CHAPTER 1

PROLIFIC PERIPHERIES: BY WAY OF INTRODUCTION

Kleanthes of Assos (331 to 232 BC) left behind one remarkable piece of work—in fact the only piece of work of his that survived the past twenty-three centuries: the *Hymn to Zeus*. The remaining fragments are commonly taken to constitute the first attempt to express philosophical thought through poetic means. This dissertation is neither philosophical nor poetic in nature, and it does not even pretend to be, but it contains ideas which may be paralleled to the *Hymn to Zeus* in the sense that it presents old ideas in new form and packaging. For his dullness, Kleanthes was also called by his peers an ass.

1.1 Locality

Locality has been a pervasive issue in generative grammar since the early days.[1] As much of this dissertation has bearings on locality in one way or another, let us embark on a brief and incomplete tour of some of the issues.

[1] Locality defined structural relations already in Chomsky (1955), was relevant for Affix Hopping in the original framework of Chomsky (1957), it played a role in Standard Theory (Chomsky 1964, 1965, and certainly Ross 1967), and especially came into everyone's

(continued…)

We commonly take some version of locality to be relevant to the formation of all dependencies, which we can break up into two general types. One amounts to the result of "rules of transformations," one way of expressing displacement, that famous property of natural language. The other one, by what we might call "rules of construal," establishes a relation between two or more positions through means other than movement (coindexation, predication and so on). Displacement, viz. transformation (henceforth, movement), takes a constituent from one position in the phrase-marker to another, linking the new position (landing site) and the original starting point (launching site). This dependency has several properties, however formalized, such as anaphoric relations between launching and landing sites, and other licensing conditions on the launching site (on the trace/copy). As such, movement-derived structures may be contrasted with base-generated relations, which can be seen as instantiations of a rule of construal, creating a dependency between two or more relevant positions without resorting to movement.

1.1.1 Movement and Locality

Let us start with a brief discussion of the traditional realm of locality, transformational dependencies that relate two positions by movement.[2] All movement is connected in one way or another to locality conditions, some of which are stricter than others. Paradigmatic examples are raising and expletive constructions, passives and other instances of A-movement, as well as Wh-movement, topicalization and other types of A'-movement, arguably different in nature from A-movement, as well as head movement, distinct from phrasal movement.

---

view in Chomsky (1973, 1977) for possibilities of binding, instances of A-movement and the formulation of A'-movement in Extended Standard Theory.

[2] For reasons that will become obvious soon, I choose the term "dependency" rather than something more technical (such as "chain"). I will stick to this terminology throughout.

Raising is one instance of A-movement. It involves what appears to be successive-cyclic movement of one argument noun phrase, subject to locality restrictions:

(1) a. John$_i$ seems [t_i to t_i like Mary].
b. John$_i$ is believed [t_i to be likely [t_i to seem [t_i to t_i like Mary]]].
c. * John is believed that [it is likely [t_i to seem [t_i to t_i like Mary]].
d. * John is believed [t_i to be likely [it seems [t_i to t_i like Mary]].

By standard convention we assume that the lower trace of *John* in (1a) is base-generated in the thematic subject position of the embedded verb, from which it receives a thematic role, or theta-role (θ-role). We further assume that the higher trace marks the grammatical subject position of the embedded clause, where it satisfies the Extended Projection Principle, the condition that every sentence must have a subject. The raising verb *seem* does not assign an external θ-role (and neither does the predicate *be likely*). *John* then moves from the embedded to the matrix subject position. Thus the logical or thematic subject of the embedded clause is the grammatical subject of both the embedded and the matrix clause, while the matrix clause does not have a thematic subject.

