal nominative CQPs have the same scope as accusative CQPs.

- (71) a. En Pere no coneix *més de dos lingüistes* (accusative CQP)
 ART Pere neg know more than two linguists
 The number of linguists Pere does not know is 2 or 1
 b. No coneixen en Pere *més de dos lingüistes* (postverbal nominative CPQ)
 neg know ART Pere more than two linguists
 The number of linguists that do not know Peter is 2 or 1
 c. *Més de dos lingüistes* no coneixen en Pere (preverbal nominative CPQ)
 more than two linguists neg know ART Pere
 The number of linguists that do not know Peter is 2, 3, ...

This strongly suggests that there is no covert movement of these nominative quantifiers to SPEC-T, supporting the course developed throughout this chapter: A-movement to SPEC-T and V-to-T movement are alternative strategies of instantiating φ -features on T; if a finite lexical V instantiates them, remerge of a DP in SPEC-T is unnecessary and thus impossible unless other features (topicality/referentiality) are to be matched. No redundant operation takes place.

5.4 Other conundrums to be solved on the basis of structural minimization

5.4.1 Double agreement dialects

The observation that a finite lexical V leaves the vP in V2 Germanic languages such as Swedish or Dutch only when there is no overt complementizer has been considered to suggest that, when V leaves the vP, it always moves to C, and never to T. However, a very important conclusion drawn by Zwart (1993) is that, in some Germanic dialects that display

this complementarity effects between overt complementizers and V-movement, verbal inflection reflects that V is sometimes on C and sometimes on T, although it appears in a second position. Therefore, V2 patterns do not necessarily involve V-to-C movement, contrary to the standard analysis since den Besten's (1977/1981) seminal work.

In the so-called double agreement dialects, both complementizers and verbs are inflected for φ -features. For instance, West Flemish displays the following complementizer agreement paradigm, introduced in chapter 3 and borrowed from Haegeman 1990).

- | | | | | |
|------|----|---------------------|---------------|---------------------|
| (72) | a. | da-n-k | ik komen | |
| | | that-1SG-I | I come-1SG | |
| | b. | da- \emptyset -j | gie komt | |
| | | that-2SG-you | you come-2SG | |
| | c. | da-t-j | ij komt | |
| | | that-3SG-he | he come-3SG | |
| | d. | da- \emptyset -se | sij komt | [$\emptyset < t$] |
| | | that-3SG-she | she come-3SG | |
| | e. | da- \emptyset -me | wunder komen | [$\emptyset < n$] |
| | | that-1PL-we | we come-1PL | |
| | f. | da- \emptyset -j | gunder komt | [$\emptyset < t$] |
| | | that-2PL-you | you come-2PL | |
| | g. | da-n-ze | zunder komen | |
| | | that-3PL-they | they come-3PL | |

Zwart's observation (1993: 174, 179-180) is that "the verb in these dialects has verbal agreement in subject initial main clauses, and complementizer agreement in subject inversion constructions". Thus, in West Flemish, a 2SG verb adds the ending '- \emptyset -j', *kom-- \emptyset -j*, as the complementizer *da-*, *da- \emptyset -j*, when it is spelled out in C (in verb-subject inversion contexts), whereas it adds the ending *-t* when it is spelled out in T (with no verb-subject inversion) or when it appears in situ. Accordingly, verbal agreement transparently manifests whether or not V is fronted to C. The same pattern is attested in East Netherlandic and Brabantish, also borrowed from Zwart (1993).

- | | | | |
|------|----|---|-------------------|
| (73) | a. | Gie kom-t/*- \emptyset | West Flemish |
| | | you come 2SG _{V/C} | |
| | b. | Kom- \emptyset -j/*t-j-gie? | |
| | | come 2SG _{C/V} you you | |
| | | "Are you coming?" | |
| | c. | da- \emptyset -j gie kom-t | |
| | | that 2SG _C you you come 2PL _V | |
| | | "... that you are coming" | |
| (74) | a. | Wij speul-t/*-e | East Netherlandic |
| | | We play 1PL _{V/C} | |
| | | "We are playing" | |
| | b. | Waar speul-e/*-t wij? | |

- Where play 1PL_{C/V} we
 “Where do we play?”
- c. Speul-e/*Speul-t wij?
 Play 1PL_C / play 1PL_V we
 “Are you playing?”
- d. ...datte wij speult
 that-1plC we play 1PL_V
 “... that we are playing”
- (75) a. Gullie kom-t/*-de
 you come 2PL_{V/C}
 “You are coming”
- b. Wanneer kom-de/*-t gullie?
 When come 2PL_{C/V} you
 “When do you come?”
- c. dadde gullie kom-t
 that-2PL_C you come-2PL_V
 “... that you come”
- Brabantish

Thus, descriptively, a finite V does not target the C-T discontinuity when there is an overt complementizer (74.d), i.e., when C is [+embedded]; when there is no overt complementizer, i.e., when C is [-embedded], V moves to T if the first component is a subject (74.a) and to C when there is a π -feature to be matched in C, a total *wh*-feature, which requires V1 patterns (74.a), a partial *wh*-feature (74.b), or a topic/focus feature.

The best account I can figure for such a dynamic pattern is again in accordance with the idea that structural minimization matters for syntactic computations. But let us first consider a preliminary question.

V-movement to the C-T discontinuity does not instantiate ϕ -features, as V2 languages are non-NSLs; in fact, the number of person-number distinctions expressed by verbal inflection in these languages is quite low: for instance, Dutch verbal inflection displays 3/6 distinctive forms for present indicative, the same as French, another non-NSL, and Swedish verbal inflection displays only one.

(76) *Present indicative*

	<i>Dutch</i>	<i>French</i>	<i>Swedish</i>
1SG.	kus	embrass-e	kyss-er
2SG.	kus-t	embrass-es	kyss-er
3SG.	kus-t	embrass-e	kyss-er
1PL.	kuss-en	embrass-ons	kyss-er
2PL.	kuss-en	embrass-ez	kyss-er
3PL.	kuss-en	embrass-ent	kyss-er
	(3/6)	(3/6)	(1/6)

This is at least the general case, but there are some interesting complications. For instance, referential null subjects in Frisian are possible only in constructions that display complementizer agreement on C or on V (77), and null subjects in West Flemish are legitimate only when subject clitics are attached on C or on V (78). See Hoekstra & Marácz 1989 and Zwart 1993, from whom this evidence is borrowed. In these languages, the presence of C-agreement/clitics is thus a necessary condition for referential null subjects.

Frisian

- (77) a. Komst (do) jûn?
 come.2SG you tonight
 ‘Do you come tonight?’
- a’. ... datst (do) jûn komst?
 that.2SG you tonight come.2SG
 ‘... that you come tonight’
- b. komst *(er) jûn?
 Come.3SG he tonight
 ‘Is he coming tonight?’
- b’. ... dat *(er) jûn komt?
 that he tonight come.3SG
 ‘... that he comes tonight’

West Flemish

- (78) a. Goa-n-ze (zunder) werk een?
 Go-3PL-CL they work have
 ‘Are they going to have a job?’
- a’. Goa-n (*zunder) werk een?
 Go-3PL they work have
 ‘Are they going to have a job?’
- b. ... da-n-ze (zunder) goan werk een
 that-3PL-CL-3PL they go-3PL work have
 ‘... that they are going to have a job’
- b’. ...da-n (*zunder) goan werk een
 that-3PL they go-3PL work have
 ‘... that they are going to have a job’

This could be taken to suggest that C-agreement on C suffices to instantiate \varnothing -features, thereby allowing referential null subjects. However, according to Hoekstra & Marácz (1989),

Zurich German shows referential null subjects and no C-agreement (79) and, according to Zwart (1993), some Hollandic dialects show C-agreement and no referential null subjects (80). Thus, as Zwart (1993: 167) concludes, “there seems to be no significant correlation between overt complementizer and referential pro-drop in the Germanic dialects”.

- (79) a. ... dass (d/du) in Züri wohnsch Zurich German
 that you in Zurich live.2SG
 ‘... that you live in Zurich’
- b. ... öb (d/du) nach Züri chunnsch
 whether you to Zurich come.2SG
 ‘... whether you come to Zurich’
- (80) a. Komme *(ze)? South Hollandic
 come.PL they
 ‘Are they coming?’
- b. ... ovve *(ze) komme?
 whether they come.PL
 ‘... whether they are coming’

I do not have enough information to try account for these microvariation patterns in Germanic dialects, which seem of some interest for a better understanding of the mechanisms that license referential null subjects. Be that as it may, I shall assume, as a working hypothesis, that the trigger of V-movement in V2 languages is tense, not φ -features, and thus, the general distinction between V2 languages and NSLs is that, in NSLs, a finite V richly inflected for φ -features instantiates φ -features, while in V2 languages a finite V poorly inflected for φ -features does not instantiate φ -features (and thus A-movement is necessary in most V2 languages). This assumption, which is no more than an attempt to define the trigger of V-movement to the C-T discontinuity with some foundation, is not essential for the analysis of V2 effects we now turn to.

Let us take as a starting point the analysis proposed in chapter 4 regarding the C-Infl connection: [+clause typing] positions (C) and [-clause typing] positions (Infl/T) constitute a discontinuous unit split into two positions or poles, and Infl-like feature are discontinuous features in syntactic representations, with an occurrence in each pole. In this case, the Infl-like feature of tense is simultaneously merged into two positions, C and Infl/T. There are thus two occurrences of tense, or rather, two co-variations of the feature of tense.

- (81) C_[tense]...T/Infl_[tense]

A priori, it is conceivable that tense is instantiated in either C or in T, or even that, if tense is instantiated in one pole, it need not, and hence must not, be instantiated in the other pole. For instance, if tense is matched in C by a finite complementizer it needs not, and hence must not, be redundantly instantiated in T by a finite verb. This is what we see in V2 languages: when a complementizer is inserted in [+embedded] clauses, V does not leave the v-V discontinuity,

since the tense feature has been satisfactorily instantiated by complementizer insertion in C. When C is [-embedded] and displays features such as topic/focus or *wh*-features, no complementizer is inserted, since V2 languages do not have [-embedded] complementizers, and V must move to C to instantiate tense features. In other words, a type tense feature is alternatively instantiated by complementizer insertion or by V-movement in the so-called V2 languages.

An important question is: what prohibits V from matching tense on C and forces V to match tense on T when the constituent occurring in an initial position is a subject? First of all, there is no reason to think that a non-focalized or a non-topicalized subject moves to SPEC-C; such a subject should be in SPEC-T, matching ϕ -features. Thus, if tense is matched in the TP when there is no π -feature to be matched in the CP then all computations are executed in a sole projection thereby reducing the structural span: DP-movement to SPEC-T instantiates ϕ -features and V-to-T movement instantiates tense features. If tense were instantiated by V-to-C movement when the first constituent is a subject matching ϕ -features on SPEC-T, two different projections would be used, not taking advantage of the availability of the head T to minimize the span of structure.

Therefore, in V2 languages, tense takes advantage of its discontinuous syntactic nature to cooperate both with clause typing features/ π -features and ϕ -features. Tense is matched on C only if there are clause typing features or π -features to be matched on C: if a [+embedded] feature needs to be matched by a complementizer, then tense is also matched during the same operation and in the same position by a finite complementizer, rendering V-to-T movement unnecessary and hence impossible; if *wh*-features and topic/focus features are to be matched in a matrix C, then V raises to C in order to instantiate tense, since there is no complementizer for matrix clauses. Otherwise, if no clause typing feature or π -feature is to be instantiated in C, and the first constituent is a subject, tense is matched in T: adhering to the MMEP, the syntactic component opportunistically prefers to exploit a sole projection (TP) to initiating another one (CP) to match a sole feature that can be indistinctively matched both in C and in T, due to its discontinuous syntactic nature. Importantly, V2 effects do not entail general V-to-C movement, but cooperation between the discontinuous feature of tense and ϕ -features and clause typing features or π -features in order to avoid initiating a CP for a sole matching relation or moving a V when a complementizer instantiates tense.

It must be noted that Zwart's account for V-to-C movement differs significantly from the one proposed here. Zwart argues that subject agreement features are not inherent to C, but rather a morphological reflex of functional head movement, in his terms, of AgrS-to-C movement. According to such an implementation, AgrS-to-C is a necessary step in eliminating N features of AgrS. However, since N-features of AgrS are eliminated by A-movement, i.e., by moving a maximal category to SPEC-AgrS, it is not self-evident that AgrS-to-C is necessary to eliminate N-features of AgrS.

Zwart's solution is that a feature α may be 'automatically present in the projection of α ' or not, in his terms, a feature may be [+accessible] or [-accessible], and that languages vary in how they subspecify α as [\pm accessible]. In the case of the N-feature of AgrS, if it is inherently [+accessible], then check under sisterhood can take place, i.e., XP-movement to SPEC-AgrS

suffices to match N; but if AgrS is [-accessible], then “something has to happen to AgrS to make the N-feature of AgrS spread to the projection of AgrS (Zwart 1993: 178)”. The N-feature of AgrS, according to Zwart, can be made accessible by two different operations: (a) AgrS-to-C movement in verb-subject inversion and (b) V-to-AgrS movement when the first constituent in a matrix clause is a subject, in which case “clauses are not expanded up to the CP level” (Zwart 1993: 183).

Strictly speaking, a projection does not exist unless a feature projects, i.e., unless a feature is accessible. This is especially clear in bare phrase structure, where syntactic representations are made exclusively of terminals and labels, thereby satisfying the Inclusiveness Condition, and X' -levels have no status. Thus, there can be no projection without a projecting feature, or a [+accessible] feature; indeed, the idea that features may be present in syntax and nonetheless [-accessible] would not be easy to support on independent grounds.

The devices proposed to make AgrS [+accessible] are also suspicious: why should AgrS-to-C favor that N became accessible in the abandoned projection?, why should V-to-AgrS, an operation triggered by the V-feature of AgrS in Zwart’s analysis, favor that N-features became accessible? Note that the same problems would arise in developing the idea that AgrS-to-C movement is necessary to make π -features accessible. As far as I can see, the simplest account for V-to-C movement is that Infl-like features are generated not only in [-clause typing] positions but also in [+clause typing] positions, which has been our starting point.

Aside from these formal difficulties, there is a potential empirical problem. In matrix clauses when the first constituent is a subject AgrS-to-C can be banned by assuming that C is absent, but, if I have not missed any crucial point of Zwart’s 1993 analysis, nothing would ban AgrS from being made accessible by V-to-AgrS movement in the case where the subject does not occur in an initial position of matrix clauses, which would predict that V occurs in a non-second position.

Note that it is not necessary for C to be absent in matrix clauses where the initial position is occupied by a subject; again, I would not know how to motivate this claim. What is necessary is that V-to-C does not occur in matrix clauses when no π -feature is matched in C, and it is undecidable whether this entails that C is absent or present when it has a default value.

Clearly, not all languages are V2: for instance, in Catalan, the presence of an overt complementizer does not ban V-to-T movement (82.a) and V-to-C movement is not required when some constituent undergoes focus fronting (82.b). Therefore, a conclusion that seems unavoidable is that V-movement does not universally cooperate with clause typing or π -features: in some languages, V-to-T movement is necessary although a complementizer is inserted and V-to-C is illegitimate although clause typing features and π -features are matched on C. The issue of whether a language is parameterized as V2 or non-V2, or whether tense cooperates with π -features or clause typing, seems to me non-derivable from other factors, such as morphological factors.

- (82) a. Diu que la Maria va sovint a la piscina
 say.3sg that art Maria go often to the swimming pool
 ‘He/she says that Maria often goes to the swimming pool’
- b. A LA PISCINA la Maria no va sovint
 to the swimming pool art Maria not go often
 ‘Maria does not often go TO THE SWIMMING POOL’

The MMEP predicts that, in languages where V can match tense both in C and in T, the occurrence of T is chosen when no other feature is matched in C, thereby avoiding the activation of CP for a sole projection, and hence reducing the span of structure used to carry out the necessary computations. Consequently, the MMEP is contradicted only if, in a language where V can match tense both in C and in T, V matches tense in C when no clause typing feature or π -feature is matched in C. Accordingly, the pattern provided in (82) does not contradict the MMEP: V does not raise to C, independently of whether clause typing or π -features are matched in the CP. The Catalan pattern (82) only shows that it is not a property of UG that V-movement cooperates with clause typing or π -features. Let’s now turn to address whether a further pattern violates the MMEP.

In old Irish, according to Carney, Pyatt & Harley (1994), verbal inflection takes two mutually excluding forms, the absolute form and the conjunct form.

- | | | | |
|------|-----------------|-----------------|---------------|
| (83) | <i>Absolute</i> | <i>Conjunct</i> | |
| | berid | -beir | ‘he carries’ |
| | berait | -berat | ‘they carry’ |
| | marbfa | -marbub | ‘I will kill’ |
| | midimmir | -midemmar | ‘we judge’ |

In Carney, Pyatt and Harley’s terms, “[T]he absolute form is used when the verbal root is in absolute first position in the sentence, that is when the inflected verb is not preceded by any conjunct particles, preverbs or pronouns”, where conjunct particles are negative markers, question markers and other clause typing conjunctions (84). All these preverbal particles are taken to occur in C. When V is preceded by any category, for instance a negative marker (85) or a fronted maximal category in some poetic texts, “where strict VSO order is not obligatory”, the conjunct form is used (86).

- (84) Berid in fer in claideb (Absolute)
 Carries-3s-abs the man the sword
 ‘The man carries the sword’
- (85) Ní beir/*beirid in fer in claideb (Conjunct)
 Neg carries-3s-conj/*abs the man the sword
 ‘The man does not carry the sword’
- (86) ... srethaib sluag soí Crimthan Coscrach cing cét catha, ...
 ... with lines of hosts won Crimthan victorious hero hundred battles

‘With lines of hosts, Crimthan the victorious hero, won a hundred battles’
(absolute: ***soid**)

Thus, verbal inflection in old Irish reflects whether V is in an absolute first position or not, as verbal inflection in double agreement dialects reflects whether V is on C or not. But there are two crucial differences between old Irish and V2 languages. First, C must always be filled in old Irish by a verb in absolute form or by other categories, but not in V2, where no category appears on C when the first constituent is a subject. And second, V-to-C in V2 languages alternates only with overt complementizers, but in old Irish, V-to-C alternates with any kind of categories occurring in C (focus/topic elements, question markers, negative markers, object clitics, etc).

It is conceptually clear that a language would violate the MMEP in discontinuous patterns if it manifested the following two properties:

- (a) V raises to C when π -features are matched in C, and
- (b) V raises to C when no π -feature is matched in C, initiating the CP for a sole computation instead of taking advantage of TP, where other features are matched

Catalan does not violate the MMEP in discontinuous patterns, since it does not manifest (a); double agreement dialects do not violate the MMEP, since they do not manifest (b); and finally, neither does old Irish violate the MMEP, because, although it satisfies (b), it does not satisfy (a): V does not raise to C when π -features are matched in C.

The difference between Catalan and old Irish is that, in old Irish, C must always be filled, by a question marker, a negative marker, or by V-movement. This indicates that V-to-C movement in old Irish, and not in V2 languages, is necessary to instantiate clause typing features, thereby defining a clause as an assertion, in opposition to a denial, a query, etc, and hence, a category must instantiate clause typing features although they have an unmarked value.

Summing up, the path to explore whether the MMEP plays any role in the computations of the syntactic component is to find out whether syncretic syntactic patterns exist and determine the special syntactic circumstance that favors a structural minimization principle to the detriment of the C-I legibility requirement of creating analytic syntactic patterns and discontinuous syntactic patterns. The first pattern taken under consideration has been the TP in NSLs, where both ϕ -features and π -features (topic/referentiality features) are matched. In this case, the syntactic circumstance favoring syntactic syncretism is that ϕ -features are matched in the head T, leaving SPEC-T available for further computations. The second pattern is provided by double agreement dialects, where the head T is opportunistically used to match tense when no π -feature is matched in matrix C, thus avoiding the activation of the CP to match a feature that can be indistinctively matched in C or in T, due to its discontinuous syntactic nature. In this case, the syntactic circumstance favoring syntactic syncretism to the detriment of a strict ordering of type features (with a one-to-one relation between type features and heads of the syntactic hierarchy of functional features) is the

discontinuous syntactic nature of Infl-features, particularly, of tense: in those languages where tense triggers V-to-C movement when π -features are matched in C, syntax prefers to match tense in T when the subject occurs in the initial position of the clause, i.e., when the CP is not active (no π -feature is matched) and TP is independently active (since the subject instantiates φ -features), rather than activating the CP for a sole computation. Indirectly, the view that C and T are at some level a unit split into two poles is further supported by the dynamic cooperation of the discontinuous feature of tense in double agreement dialects and more generally in V2 languages, where verbal inflection is silent on whether V moves to C.