(1b) illustrates that such DP-raising is not restricted to only one occurrence. Given that *John* moves in (1a) to receive Case from finite INFL (nominative case),[3] it must raise further if the next highest INFL is non-finite; the result is a successive-cyclic application of raising. Applying the logic behind the approach to (1a), we expect the relevant parts of the structure to look as shown, namely with a trace of *John* in the thematic subject position of the most deeply embedded lexical verb, a trace in the grammatical subject position of the most deeply embedded sentence, and a trace in the grammatical subject position of

[3] Following by now standard conventions, I capitalize "Case" when it refers to an abstract feature or property of a DP, and reserve the spelling "case" for specific instances, such as nominative, accusative etc. (See Chomsky 1981 where it is also suggested that non-finite INFL does not have Case-assigning properties.)

the next highest embedded sentence (and so on)—where all embedded sentences are non-finite, hence cannot assign nominative case, and involve predicates which do not select an external argument.

The highest embedded sentence in (1c) is finite, and this property gives the story a little twist in the direction of locality. If a DP could move from any (non-finite) subject position to any finite one, why is (1c) ungrammatical? (Successive-cyclic) DP-movement is subject to certain locality conditions. In particular, Move DP cannot skip a potential landing site: in (1b), *John* passes through all available subject positions (of the infinitival clauses), and in (1c) it fails to do so, it would have to move over *it*. Reversing the point when the embedded finite clause is added does not make the sentence any better; cf. (1d). The for us relevant notion of locality is thus that the raising subject can and must move through all subject positions on its way to a finite position, and finite positions may not interfere.

Another type of phrasal movement, also subject to "locality," yet quite different from Case-driven A-movement, is A'-movement. One difference is that the A'-moved element, sticking to DPs, is already Case-marked. (2) is an appropriate illustration:

(2) a. Who$_i$ does John like t$_i$?
b. Who$_i$ does Bill believe [t$_i$ (that) Peter said to Jane [t$_i$ (that) John likes t$_i$]]?
c. * Who$_i$ does Bill believe [who$_j$ Peter said to t$_j$ [t$_i$ (that) John likes t$_i$]]?

In a classic example of A'-movement, movement of a Wh-phrase in languages like English, the Wh-phrase is fronted to the beginning of the sentence, the COMP position. Whether *who* in (2a) receives accusative case in situ or in an intermediate position, it ends up in SpecCP—and this movement seems to be able to take place more than once, as (2b) shows. In (2b) the Wh-phrase moves, again successive-cyclically, through all intermediate SpecCP positions. That the dependency between the fronted *who* and the

original trace might indeed be of successive-cyclic nature can be witnessed in (2c): here one of the intermediate SpecCP positions is filled by another Wh-phrase, and the lower *who* may not move across an "intervening" *who*. The successive-cyclic nature of Wh-movement is not undeniable, but suggestive.

Zooming in on locality, regardless of whether Wh-movement is indeed successive-cyclic or unbounded,[4] it is blocked in some cases. Islands form a famous class of such blockers. The long and short of an island is that it constitutes a barrier to movement:

(3) a. * Who_i does Bill believe [$_{DP}$ the claim [$_{CP}$ t_i that [$_{TP}$ John likes t_i]]]?
b. * Who_i does John read [$_{DP}$ the book [$_{CP}$ (t_i) which Mary gave t_i]]?
c. * Who_i was [$_{DP}$ a friend of t_i] known to kiss Bill?
d. * Who_i was [$_{CP}$ t_i that John likes t_i] expected?

(3a) illustrates the Complex Noun Phrase Constraint with a sentential complement and (3b) with a relative clause. In both instances it is illicit to extract a Wh-phrase out of the noun phrase. Regardless of how we want to account for the impossibility of such extraction, it violates some sort of locality condition. (3c) is an instance of attempted extraction out of a subject. This so-called Subject Condition receives one or two question marks rather than an asterisk in much of the literature, but the type of Wh-element presumably plays a role for this judgement also (e.g. *which girl* is often taken to be better