5.4.2 From the Vacuous Movement Hypothesis to syntactic syncretism

Force features of C such as interrogative and affective features require auxiliary insertion and auxiliary-subject inversion in English. This indicates that fronted interrogatives and affective categories have undergone movement to a SPEC-C position and that the auxiliary occurs in a C-head position, preceding the subject.

- (87)
- a. What did you read?
 - b. Did you come yesterday?
 - c. With no jobs would Bill be happy
 - d. With few jobs would Bill be happy

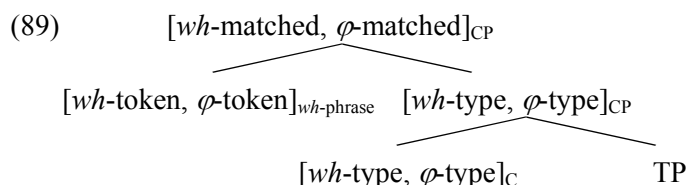
However, as is well-known, there is a clear asymmetry between non-nominative *wh*-phrases and nominative *wh*-phrases: the former require auxiliary insertion and auxiliary-subject inversion (88.a, b), whereas the latter neither require nor allow auxiliary insertion (88.c), unless the auxiliary conveys emphatical affirmation (87.d).

- (88)
- a. Where *(do) you live?
 - b. What *(did) you say?
 - c. Who (*did) came yesterday?
 - d. Who DID come yesterday?

The absence of *do*-insertion with nominative *wh*-phrases, as Chomsky (1986: 48) observes, is consistent with the idea that *wh*-movement to SPEC-C does not take place in the case of nominative *wh*-phrases. *Do*-support indicates that movement of a non-nominative *wh*-phrase or an affective phrase has taken place in matrix clauses, but its absence raises the possibility that nominative *wh*-phrases do not target SPEC-C, but SPEC-T.

But this is not the only conceivable choice. Suppose that, in English, a finiteness feature must be matched on C when an interrogative or an affective feature is active on C. Such a finiteness feature may be understood as a conglomerate of tense and φ -features. This finiteness feature of C can be matched by an inflected V (an English auxiliary, but not an English lexical verb) or by a nominative *wh*-phrase (in point of fact, the crucial difference between nominative *wh*-phrases and non-nominative *wh*-phrases is that the former bear not only a token of *wh*-features but a token of φ -features associated with the type of φ -features of the C-T discontinuity). Accordingly, if the bearer of *wh*-features is also the bearer of φ -features, a type of finiteness features, *do*-insertion is unnecessary to match finiteness, and

hence impossible. This would mean that both φ -features and π -features are syncretically matched in SPEC-C, which becomes a mixed A/A'-position, following the MMEP. Non-nominative *wh*-phrases do not have a suitable token for φ -features of the C-T discontinuity, and hence *do*-support is necessary to instantiate finiteness. A similar analysis is developed by Pesetsky & Torrego (2001, 2004a, b), putting aside some important differences in technical implementation.



Thus the question is whether the absence of *do*-insertion indicates that the nominative *wh*-phrase is in SPEC-T or that the nominative *wh*-phrase matches both finiteness and *wh*-features in C (rendering *do*-insertion unnecessary).

To the best of my knowledge, two arguments have been given for the idea that a nominative *wh*-phrase is in SPEC-C, and not in SPEC-T, in overt syntax (see Boeckx 2003, note 8 of his chapter 3).

Based on the observation that an interrogative in situ cannot combine with “the hell” ((90.a) vs. (90.b)), Pesetsky has argued that the nominative interrogative in (90.c) must be in SPEC-C (see Pesetsky & Torrego 2001, note 9).

- (90)
- a. What the hell did you see?
 - b. What did you give to who (*the hell)?
 - c. Who the hell bought the book?

In the degraded (90.b), the interrogative [*wh* + “the hell”] is arguably in its base positions; in the well-formed (90.a), *what the hell* is in SPEC-C, as auxiliary-subject inversion indicates. Thus, as Pesetsky argues, it seems that [*wh* + “the hell”] must move to SPEC-C, and it cannot be spelled out in its base position. But, since SPEC-T is not the base-position for nominative subjects, the data given in (90) is not relevant in deciding whether the nominative [*wh* + “the hell”] is in SPEC-C or in SPEC-T, merely reflecting that it is not in the base position SPEC-v*. Note that the patterns of subject islands discussed by Chomsky provide a strong argument (among many others) for the widely accepted idea that agentive subjects are generated inside the vP, and not in SPEC-T.

The second argument is based on the generally accepted assumption that sluicing is TP-deletion (Lasnik 1999, Merchant 2001). Accordingly, since nominative *wh*-phrases can appear in sluicing contexts, they should be in SPEC-C, and not in SPEC-T.

- (91)
- a. Someone left
 - b. I wonder [_{CP} who [_{TP} ~~left~~]]

But the *wh*-phrase that appears in sluicing contexts is always given focal stress, which might force movement of the *wh*-phrase to SPEC-C independently of *wh*-features. Accordingly, it may be that a nominative *wh*-phrase is spelled out in SPEC-T in matrix questions and in SPEC-C in sluicing contexts.

Therefore, there seems to be no compelling reason to think that a nominative *wh*-phrase moves to SPEC-C in matrix questions.

A different answer is provided by Chomsky (1986: 48-54), in his elaboration of an observation from Chung & McCloskey (1983). If a nominative *wh*-phrase remained in embedded SPEC-T, embedded SPEC-C would be available and “*wh*-islands effects will be removed for embedded *wh*-subjects (p. 48)”, as in (92.a).

- (92) a. ?What do you wonder [_{CP} who saw t]?
b. *What did she wonder [_{CP} where John put t]?

However, it seems that *wh*-island effects still emerge in (92.a), although they are weaker than those induced by a non-nominative *wh*-phrase (92.b). Chomsky assigns no question mark to (92.a), but admits “the persistence of weak island effects even with *wh*-subjects, as in [(92.a)]”.

A similar contrast is obtained with subject relativization and non-subject relativization (from Chung & McCloskey 1983): extraction of a *wh*-phrase from a subject relative yields much more acceptable outputs than extraction from a non-subjective relative, an unexpected contrast if both relatives were in SPEC-C.

- (93) a. That’s one trick that I’ve known a lot of people who’ve been taken in by
a’. *That’s one trick that I know a lot of people who no one will admit have been taken in by
b. Isn’t that the song that Paul and Stevie were the only ones who wanted to record?
b’. *Isn’t that the song that Paul and Stevie were the only ones who George said could record?

Chomsky provides both conceptual and empirical arguments to think that, although nominative *wh*-phrases stay in SPEC-T in overt syntax, they covertly move to SPEC-C.

The main empirical argument is primarily concerned with adjunct *wh*-movement. The contrast illustrated in (94) was accounted for by appealing to the ECP: (94.a) is an acceptable output because the trace of the non-adjunct ‘what’ at LF does not require proper government, whereas the trace of the adjunct ‘how’ in (94.b) must be properly governed at LF.

- (94) a. Who knows how John did what?
b. *Who knows what John did how?

Thus, it is assumed that such a contrast must be accounted for in terms of covert movement of (in situ) *wh*-phrases. Consider the acceptability contrast illustrated in (95).

- (95) a. ?What do you wonder [_{CP} who saw t]?
b. *How do you wonder [_{CP} who fixed the car t]?

As Chomsky argues (1986: 49), “[I]f *how* were permitted to move through the embedded pre-IP position [i.e., SPEC-C] while *who* remained in situ [i.e., at SPEC-T] at LF, then this sentence [(95.b)] should be perfectly grammatical, with *t* properly governed by the trace of *how* in the specifier position of the embedded CP”. Accordingly, the nominative ‘who’ would covertly move from SPEC-T to SPEC-C, thereby erasing the intermediate trace of ‘what’ in (95.a) and of ‘how’ in (95.b); because the base-generated trace of ‘how’ must be properly governed, but not the base-generated trace of ‘what’, only the erasure of the former yields a degraded output.

Chomsky (1986) conceives the absence of *wh*-movement to SPEC-C in overt syntax in the case of nominative *wh*-phrases as an effect of a Prohibition Against Vacuous Movement (i.e., the Vacuous Movement Hypothesis). The enunciation proposed is mine.

- (96) *Prohibition Against Vacuous Movement*

A transformation can apply on a sequence σ of syntactic objects only if it yields a sequence σ' and $\sigma \neq \sigma'$

Such a prohibition would operate in superficial structure (or narrow syntax), but not in covert syntax: if the nominative ‘who’ did not move covertly to SPEC-C, (95.b) should be as acceptable as (95.a), since the intermediate trace of ‘how’ would not be deleted and could properly bind the trace in the base position.

Empirical arguments seem to agree with conceptual arguments: if all *wh*-phrases, independently of whether or not they are nominative, universally move to SPEC-C positions for legibility conditions such as selection, scope and absorption, languages would not present deep differences as to the position in which *wh*-phrases are interpreted, but only as to the position in which they are spelled out.

However the simplicity and the empirical and conceptual adequacy of such a proposal are merely apparent. If movement obeys the Prohibition Against Vacuous Movement formulated above, then the *wh*-feature of C must be instantiated covertly when the token of *wh*-features is a subject and overtly when it is a non-subject in the same language (say, English). This leads to a rather suspicious situation: why should a nominative interrogative be able to stay in an A-position in overt syntax, and not an object interrogative?, why should a type *wh*-feature be allowed to be instantiated covertly if the bearer of *wh*-features is a subject, and not if it is a non-subject? In other words, why should the timing of matching *wh*-features depend on the grammatical function of the bearer of *wh*-features?

The vacuous movement account presupposes that the function of movement is to alter the internal structure of a syntactic object by means of transformations, and not to yield new

semantic relations. However, as Chomsky himself has recently argued, there is good reason to think that movement serves the purpose of associating a category with different meanings by relocating it to different positions, a path that has been followed and developed in this study. Thus, the vacuity of a movement operation is not to be understood in bare formal terms as in (96), but as in (96’).

(96’) *Prohibition Against Vacuous Movement*

A movement operation is vacuous if and only if it does not yield a satisfactory type-token relation, where a type feature is instantiated and a token feature is hosted in a suitable position

It must also be noted that several well-founded internal merge applications that violate the Prohibition Against Vacuous Movement are postulated in current generative transformational syntactic analysis; for instance, A-movement of an external argument from SPEC-v* to SPEC-T arguably occurs in English for matching requirements, but the output sequence is identical to the input sequence: <external argument, V>. A further prediction of the the Prohibition Against Vacuous Movement is that, in overt syntax, an agentive subject moves to SPEC-T only if there is an explicit adverb, as in “John often kisses Mary”, where the output sequence (<John, often, kisses>) is different from the input sequence (<often, John, kisses>). Moreover, an agentive *wh*-phrase would be spelled out in SPEC-v*, like a non-interrogative external argument, whereas a non-agentive subject (interrogative or not), which is base-generated in V-COMPL, should overtly move to SPEC-T, yielding the new sequence <non-agentive subject, V>. Since all these hypothetical variety of effects lack empirical support, I conclude that internal merge, as external merge, is triggered by matching necessities, and does not care about sequential alterations in itself. See den Dikken (2006) for a different recent revision of the Prohibition against Vacuous Movement.

Therefore, the vacuous movement account proposed by Chomsky (1986) is problematic in – at least– three aspects:

- (i) it does not explain why the timing of matching *wh*-features depends on the grammatical function of the bearer of *wh*-features (why, in the same language, should a nominative interrogative be able to stay in an A-position in overt syntax, and not an object interrogative?, why, in the same language, should a type *wh*-feature be allowed to be instantiated covertly if the bearer of *wh*-features is a subject, and not if it is a non-subject?),
- (ii) it fails to grasp that the role of internal merge (as well as that of external merge) is to provide semantic richness to the interfaces, and not to formally alterate the internal structure of a syntactic object,
- (iii) it makes empirical predictions that seem to be inaccurate.

An interesting alternative emerges if the MMEP is taken into account and the belief that *wh*-features are matched by necessity in C is abandoned. The special circumstance of matrix clauses with a nominative interrogative is that the token of *wh*-features and φ -features of the

C-T discontinuity is the same category, the nominative interrogative. This favors that both type φ -features and type *wh*-features are matched in the same SPEC-position, and not by successive applications of internal merge, minimizing the span of structure where matching relations are carried out, and hence adhering to the MMEP.

Note that, according to the above reasoning, φ -features and *wh*-features could be both instantiated in SPEC-C and in SPEC-T: in both cases, two features would be matched in a sole projection, as the MMEP requires. However, the alleviation of *wh*-island effects seems to suggest that in English SPEC-T is the position in which both features are syncretically matched, not SPEC-C. Reconsider the definition of MMEP.

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

Throughout this chapter, a structural span has been measured by the number of projections where different type features are located and matched: when two type features are located and matched in a sole head (and hence, in a sole projection) a smaller span of structure is used than when two type heads are located and matched in two different heads. In other words, the span of structure of a syncretic pattern is smaller than the span of structure of an analytic pattern.

But this is not the only possible way of measuring a span of structure. Note that TP is smaller than CP in that TP is a constituent of CP and CP is not a constituent of TP. Accordingly, if TP is chosen instead of CP, the span of structure to match φ -features and π -features of the C-T discontinuity is smaller than if CP was chosen, satisfying the MMEP in a double sense: it is satisfied in an ‘internal’ sense (two matchings are carried out in one projection) and in an ‘external’ sense (in the discontinuous unit CP-TP, the smallest projection is chosen instead of the largest one).

English nominative *wh*-phrases do not offer the only possible instance of mixed A’-/A-position. Similar patterns are obtained in Standard Dutch: when the first constituent is a 2SG topic subject, V displays Infl-agreement (-*t*), whereas when it is a non-subject topic it displays C-agreement (- \emptyset). This indicates that topic constituents are in SPEC-T when they are subjects, and in SPEC-C when they are non-subjects. If the first constituent is a 2SG focused subject, V also displays Infl-agreement (Jan-Wouter Zwart, p.c.), as illustrated in (97.b). Consequently, there are no instances where the first constituent is a 2SG subject and it does not trigger Infl-agreement.

(97) a. [Jij]_{topic} gaat/*ga (Zwart 1993: 183, note 8)
 ‘You go’

a’. [Dat boek]_{topic} *kent/ken jij (Zwart 1993: 245)
 that book know you
 ‘You know that book’

b. JIJ gaat/*ga

Thus, both topic/focus-features and *wh*-feature may be matched in SPEC-T in the special circumstance that the bearer of token π -features is the same category as the bearer of token φ -features, thereby neutralizing the distinction between A'-/A-position and minimizing, in an internal and external sense, the span of structure where matching relations take place. See Miyagawa (2005) on the idea that topic features are matched in SPEC-T in Japanese, taking advantage of the non-prominence of φ -features in the syntax of this language.

Before closing this section, I must once again consider *wh*-island effects. Under Chomsky's (1986) approach, it was quite natural for (98.b) to be more degraded than (98.a) if the nominative *wh*-phrase remained in SPEC-T, but it was not clear why (98.a) was not completely acceptable (recall that the trace of 'what' does not need to be properly governed).

- (98) a. ?What do you wonder [_{CP} who saw t]?
 b. *What did she wonder [_{CP} where John put t]?

In chapter 3, I argued that *wh*-island effects emerge from the inability of C to trigger parallel movement of two *wh*-phrases to SPEC-C positions, not from the deletion of an intermediate copy, somewhat straying from the classical analysis. Thus, the crucial difference between English and Bulgarian is that in the latter C has the ability to trigger multiple *wh*-movement, thereby allowing multiple *wh*-fronting (Richards 1997, 1999, Boskovic 1999) and lacking *wh*-island effects (Richards 1997). More precisely, I claimed that an embedded C in Bulgarian can have both a *wh*-feature (hosting a *wh*-phrase which gets frozen in place) and a defective π -feature (which suffices to trigger movement of a *wh*-phrase but fails to offer it a suitable position, and thus it is available to a further probe), whereas in English an embedded C cannot have both features: if it has a *wh*-feature (i.e., if it is an embedded interrogative), then it lacks a defective π -feature, which yields *wh*-island effects.

The contrast between nominative and non-nominative embedded *wh*-phrases forces us to refine this proposal, for in both cases the embedded interrogative C fails to readily trigger movement of two *wh*-phrases (neither (98.a) nor (98.b) is fully acceptable), although the degree of acceptability increases when each *wh*-phrase can be attracted to a different projection, that is, if one of the two *wh*-phrases is nominative.

This picture is more compatible with the hypothesis that English embedded C can have both *wh*-features and defective π -features –as in Bulgarian–, though they cannot easily operate during the same derivational step; the resulting output improves when *wh*-features are matched along with φ -features in SPEC-T by a nominative *wh*-phrase, which suggests that *wh*-features and defective π -features can operate more easily in the same derivation when they trigger movement to different projections (99.a). When both defective π -features and *wh*-features are matched in C, the output is more degraded (99.b).

- (99) a. ?[_{CP} what C_[+emb; def- π] [_{TP} who T_[wh; φ] [_{VP} ~~who~~ saw ~~what~~]]]
 b. *[_{CP} what [_{CP} where C_[+emb; wh; def- π] [_{TP} John T_[φ] [_{VP} ~~John~~ put ~~what~~ ~~where~~]]]]

This analysis does not account for the contrast between extraction of an argument *wh*-phrase and an adjunct *wh*-phrase over a nominative *wh*-phrase, repeated in (100).

- (100) a. ?What do you wonder [_{CP} who saw t]?
b. *How do you wonder [_{CP} who fixed the car t]?

However, that the nominative *wh*-phrase moves covertly to SPEC-C, thereby deleting the intermediate trace of ‘what’ in (100.b), is not a necessary condition to account for the robust observation that adjunct extraction from embedded interrogatives yields worse results than extraction of arguments. See, for instance, Cinque (1990) and Rizzi (1990) for the proposal that the existence of referential indices is what alleviates island effects in the case of argument extraction.

In sum, on the basis of the alleviation of *wh*-island effects, it has been proposed that nominative *wh*-phrases match both φ -features and π -features in SPEC-T, which becomes a mixed A’-/A-position. In this case, the special circumstance favoring syntactic syncretism to the detriment of the semantic requirement of bringing semantically devoted positions into syntactic representations is that the bearer of token *wh*-features is the same as the bearer of φ -features, a nominative interrogative. *Wh*-island effects always emerge, since *wh*-features and defective π -features cannot readily probe during the same derivational step in English; however, a *wh*-phrase undergoing long-movement can target SPEC-C more easily if the type *wh*-feature is syncretically matched in SPEC-T by a nominative *wh*-phrase. It has been argued that in Dutch π -features such as topic and focus are also syncretically matched in SPEC-T along with φ -features.

5.5 Contraction

In this section I shall tentatively address a further question related to structural minimization: can C and T contract in some special syntactic circumstance, and thereby avoid the essential [\pm clause typing] distinction of the C-T discontinuity?

Consider the following standard instances of the so-called comp-trace effect, where extraction of a nominative *wh*-phrase (interrogative or relative) from an assertive embedded clause is highly degraded unless the embedded complementizer ‘that’ is absent. If the extracted *wh*-phrase is non-nominative, the insertion of ‘that’ is optional.

- (101) a. Who do you think (*that) came?
a’. Who do you think (that) Peter saw?
a’’. Where do you think (that) Peter lives?

b. The car which I think (*that) was stolen yesterday
b’. The car which I think (that) someone stole yesterday
b’’. The place where I think (that) Peter lives

There is an interesting empirical complication: if an adverbial phrase like ‘next year’ or ‘during an operation’ appears interpolated between the complementizer and the embedded

subject, in a topical peripheral position, then the overt presence of ‘that’ yields a more acceptable output, as originally observed by Bresnan (1977: 194).

- (102) a. An amendment which they say (*that) will be law next year
 b. An amendment which they say that, next year, will be law
- (103) a. Which doctor did you tell me (*that) had had a heart attack during an operation?
 b. Which doctor did you tell me that, during an operation, had had a heart attack?