[4] An unbounded dependency could then involve only the landing site and the launching site. This would naturally lose any relation to A-movement. Note that such a state of affairs would neither be necessarily detrimental (as the two types of movement differ anyway) nor be unidirectional: one can imagine a similar approach to A-movement, namely one where intermediate subject positions of infinitival clauses are not filled. The latter approach has recently been advocated, among others, by Castillo, Drury & Grohmann (1997), Epstein & Seely (1999), Grohmann, Drury & Castillo (2000), Nasu (2000), and to some extent Boeckx (2000a), Bošković (2000)—a direction continued here, at least in spirit, and refuted as "Alternative II" without discussion in Chomsky (1999), for reasons which Castillo, Drury & Grohmann (1999) address. Whatever the outcome, successive cyclicity can be divorced from locality issues which remain regardless (see chapter 6).

than *who*). A possibly better case to illustrate is the Sentential Subject Condition in (3d), where we try to extract a Wh-element out of a sentential subject.[5]

Not all movement targets phrases. A crucial ingredient of generative analyses at least since Baker (1988) has been movement of heads. It is standardly assumed that rather than lowering an affix to the desired verbal position, the verb itself moves. Consider (4):

(4) a. Can_i Bill t_i have kissed Mary?
b. * $Have_i$ Bill can t_i kissed Mary?
c. * $Kissed_i$ Bill can have t_i Mary?

Assuming for the sake of discussion that *yes/no*-questions involve minimally inversion of the subject and the inflected verbal element, at least one verbal head must have moved, even in the grammatical (4a). One possible way to rule out (4b,c) is to say that *can* is closer to C, the locus of inversion, than *have* or *killed*, and head movement targets the closest available head. Well, this amounts to saying that head movement is local and as such restrained by locality conditions.

As a general tendency, we might think of locality as a condition which forces a certain element, head or phrase, to move to the next closest position, or to move locally. An influential theory of locality is Rizzi's (1990) Relativized Minimality, and the minimalist framework adopted here has various technical ways of incorporating it. Apart from this upper bound on the distance in a given dependency, we will consider cases which seem to require a lower bound also, a minimal distance that must separate two positions.

---

[5] It has been debated whether the bracketed constituent really is a sentential subject or rather an adjunct; see, for instance, Koster (1978). Should it be correct, though, to say that the sentential constituent is "only" an adjunct, we can still account for the ungrammaticality in (3d), given that adjuncts also invoke islands. Trying to extract out of an Adjunct Island is thus just as bad, regardless of whether we can employ more local movement steps, as indicated in parentheses in (i):

(i) * Who_i did Bill kiss Mary [(t_i) after John told him (t_i) that he liked t_i]?

1.1.2 Construal and Locality

Now that we have set the stage for locality in movement, let us see whether we find similar conditions on the distance between dependencies in constructions that have standardly been assumed to not involve movement. The reasoning to do so is two-fold: first, it is not so clear that such dependencies are always really the results of construal rather than movement; secondly, and independently of the first point, a dependency is a dependency, and, as much of the work done over the past five decades aims to generalize and unifying grammatical (often, apparently unconnected) phenomena, it is only natural to think about the dependencies created between moved and base-generated objects in similar ways, especially from the broad perspective of "locality." Traditional domains of construal are some of the modules postulated in GB, such as the Theta Module, the Binding Module, or the PRO Module; see, for instance, Chomsky (1981, 1986b).

A brief discussion of binding and pronouns will suffice to illustrate locality:

(5) a. John$_i$ likes himself$_i$.
c. John$_i$ likes him$_{*i/j}$.
c. * He$_i$ thinks that Mary/himself$_{i/j}$ likes John$_i$.

These data illustrate some basic aspects of Binding Theory which we can summarize very crudely as follows: anaphors must be bound in their domain (Condition A), pronouns must be free in their domain (Condition B), and R-expressions must be free, period (Condition C). Leaving aside an appropriate definition of "domain," the anaphor *himself* in (5a) is bound in its domain by *John*, it satisfies Condition A, and the structure is grammatical. By parity of reasoning, whatever constitutes the relevant domain for binding in (5a) should be the same in (5b) and as such, the pronoun *him* would be bound, hence violating Condition B. Given that the entire sentence (5c) seems to be bigger than

the required domain for binding of anaphors, or lack thereof in the case of pronouns (cf. (6a,b)), the R-expression *John* is not bound in its relevant domain, but it is not entirely free either, as the contrast with (6c) shows. (The alternative in (5c), with the reflexive as the embedded subject, shows that *he* cannot serve as the antecedent of *himself* because anaphors must be bound in a more restricted domain than across clause boundaries, and that its antecedent cannot follow, nor be bound by it, if we take that to be *John*.)