Rizzi (1997) accounts for the contrast illustrated in (102) and (103) by combining an economy principle that requires structural minimization and the ECP. Force and finiteness can be “specified on distinct heads of the complementizer system or can be specified syncretically, on a single head” (Rizzi 1997: 311). The topic-focus field is assumed to be present in the structure only when it is active, i.e., only when topical or focal categories are to be hosted. When the topic-field is present, it is sandwiched between the Force head and the Fin head, which “must terminate the C-system upward and downward (p. 288)”. The economy principle that avoids structure would require them to be specified syncretically when the topic-focus field is absent, but, in (102) and (103), where an adverbial is preposed, force and finiteness are two different heads, as in the representation (104), where ‘that’ occurs in Force and a null category in Fin.

- (104) ..._{[ForceP} that-Force _{[TopicP} next year Topic _{[FinP} ∅-Fin _[John will win the prize]]]

According to Rizzi, what licenses subextraction when ‘that’ is omitted is the ability of a null Fin to properly govern the trace of a subject; since, for reasons of economy, null Fin is only present when the split of force-finiteness is required, a nominative *wh*-phrase can be extracted when ‘that’ is overt only if a preposed category requires Force-Fin to split.

Here, I shall follow Rizzi’s intuition that structural minimization is an important factor in understanding comp-trace effects (101-103, 104), but I shall investigate whether a more basic explanation is possible, without assuming the ECP.

The main observation concerning (102) and (103) is that, when a nominative *wh*-phrase is extracted from an embedded CP, the insertion of ‘that’ is necessary only if a π -feature has been satisfactorily matched by a topic phrase; if this is not the case, ‘that’-insertion is not optional but impossible. Such impossibility is compatible with the hypothesis that the complementizer ‘that’ cannot be inserted simply because there is no C-position in the structure. Let us examine whether there is any reason to adopt such a view.

In section (5.4.2) it has been argued that π -features (*wh*-features, and topic/focus-features) are matched along with φ -features in SPEC-T when the token of both type features is the same category. If, otherwise, the tokens for *wh*-features and φ -features are in different categories, π -features and φ -features are matched in different projections, yielding analytic syntactic objects or cartographic effects.

If defective π -features triggering intermediate positions of long A'-movement behave as non-defective π -features, which is at least the neutral hypothesis, then a very simple account for the standard cases of comp-trace effects (101) is possible. A nominative *wh*-phrase moves to embedded SPEC-T attracted by defective π -features and φ -features; it satisfactorily matches φ -features, but defective π -features cannot offer a suitable position for it, and thus it remains accessible to a further π -probe. If no π -feature is matched in CP, which is consequently freed of π -features, and the label of TP manifests both φ -features and π -features, then the semantic distinction between [+clause typing] positions and [-clause typing] positions is neutralized, and accordingly, C and T should contract, making *that*-insertion impossible. In the case that a non-nominative is extracted from an embedded CP, the π -features that trigger intermediate positions of A'-movement operate in CP, and thus such a projection is active in syntax, and cannot contract, whereby 'that' can be inserted.

C and T are two elements of the same unit each specified for a different polarity of the same feature ([\pm clause typing]); thus what licenses the C-T distinction is that C becomes a [+clause typing] position and T a [-clause typing] position. If it turns out that, due to a certain syntactic circumstance that favors structural minimization, both π -features and non- π -features are matched in T, then there is no longer any reason to maintain the C-T distinction in a syntactic representation. Once π -features are matched in TP, CP and TP contract, and there is no structural space to insert a complementizer.

Interestingly, there are some cases in which 'that'-insertion is neither possible nor impossible, but necessary: when the embedded CP contains an active feature such as an affective feature (105a, b) or a topic feature (105.c), instead of a defective π -feature.

- (105) a. She swore/insisted/thought *(that) never in her life would she accept this solution (Grimshaw 1997)
 b. He said *(that) under no circumstances would he do it. (Rizzi & Roberts 1989)
 c. I think *(that) next year, John will win the prize (Rizzi 1997)

This indicates that an embedded complementizer must be omitted when defective π -features are matched in TP, and it can be inserted when defective π -features are matched in CP. In the former case, CP and TP contract, because TP contains not only φ -features but π -features as well. But when a non-defective π -feature is matched in CP, then 'that' is necessary. Thus, according to the proposed analysis, in (102.b) and (103.b) 'that' insertion is not allowed because C and T have contracted; however 'that' must be inserted in (102.a) and in (103.a) because CP and TP do not contract and a topic feature matched in the CP requires 'that' insertion, as in (105).

Independently of the merits of this proposal, based on the hypothesis that two poles of the C-T discontinuity may contract when both π -features and φ -features are computed in TP, there always remain two important questions with regard to comp-trace effects: why do NSLs not display comp-trace effects (Perlmutter 1971)?, and why do certain English dialects lack comp-trace effects, though they are virtually identical to those that manifest comp-trace effects (Sobin 1987)? This may suggest that the circumstance whereby both π -features and φ -

features are matched in T does not universally force C-T contraction. In the end, comp-trace effects emerge rather randomly.

5.6 Conclusion

This chapter has explored whether the Maximize Matching Effects Principle, a genuine principle of structural minimization, plays any role in the computations of the C_{HL} .

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

The path of exploring the MMEP has been to determine whether syncretic syntactic patterns exist and define the special syntactic circumstance that favors a structural minimization principle to the detriment of the C-I legibility requirement of creating analytic syntactic patterns and discontinuous syntactic patterns.

Thus, whereas analytic and discontinuous syntactic patterns (or cartographic effects) have been attributed to the C-I requirement of bringing semantic distinctions into syntactic representations, syncretic syntactic patterns (or anti-cartographic effects) derive from a principle of structural minimization.

The discussion of the nature of V-movement and verbal inflection has concluded that:

- I. Both movement of maximal categories and minimal categories take place at narrow syntax; thus V-movement to T/C and DP-movement to T take place at narrow syntax.
- II. Inflected verbs contain not only inherent features (θ -features and Aktionsart features), but also non-inherent functional features (tense, φ -features, mood, aspect, etc.). Functional features attached on V convert V into a multicategorical word, which does not mean that they are uninterpretable on V or that they must be deleted.
- III. The position of V is determined by the non-inherent functional features attached on V; particularly, V leaves the vP moving up to C or to T if it is finite.
- IV. A language L is a NSL if and only if verbal inflection is rich in L ; L is a NAL if and only if L displays free discourse indexing mechanisms, independently of verbal inflection.

Three possible instances of syncretic syntactic patterns have been discussed:

- I. In NSLs, TP is a syncretic syntactic pattern: in the head T, type φ -features are matched by the finite verb, and at SPEC-T, π -features such as topicality

and referentiality are matched by a preverbal subject or a CLLD element. Thus, two matching relations take place in a sole projection, thereby opportunistically using SPEC-T as an A'-position.

The syntactic circumstance that favors syncretism in the TP of NSLs is that the finite V satisfactorily instantiates φ -features on T, which makes A-movement to SPEC-T redundant.

- II. In V2 languages, the type feature of tense cooperates with both clause typing/ π -features and φ -features: (a) if C is [+embedded], a complementizer is inserted in order to match clause typing and tense on C, which renders V-movement unnecessary; (b) if C is [-embedded], V-movement is necessary to match tense.

When a is active on C in a [-embedded] clause, tense cooperates with such a π -feature by triggering V-to-C movement. When the first constituent of a [-embedded] clause is a non-topicalized/non-focalized subject, no π -feature is active on C, and the syntactic component opportunistically makes use of TP to minimize the structural span of structure where matching relations take place: both φ -features and tense are matched in the TP, which avoids initiating the CP for a sole computation. Whether or not V is on C is reflected in verbal inflectional morphology in double agreement dialects.

The syntactic circumstance that favors this case of syntactic syncretism is that of the discontinuous syntactic nature of tense, and more particularly, the fact that in the so-called V2 languages tense cooperates both with φ -features and π -features. In NSLs or in old Irish, this circumstance is not given, as properties (a) and (b) expressed in (3.2.0) are not given.

- III. On the basis of the alleviation of *wh*-island effects, it has been proposed that, in English, nominative *wh*-phrases match both φ -features and π -features in SPEC-T, which becomes a mixed A'-/A-positions. *Wh*-island effects always emerge, since *wh*-features and defective π -features cannot readily probe in the same derivational step in English; however, a *wh*-phrase undergoing long-movement can target SPEC-C more easily if the type *wh*-feature is syncretically matched in SPEC-T by a nominative *wh*-phrase. It has been argued that in Dutch π -features such as topic and focus are also syncretically matched in SPEC-T along with φ -features.

The special circumstance that favors mixed A'-/A-positions, a case of syntactic syncretism, to the detriment of the semantic requirement of bringing semantically devoted positions into syntactic representations is that the bearer of token π -features is the same as the bearer of φ -features, a nominative interrogative, focus or topic.

What licenses the C-T distinction is that C becomes a [+clause typing] position and T a [-clause typing] position. If, due to a certain syntactic circumstance that favors structural minimization, both π -features and non- π -features are matched in T, then there is no longer any reason to maintain the C-T distinction in a syntactic representation. Once π -features are matched in TP, CP and TP may be forced to contract in certain languages, in which case there is no structural space in which to insert a complementizer.

PART III. CONCLUSION

Chapter 6. On the emergence of order in syntax

According to the biolinguistic-generative grammar amalgam (see especially Chomsky 2005a, and references cited therein), one's language is a set of emergent properties of an organical structure, the brain. According to Chomsky (2005a), three factors are expected to enter into the growth of language: Factor I (experience), Factor II (genetic endowment) and Factor III (general principles).

Factor I is the experience that the organism faces during the process of maturation, which "leads to variation, within a fairly narrow range, as in the case of other subsystems of the human capacity and organism generally (Chomsky 2005a: 6)".

Factor II is usually called Universal Grammar (UG) and is responsible for constraining the set of possible growing paths. The genetic endowment of an organism, along with the properties of its sensory apparatus, determines the world of experience, or the 'umwelt', to use Uexküll's term (1934/1957). Factor II thus interprets external data as experience, which distinguishes humans as the knowledge of language and bees as the ability to infer the position of the sun's azimuth from the polarization pattern of the sky when they cannot see the sun.

Factor III involves principles that are not specific to UG, such as principles of architecture, development constraints, principles of computational efficiency and principles of analysis and data processing.

To the best of my knowledge, it is fair to claim that the major guiding idea in defining the content of Factor II during these fifty years of generative grammar has been that linguistic objects are made of a particular class of basic units which are organized in a specific type of representations (see, again, Chomsky 2005a for a very clear and personal exposition of the main trends in the biolinguistic-generative grammar amalgam). The basic units are called features (which are coded in lexical items contained in the lexicon) and are interpreted as useful instructions at the levels of interpretation of the Faculty of Language. Representations are claimed to be hierarchical, a property that derives from successive Merge, as argued in chapter 1; accordingly, no language-or-construction-specific rules and no abstract format needs to be attributed to Factor II to derive the relevant hierarchical structure from linguistic expressions. All that is required is the knowledge of the relevant instructions ('features'), the knowledge of a basic combinatorial operation (successive internal and external Merge) and the knowledge of the relevant principles of analysis and data processing to determine constituency structure and distance relations (see Yang 2002, 2004, Gambel & Yang 2003 for the relevance of principles of analysis).

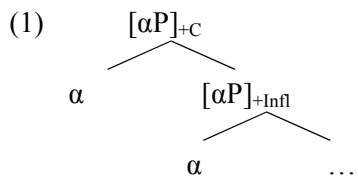
This suggests that no specific elements need to be postulated for the knowledge of grammar, which is a rewiring of the most basic elements independently in place (instructions, successive Merge and principles of analysis). Thus, in some sense, Factor II is a rewiring of elements belonging to Factor III.

The aim of this work has been to extend this reasoning to a relatively detailed study of the intricacies of syntactic patterns (or how features are ordered).

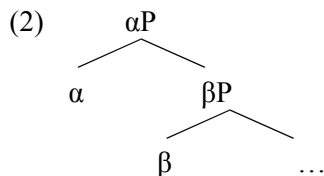
First, it has been proposed that several Infl-like features such as φ -features, tense, mood, modality and negation are assigned both to a [+clause typing] position (C) and to a [-clause typing] position (Infl), trivially accounting for the C-Infl link:

- (I) Complementizers can ‘replicate’ Infl-like features
- (II) There is a correlation between the characteristics of features surfacing on Infl and the choice of C
- (III) Infl-like features are involved in triggering V-to-C

The semantic instruction provided by Infl-like features (roughly, to trigger referential displacement in several dimensions) is orthogonal to the [\pm clause typing] distinction. Our specific proposal has been that C and Infl are two polarities of the same feature, and in that sense, they constitute a discontinuous syntactic pattern. This orthogonality has been argued to be the source of the C-Infl link: when Infl-like features are introduced into the discontinuous template ‘[+clause typing]_C...[-clause typing]_{Infl}’, they are merged in both poles, splitting into a [+clause typing] occurrence and a [-clause typing] occurrence. Thus, an Infl-like feature α is associated with two positions of a discontinuous syntactic pattern, and hence is a discontinuous feature in syntax.



Second, following the rationale of cartographic studies, it has been argued that, in general, the C-I system requires analytic syntactic patterns, i.e., syntactic representations with a one-to-one relation between type features and positions.



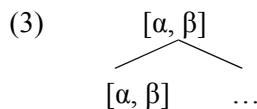
Several ordering restrictions in the Infl-area have been attributed to a semantic legibility condition, the Full Interpretation Condition.

Full Interpretation Condition

An object generated by a grammatical component must be constituted only of useful elements for a particular level of interpretation when it attains such a level of interpretation

Discontinuous syntactic patterns do not contradict the idea that the C-I system requires articulated syntactic representations with semantically devoted positions, but rather adhere to it: both occurrences of an Infl-like feature α are minimally distinguished, and thus, they are not the same element in two positions but two covariations. Thus, the creation of [\pm clause typing] covariations (or discontinuous syntactic objects) is at the service of the creation of syntactic representations with semantically devoted positions or (analytic syntactic objects). Discontinuous syntactic patterns are a subtype of analytic syntactic patterns. As argued, the observed connections between P and K and v and V offer two other plausible discontinuous syntactic patterns.

And third, it has been claimed that syncretic syntactic patterns (where two features are matched in the same projection) also exist.



Whereas analytic and discontinuous syntactic patterns (or cartographic effects) have been attributed to the C-I requirement of bringing semantic distinctions into syntactic representations, syncretic syntactic patterns (or anti-cartographic effects) have been derived from a principle of structural minimization, the Maximize Matching Effects Principle.

Maximize Matching Effects Principle (MMEP)

Instantiate as many type features as possible using the smallest span of structure

The MMEP does not apply arbitrarily, but under special syntactic circumstances that favor structural condensation to the detriment of the C-I legibility requirement of creating analytic syntactic patterns and discontinuous syntactic patterns, where a linear strict order is obtained.

Aside from the Full Interpretation Condition and the MMEP, a further factor has been regarded as important in defining the shape of syntactic objects: inflectional morphology. Our case study has been the observation that A-movement to SPEC-T is possible and necessary only when the verbal inflection does not suffice to instantiate ϕ -features on T. It is tempting to speculate that, if our morphological and syntactical analysis were accurate enough, we would see that the morpholexical characterization of words is the seed of syntactic objects.

Some ideas that would deserve further research have been proposed: (i) the C-Infl discontinuity may contract in some languages under some special circumstances, thereby rendering comp-insertion impossible, and (ii) the Relativized Opacity Principle, an alternative to the Phase Impenetrability Condition to account both for subject islands and *wh*-islands.

We have come across several syntactic devices postulated in the literature, such as the Linear Correspondence Axiom, suicidal greed and the existence of uninterpretable features, the Extended Projection Principle and the existence of merely occurrential features (EPP/OCC) and the Vacuous Movement Hypothesis, and we have concluded that they are inaccurate and unnecessary. No more than a recursive procedure responsible for creating nests by successive Merge and for ordering and instantiating type features driven by the morpholexical characterization of words and constrained by the Full Interpretation legibility condition and the Maximize Matching Effects principle of structural efficiency is required, thereby supporting the idea that the syntactic component is an optimal solution to legibility conditions.

References

- Acedo, Víctor & Jordi Fortuny. 2006. Exploring the Strongest Minimalist Thesis. In Joana Rosselló & Jesús Martín, eds., *The Biolinguistic Turn. Issues on Language and Biology*. Barcelona: Publicacions Universitat de Barcelona.
- Adams, Marianne. 1987. From Old French to the theory of prodrop. *Natural Language and Linguistic Theory* 5: 1-32.
- Akmajian, Adrian. 1975. More Evidence for an NP Cycle. *Linguistic Inquiry* 6: 115-129.
- Alexiadou, Artemis & Elena Anagnostopoulou. 1998. *Natural Language and Linguistic Theory* 16: 419-539.
- Barss, Andrew. 2001. Syntactic reconstruction effects. In Martin Baltin & Chris Collins, eds., *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.
- Bayer, Josef. 1984. Comp in Bavarian. *The Linguistic Review* 3: 209-274.
- Beghelli, Filippo & Tim Stowell. 1997. Distributivity and Negation: The Syntax of *Each* and *Every*. In Anna Szabolcsi, ed., *Ways of Scope Taking*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Belletti, Adriana & Luigi Rizzi, eds. 2002. An interview on minimalism. In *On Nature and Language*. Cambridge, Mass.: Cambridge University Press.
- Besten, Hans den. 1977/1981. On the Interaction of Root Transformations and Lexical Deletive Rules. *Groninger Arbeiten zur Germanistischen Linguistik* 20.
- Bhatt, Rajesh. 1999. Covert Modality in Non-finite Contexts. PhD. Dissertation. University of Pennsylvania.
- Bianchi, Valentina. 2001a. *On Person Agreement*. Ms. Scuola Normale Superiore. Pisa.
- Bianchi, Valentina. 2001b. *On Finiteness as Logophoric Anchoring*. Ms. Scuola Normale Superiore. Pisa.
- Boeckx, Cedric. 2003. *Islands and Chains. Resumption as stranding*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Boeckx, Cedric. 1999. Conflicting c-command requirements. *Studia Linguistica* 53: 227-250.
- Bonet, Eulàlia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. PhD dissertation. MIT.
- Borer, Hagit. 2004a. *Structuring sense*. Vol. 1, *In name only*. New York/Oxford: Oxford University Press.
- Borer, Hagit. 2004b. *Structuring sense*. Vol. 2, *The normal course of events*. New York/Oxford: Oxford University Press.

- Boskovic, Željko. 1999. On Multiple Feature Checking. In Samuel D. Epstein & Norbert Hornstein, eds., *Working Minimalism*. Cambridge, Mass.: MIT Press.
- Bresnan, Joan. 1977. Variables in the theory of transformations. In Peter Culicover et al., eds., *Formal Syntax*. New York: Academic Press.
- Bresnan, Joan. 1972. *Theory of Complementation in English Syntax*. PhD dissertation. MIT.
- Burzio, Luigi. 1986. *Italian Syntax*. Dordrecht: Reidel.
- Bybee, Joan. 1985. *Morphology: A study of the Relation between Meaning and Form*. Amsterdam: John Benjamins Publishing Company.
- Carstens, Vicki. 2001. Multiple Agreement and Case-deletion: Against ϕ -(in)completeness. *Syntax*, 4: 147-163.
- Chomsky, Noam. 2005a. Three Factors in Language Design. *Linguistic Inquiry* 36: 1-22.
- Chomsky, Noam. 2005b. *On Phases*. Ms. MIT.
- Chomsky, Noam. 2001a. Derivation by Phase. In Michael Kenstowicz, ed., *Ken Hale: A life in Language*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2001b. Beyond Explanatory Adequacy. *MIT Occasional Working Papers in Linguistics*, 20. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In Robert Martin, David Michaels & Juan Uriagereka, eds., *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1994. Bare phrase structure. *MIT occasional papers in linguistics* 5. Department of Linguistics and Philosophy, MIT.
- Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In Noam Chomsky 1995. Originally published In Kenneth Hale & Samuel Jay Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1991. Some Notes on Economy of Representations. In Noam Chomsky 1995. Originally published in Robert Freidin, ed., *Principles and Parameters in Comparative Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