(6) a. * $John_i$ thinks that Mary likes $himself_i$.
b. $John_i$ thinks that Mary likes him_i.
c. Jane said that Mary really likes him_i. $John_i$ is very happy.

In sum, we can observe that once again, locality conditions and restrictions seem to play a major role in the licensing of pronouns, the dependency between antecedent and anaphor, and general binding (im)possibilities.

Lastly, locality is also an issue in control structures, different from raising at least in that the matrix subject bears two θ-roles in Control and only one in raising. Consider the following paradigm:

(7) a. $John_i$ hopes $PRO_{i/*j}$ to kiss Mary.
b. * $John_i$ hopes $Jane_l/Bill_k/him_j/himself_i$ to kiss Mary.
c. $John_i$ hopes that $Mary_j$ wants $PRO_{*i/j/*k}$ to kiss him.

In (7a) *John* is the hoper as well as the (potential) kisser, here indicated as coindexed PRO, the GB notation for a phonetically empty subject of infinitival control clauses. The contrast with (7b) suggests that the embedded infinitival clause may not have an overtly realized subject (regardless of presence of R-expression vs. pronominal, or gender, felicitating a context that could express "traditional values"), not even one that is coreferent with the matrix subject (reflexive), independent of other binding relations

(pronoun vs. anaphor). That the condition for coindexing the overtly realized matrix subject and the embedded subject PRO must take locality conditions into consideration, is expressed in (7c): introducing a further level of embedding, only the subject of the clause immediately dominating the one containing PRO can control PRO; neither a higher subject (such as matrix *John*) nor some other referent can do so. (The binding of *him* is irrelevant here but could be indexed "i" or "k," of course.)

1.1.3 Local Summary

In sum, locality is a pervasive issue for all types of dependencies, regardless of whether these are derived (by movement) or base-generated (by some rule of construal), and independent of further sub-classes within the two types. This is by no means a novel discovery. The discussion was important, however, not only as it sets the stage for much more to come, but also as it shows us at least two things.

Firstly, if "locality" can be viewed as an overarching theme common to any analysis of the above, different types of locality must be invoked. They have to do with defining appropriate domains for binding, regulating licensing of PRO, formalizing islands and so on. For the most part, we will not be concerned with this type of locality.

Secondly, it points to one domain where constructions standardly assumed to be the result of movement and those which are taken to result from an application of a rule of construal have properties in common. This is more interesting.

As already mentioned, I will concentrate on the latter point, namely that it seems that "locality" (or locality conditions and restrictions) plays a major role for all types of dependencies between two or more positions in the phrase-marker, and that as a general rule, this locality conforms to some version of "closeness." Adopting, and elaborating on, the intuition driving some recent work, I take it to be one desideratum of a minimalist approach that we not only ensure that unnecessary levels of representation are properly

dispensed with of the sort advocated here, but also that we attempt to find unification among apparently different constructions, if the mechanical tools allow this.

Regarding the first point, Chomsky (1993) set the agenda with the call for elimination of D- and S-structure. As it happens, much of Binding Theory (and the Control Module, for example) crucially relied on this distinction. Chomsky himself started reformulating binding conditions in interesting ways. However, the Control or PRO Module was kept pretty much intact by introducing a different type of Case, "Null Case" (cf. Chomsky & Lasnik 1993, Martin 1996). An alternative route would be to deny that thematic roles are assigned in pre-determined structural configurations, but function by and large like other formal features, the current motivation to apply Move. If an object can then "check" a theta-feature, much of the construal needed to derive control (and many other constructions, such as parasitic gaps or *easy-to-please*-constructions) would become unnecessary. If the tools are there, they should be used, and recent work by, among many others, Bošković (1994, 2000b), Lasnik (1995c), Nunes (1995), Bošković & Takahashi (1998), Castillo, Drury & Grohmann (1999), Boeckx (2000a), suggests that not only are the tools there, but once applied, they can account for many base-generated dependencies in derivational terms.[6]