- Chomsky, Noam. 1973. Conditions on Transformations. In Stephen Anderson & Paul Kiparsky, eds., *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1964a. The logical basis of linguistic theory. In Horace G. Lunt et al, eds, *Proceedings of the Ninth International Congress of Linguistics*, Cambridge, Mass., 1962. The Hague: Mouton & Co.
- Chomsky, Noam. 1964b. *Current Issues in Linguistic Theory*. The Hague: Mouton & Co.
- Chomsky, Noam & Howard Lasnik. 1993. The Theory of Principles and Parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, eds., *Syntax: An International Handbook of Contemporary Research*. Berlin/New York: Walter de Gruyter. Reprinted in Noam Chomsky 1995.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton & Co.
- Chomsky, Noam. 1955/1975. *Logical Structure of Linguistic Theory*. Chicago: Plenum.
- Chomsky, Noam & Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425-504.
- Chung, Sandra & James McCloskey. 1987. Government, Barriers and Small Clauses in Modern Irish. *Linguistic Inquiry* 18: 173-237.
- Cinque, Guglielmo. 2005. Deriving greenbergs universal 20 and its exceptions. *Linguistic Inquiry* 36: 315-332.
- Cinque, Guglielmo. 2000. "Restructuring" and Functional Structure. Ms. Università di Venezia.
- Cinque, Guglielmo. 1999b. "Restructuring" and the Universal Hierarchy of Functional Heads. Presented at the *GLOW 1999 Colloquium*. Berlin.
- Cinque, Guglielmo. 1999a. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. New York/Oxford: Oxford University Press.
- Cinque, Guglielmo. 1996. The Antisymmetric Programme: Theoretical and Typological Implications. *Journal of Linguistics* 32: 447-464.
- Cinque, Guglielmo. 1994. Evidence for partial N-movement in the Romance DP. In Guglielmo Cinque et al., eds., *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne*. Washington, D.C.: Georgetown University Press.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.
- Cole, Peter & Gabriella Hermon 2000. Partial wh-movement: evidence from Malay. In Uli Lutz, Gereon Muller, Arnim von Stechow, eds., *Wh-scope Marking*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Cormack, Annabel & Neil Smith. 2002. Modals and negation in English. In Sjef Barbiers, Frits Beukema & Wim van der Wurff, eds., *Modality and its Interaction with the Verbal System*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Corver, Norbert. 1997a. The Internal Syntax of the Dutch Extended Adjectival Projection. *Natural Language and Linguistic Theory* 15: 289-368.
- Corver, Norbert. 1997b. Much Support as a Last Resort. *Linguistic Inquiry* 28: 119-164.
- Dermidache, Hamida & Myriam Uribe-Etxebarria. 2000. The Primitives of Temporal Relations. In Roger Martin et al., eds., *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press.
- De Villiers, Jill G., Roeper, Thomas & Vainikka, Anne. 1990. The acquisition of long-distance rules. In Frazier, Lyn and Jill de Villiers, eds., *Language Processing and Language Acquisition*. Kluwer, Dordrecht.
- Dikken, Marcel den. 2006. Vacuous Movement in Focus. On the syntax of highest-subject-wh-questions and relative clauses. Presented at the 29th GLOW Colloquium. Universitat Autònoma de Barcelona.
- Drubig, Hans Bernhard. 2001. *On the Syntactic Form of Epistemic Modality*. Ms. Universität Tübingen.
- Duarte, Maria Eugenia L. 1993. Do pronome nulo ao pronome pleno. In Ian Roberts & Mary Aizawa Kato, eds., *Português Brasileiro: uma viagem diacrônica* (Homenagem a Fernando Tarallo). Campinas: Editora de UNICAMP.
- Emonds, Joseph E. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- Emonds, Joseph E. 1978. The Verbal Complex V'-V in French. *Linguistic Inquiry* 9: 151-175.
- Felser, Claudia. 2004. Wh-copying, phases and successive cyclicity. *Lingua* 114: 543-574.
- Foley, William & Robert D. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, Mass.: MIT Press.
- Freidin, Robert. 1999. Cyclicity and Minimalism. In Samuel D. Epstein & Norbert Hornstein, eds., *Working Minimalism*. Cambridge, Mass.: MIT Press.
- Freidin, Robert. 1978. Cyclicity and the Theory of Grammar. *Linguistic Inquiry* 9: 4.
- Gallego, Ángel. 2006. *Phase Sliding*. Ms. Universitat Autònoma de Barcelona/University of Maryland.
- Gallego, Ángel. 2004. *Phase Effects in (Iberian) Romance*. Ms. Universitat Autònoma de Barcelona/Universidad del País Vasco-Euskal Herriko Unibertsitatea.
- Gambell, Timothy & Charles D. Yang. 2003. *Scope and Limits of Statistical Learning in Word Segmentation*. Ms. Yale University.

- Giorgi, Alessandra & Fabio Pianessi. 1997. *Tense and Aspect: From Semantics to Morphosyntax*. New York/Oxford: Oxford University Press.
- Giorgi, Alessandra & Fabio Pianessi. 1996. Verb movement in Italian and Syncretic categories. *Probus* 8: 137-160.
- Grimshaw, Jane. 1997. Minimal projection, heads and optimality. *Linguistic Inquiry* 28: 373-422.
- Haegeman, Liliane. 1995. *The Syntax of Negation*. Cambridge: Cambridge University Press.
- Haegeman, Liliane. 1992. *Some speculations on argument shift, clitics and crossing in West Flemish*. Ms. Université de Genève.
- Hale, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In Kenneth Hale and Samuel Jay Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.
- Hendrick, Randall. 2000. Celtic Initials. In Andrew Carnie & Eithne Guilfoyle, eds., *The Syntax of Verb Initial Languages*. New York/Oxford: Oxford University Press.
- Higginbotham, James & Robert May. 1981. Questions, Quantifiers, and Crossing. *The Linguistic Review* 1: 5-47.
- Holmberg, Anders & Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. New York/Oxford: Oxford University Press.
- Horn, Laurence R. (1989). *A natural history of negation*. New York/Chicago: The University of Chicago Press.
- Horn, Laurence R. & Yasuhiko Cato. 2000. *Negation and Polarity*. New York/Oxford: Oxford University Press.
- Huang, Cheng-Teh James. 2006. The macro-history of Chinese Syntax and the theory of change. Invited talk given at the University of Chicago and conference given at the University of Toronto.
- Huang, Cheng-Teh James. 2005. Syntactic Analyticity and the Other End of the Parameter. Summer Linguistic Institut. Harvard-MIT.
- Huang, Cheng-Teh James. 1982. *Local Relations in Chinese and the Theory of Grammar*. PhD dissertation. MIT.
- Iatridou, Sabine & Kai von Stechow. 2003. Epistemic Containment. *Linguistic Inquiry* 34: 173-198.
- Jayaseelan, Karattuparambil A. 2001. IP-Internal Topic and Focus Phrases. *Studia Linguistica* 55 (1): 59-75.
- Kato, Mary Aizawa. 1999. Strong and weak pronominals in the null subject parameter. *Probus* 11: 1-37.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

- Kayne, Richard S. 1989. Notes on English agreement. Reprinted in Richard S. Kayne. 2000. *Parameters and Universals*. New York/Oxford: Oxford University Press.
- Kayne, Richard S. 1975. *French Syntax. The Transformational Cycle*. PhD dissertation. MIT.
- Kean, Mary-Louise. 1974. The Strict Cycle in Phonology. *Linguistic Inquiry* 5: 179-203.
- Kim, Yookyung. 1994. A non-Spurious Account of "Spurious" Korean Plurals. In Youn-Key Kim-Renaud, ed., *Theoretical Issues in Korean Linguistics*. Stanford: CSLI.
- Klima, Edward S. 1964. Negation in English. In Jerrold Katz & Jerry Fodor, *The Structure of Language*. Englewood Cliffs, New Jersey: Prentice Hall.
- Koizumi, Masatoshi. 1995. *Phrase Structure in Minimalist Syntax*. PhD dissertation. MIT.
- Koopman, Hilda. 1993. *The Structure of Dutch PPs*. Ms. University of California at Los Angeles.
- Koster, Jan. 2003. *All Languages are Tense Second*. Ms. University of Groningen.
- Koster, Jan. 1975. Dutch as an SOV Language. *Linguistic Analysis* 1: 111-136.
- Kratzer, Angelika. 1996. Severing the External Argument from its Verb. In Johan Rooryck & Laurie Zaring, eds., *Phrase Structure and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Kratzer, Angelika. 1995. Stage-Level and Individual-Level Predicates. In Carlson, Gregory N. & Francis J. Pelletier, eds., *The Generic Book*. Chicago/London: The University of Chicago Press.
- Kratzer, Angelika. 1986. Conditionals. *Chicago Linguistics Society*, 22.2: 1-15. Reprinted in Arnim von Stechow & Dieter Wunderlich, eds., 1991, *Semantics: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter.
- Kratzer, Angelika. 1981. The Notional Category of Modality. In H.J. Eikmeyer & H. Rieser, eds., *Words, Worlds, and Contexts. New Approaches in Word Semantics*. Walter de Gruyter: Berlin. Reprinted in Paul Portner & Barbara Partee, eds., 2002, *Formal Semantics: The Essential Readings*. Oxford: Blackwell.
- Kratzer, Angelika. 1976. Was "können" und "müssen" bedeuten können müssen. *Linguistische Berichte* 42.
- Kuratowski, Kazimierz. 1921. Sur la notion de l'ordre dans la théorie des ensembles. *Fundamenta Mathematicae* 2: 161-171.
- Laka, Miren Itziar. 1990. *Negation in Syntax: On the Nature of Functional Categories and Projections*. PhD dissertation. MIT.
- Lasnik, Howard. In press. Conceptions of the Cycle. In Lisa Lai Shen Cheng and Norbert Corver, eds., *Wh-movement on the Move*. Cambridge, Mass.: MIT Press.
- Lasnik, Howard. 2002. Clause-mate conditions revisited. *Glott International* 6(4), April.

- Lasnik, Howard. 2001. Subjects, objects and the EPP. In Davis, William D. and Stanley Dubinsky, eds., *Objects and other subjects: Grammatical functions, functional categories, and configurationality*. Dordrecht: Kluwer Academic Publishers.
- Lasnik, Howard. 1999. On feature strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30: 197-217.
- Lasnik, Howard. 1972. *Analysis of Negation in English*. PhD dissertation. MIT.
- Lasnik, Howard & Mamoru Saito. 1992. *Move α* . Cambridge, Mass.: MIT Press.
- Lasnik, Howard & Mamuro Saito. 1991. On the Subject of Infinitives. In Lise M. Dobrin et al., eds., *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. Reprinted in Howard Lasnik. 1999. *Minimalist Analysis*. Oxford: Blackwell.
- Lasnik, Howard & Mamoru Saito. 1984. On the Nature of Proper Government. *Linguistic Inquiry* 15: 235-289.
- Lees, Robert B. 1960. *The Grammar of English Nominalizations*. The Hague: Mouton & Co.
- Legate, Julie Anne. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34: 509-516.
- Lewis, David. 1975. Adverbs of Quantification. In Edward Keenan, ed., *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press.
- Marantz, Alec. 1997. No scape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams, eds., *Proceedings of the 21st Annual Penn Linguistics Colloquium*. Pennsylvania Working Papers in Linguistics 4.2. Philadelphia: Penn Linguistics Club, University of Pennsylvania.
- Mascaró, Joan. 1976. *Catalan Phonology and the Phonologic Cycle*. PhD dissertation. MIT.
- McCloskey, James. 2000. Quantifier Float and wh-movement in an Irish English. *Linguistic Inquiry* 31: 57-84.
- McDaniel, Dana, Bonnie Chiu & Thomas Maxfield. 1995. Parameters for wh-movement types: evidence from child English. *Natural Language and Linguistic Theory* 13: 709-753.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. New York/Oxford: Oxford University Press.
- Miyagawa, Shigeru. 2005. On the EPP. In Norvin Richrads & Martha McGinnis, eds., *Perspectives on Phases*. Cambridge, Mass.: MIT Working Paper in Linguistics.
- Ordóñez, Francisco. The inversion construction in interrogatives in Spanish and Catalan. In José Lema & Esthela Treviño, eds., *Theoretical analyses on Romance languages*. Amsterdam: John Benjamins Publishing Company.
- Perlmutter, David M. 1971. *Deep and Surface Constraints in Syntax*. New York: Holt.

- Pesetsky, David & Esther Torrego. 2004a. Tense, Case, and the Nature of Syntactic Categories. In Jacqueline Guéron & Jacqueline Lecarme, eds., *The Syntax of Time*, Cambridge, Mass.: MIT Press.
- Pesetsky, David & Esther Torrego. 2004b. *The Syntax of Valuation and the Interpretability of Features*. Ms. MIT/University of Massachusetts.
- Pesetsky, David & Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In Michael Kenstowicz, ed., *Ken Hale: A Life in Language*. Cambridge, Mass.: The MIT Press.
- Picallo, M. Carme. 1998. On the Extended Projection Principle and null expletive subjects. *Probus* 10: 219-241.
- Picallo, M. Carme. 1985. *Opaque Domains*. PhD dissertation. City University of New York.
- Platzack, Christer & Anders Holmberg. 1989. The Role of AGR and Finiteness in Some European VO Languages. Paper presented at *GLOW 1989 Colloquium*, Utrecht.
- Pollock, Jean-Yves. 1989. Verb-movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry* 20: 365-424.
- Postal, Paul. 1974. *On Raising*. Cambridge, Mass.: MIT Press.
- Progovac, Ljiljana. 1988. *A Binding Approach to Polarity Sensitivity*. PhD dissertation. UCLA.
- Quer, Josep. In press. Subjunctives. In Martin Everaert & Henk van Riemsdijk, eds., *Companion to Syntax*. Oxford: Blackwell.
- Quer, Josep. 2003. Interpreting mood. *Probus* 13: 81-111.
- Quer, Josep. 1998. *Mood at the Interface*. PhD dissertation. UiL OTS/University of Utrecht. The Hague: Holland Academic Graphics.
- Raposo, Eduardo P. 1987. Case Theory and Infl-to-Comp: The Inflected Infinitive in European Portuguese. *Linguistic Inquiry* 18: 85-109.
- Richards, Norvin. 1999. In Samuel D. Epstein & Norbert Hornstein, eds., *Working Minimalism*. Cambridge, Mass.: MIT Press.
- Richards, Norvin. 1997. *What Moves Where in Which Language*. PhD dissertation. MIT.
- Riemsdijk, Henk van. 1996. *Categorial Feature Magnetism: The Extension and Distribution of Projections*. Ms. Universiteit van Tilburg.
- Riemsdijk, Henk van. 1990. Functional Prepositions. In Harm Pinkster & Inge Genée, eds., *Unity in Diversity*. Dordrecht: Foris.
- Riemsdijk, Henk van. 1978a. On the Diagnosis of Wh Movement. In Samuel Jay Keyser, ed., *Recent Transformational Studies in European Languages*. *Linguistic Inquiry* 3. Cambridge, Mass.: MIT Press.
- Riemsdijk, Henk van. 1978b. *A Case Study in Syntactic Markedness*. Lisse: Peter de Ridder Press.

- Rigau, Gemma. 1995. The properties of the temporal infinitive constructions in Catalan and Spanish. *Probus* 7: 279-301.
- Rizzi, Luigi. 2004. *On the Form of Chains: Criterial Positions and ECP Effects*. Ms. Università di Siena.
- Rizzi, Luigi. 2002. Locality and left Periphery. In Adriana Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures*, vol. 3. New York/Oxford: Oxford University Press.
- Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman, ed., *Elements of Grammar*. Dordrecht: Kluwer Academic Publishers.
- Rizzi, Luigi. 1991. Residual verb second and the Wh criterion, *Technical Reports in Formal and Computational Linguistics 2*, Université de Genève.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17: 501-557.
- Rizzi, Luigi. 1982a. Comments on Chomsky's chapter "On the representation of form and function". In Jacques Mehler, Edward C.T Walker & Merrill Garrett, eds., *Perspectives on mental representations: experimental and theoretical studies of cognitive processes and capacities*. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers.
- Rizzi, Luigi. 1982b. *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, Luigi & Ian Roberts. 1989. Complex Inversion in French. *Probus* 1: 1-30.
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.
- Rochette, Anne. 1988. *Semantic and Syntactic Aspects of Romance Sentential Complementation*. PhD dissertation. MIT.
- Rosenbaum, Peter Steven. 1965. *The Grammar of English Predicate Complement Constructions*. PhD dissertation. MIT.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. PhD dissertation. MIT.
- Rosselló, Joana. 2000. A Minimalist Approach to the Null Subject Parameter. *Catalan Working Papers in Linguistics* 8: 97-128.
- Rosselló, Joana. 1986. *Gramàtica, Configuracions i Referència*. PhD dissertation. Universitat de Barcelona.
- Saddy, James Douglas. 1991. *Investigations into Grammatical Knowledge*. PhD dissertation. MIT.
- Sigurðsson, Halldór Ármann. 2004. The syntax of Person, Tense and speech features. *Italian Journal of Linguistics/Rivista di Linguistica*.
- Sigurðsson, Halldór Ármann. 2003. *Meaningful silence, meaningless sounds*. Ms. Lunds Universitet.

- Sigurðsson, Halldor Ármann. 1993. Argument-drop in Old Icelandic. *Lingua* 89: 247-280.
- Sobin, Nicholas. 1987. The Variable Status of Comp-Trace Phenomena. *Natural Language and Linguistic Theory* 5: 33-60.
- Solà, Jaume. 2002. Clitic Climbing and Null Subject Languages. *Catalan Journal of Linguistics* 1: 225-257.
- Solà, Jaume. 1996. Morphology and Word Order in Germanic Languages. In Abraham Werner et al., eds., *Minimal Ideas. Syntactic Studies in the Minimalist Framework*. Amsterdam: John Benjamins Publishing Company.
- Solà, Jaume. 1992. *Agreement and Subjects*. PhD. dissertation. Universitat Autònoma de Barcelona.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425-449.
- Stowell, Timothy Angus. 1981. *Origins of Phrase Structure*. PhD dissertation. MIT.
- Suñer, Margarita. 1994. V-Movement and the Licensing of Argumental *Wh*-Phrases in Spanish. *Natural Language and Linguistic Theory* 12: 335-372.
- Svenonius, Peter. 2002. Subject Positions and the Placement of Adverbials. In Peter Svenonius, ed., *Subjects, Expletives and the EPP*. New York/Oxford: Oxford University Press.
- Takahashi, Daiko. *Minimality of Movement*. PhD dissertation. University of Connecticut.
- Taraldsen, Knut Tarald. 1980. *On the Nominative Island Condition: Vacuous Application and the that-trace filter*. Indiana University Linguistics Club.
- Thornton, Rosalind. 1990. *Adventures in long distance moving: the acquisition of complex wh-questions*. PhD dissertation. University of Connecticut.
- Torrego, Esther. 1985. *On Empty Categories in Nominals*. Ms. University of Massachusetts.
- Uexküll von, Jakob. 1934/1957. A Stroll Through the Worlds of Animals and Men. In Claire H. Schiller, editor and translator, *Instinctive Behaviour*. New York: International Universities Press.
- Ura, Hiroyuki. 2000. *Checking theory and grammatical functions in universal grammar*. New York/Oxford: Oxford University Press.
- Uriagereka, Juan. 2004. *Spell Out consequences*. Ms. University of Maryland.
- Vikner, Sten. 1994. Finite verb movement in Scandinavian embedded clauses. In David Lightfoot & Norbert Hornstein, eds., *Verb Movement*. Cambridge: Cambridge University Press.
- Villiers, de, Jill G., Thomas Roeper & Anne Vainikka. 1990. The acquisition of long-distance rules. In Lyn Frazier & Jill de Villiers, eds., *Language Processing and Language Acquisition*. Dordrecht: Kluwer.

- Yang, Charles D. 2004. Universal Grammar, Statistics or both? *Trends in Cognitive Sciences*, vol.8, n.10.
- Yang, Charles D. 2002. *Knowledge and Learning in Natural Language*. New York/oxford: Oxford University Press.
- Zagona, Karen. 1998. Structural case and case construal. In Jean-Marc Authier, Barbara E. Bullock & Lisa A. Reed, eds., *Formal Perspectives on Romance Linguistics*. Amsterdam: John Benjamins Publishing Company.
- Zamparelli, Roberto. 1995. *Layers in the Determiner Phrase*. PhD dissertation. University of Rochester.
- Zanuttini, Raffaella. 1997. *Negation and Clausal Structure. A Comparative Study of Romance Languages*. Nova York/Oxford: Oxford University Press.
- Zwart, Cornelius Jan Wouter. 2001. Syntactic and Phonological Verb Movement. *Syntax* 4: 34-62.
- Zwart, Cornelius Jan-Wouter. 1993. *Dutch Syntax. A Minimalist Approach*. University of Groningen: Groningen Dissertations in Linguistics.