It should be obvious that the types of constructions above and the types of locality conditions involved may overlap, i.e. my decision to illustrate the role of islands with A'-movement does not imply that we do not find island effects in other types of movement, or even in construed constructions. In fact, islands have always been assumed to be a

[6] I hasten to point out that I use the term "derivational" purely for expository reasons, contrasting it with "base-generation." That is to say, whether the syntactic computation proceeds really derivationally, as opposed to representationally, shall not be the concern of this work, despite any personal conviction one way or another (see e.g. Lasnik 2001 for discussion). As far as I can tell, the arguments for one over the other go both ways and I am not prepared to enter the discussion at this point. See especially the proposal by Epstein et al. (1998), and references cited. Likewise, I will leave out of the discussion entirely whether the "derivation" really proceeds bottom-up, as assumed here, or top-down (see e.g. Phillips 1996, to appear, Drury 1999, Guimarães 2000, Richards 2000).

good diagnostic for whether movement has taken place or not. All things being equal, this really is a good diagnostic to tease apart different types of dependencies, but the issues are sometimes blurred, especially under the approach just endorsed, as we will soon see.

This dissertation tackles locality from the following angle. Acknowledging that locality conditions are needed to allow some and rule out other dependencies, we will look at how closely related two positions in a given dependency must be (or can be!). We will ask ourselves whether there is a structural relation between two positions which could be considered to be "too local." Indeed, I will discuss some instances in which a relationship between two objects in the phrase-marker can be thought of as being exactly this, too local to be licensed, or *anti-local*, so to speak.

Anti-locality in the phrase-marker is going to be one major issue raised here. I will explore one possible way to capture what we might call "idiosyncratic discrepancies" of the sorts illustrated above, such as additional filters and conditions as part of the computational system (Theta Theory, Binding Theory, Case Filter, PRO Module etc.). My major aim, however, is to unify anti-local configurations within a derivationally construed computational system. In order to meet this goal, I will dispense with certain proto-typical notions of construal relations (ancillary mechanics such as empty operators or coindexation), motivate their elimination, and show empirically why we would do so.

This task depends on a theoretical extension of the minimalist framework in directions that allow for a definition of (anti-)locality. I propose a tripartite clausal structure, split into three *Prolific Domains* to state (anti-)locality effects. I provide initial empirical support for this proposal in chapter 2 and concentrate on it in chapters 3, 4 and 5. The remainder of this chapter deals with the presentation of the core ingredients of the Minimalist Program that I adopt and an introduction to some ideas that I extend. Finally I sketch the organization, content and main ideas of this work.

1.2 Minimalist Inquisitions

Descriptions of natural language phenomena are one thing; it is something else entirely to account for them in an adequately explanatory fashion. Analytical tools to do so have changed dramatically in the history of generative grammar, and the proposals regarding structures, derivations and other analytical parts that I present here need to be classified somehow. I adopt recent incarnations and developments of the Principles and Parameter model (of Chomsky 1981, 1986a and much subsequent work), the so-called Minimalist Program ("minimalism"), the culmination of work by Chomsky (1993, 1994, 1995a, 1998, 1999) and many other scholars.[7] In its brief history, minimalism has undergone quite a number of changes; not as much in its basic structure and premises as in its mechanics. As such, referring to "the" minimalist framework is more like referring to "the" nightlife of a reasonably developed city: it comes in many flavours, yet shares basic premises. I will tease the different concepts apart in this section, in as far as they are relevant for the present study, and explore certain aspects in the next five chapters.

1.2.1 The Structure of the Grammar

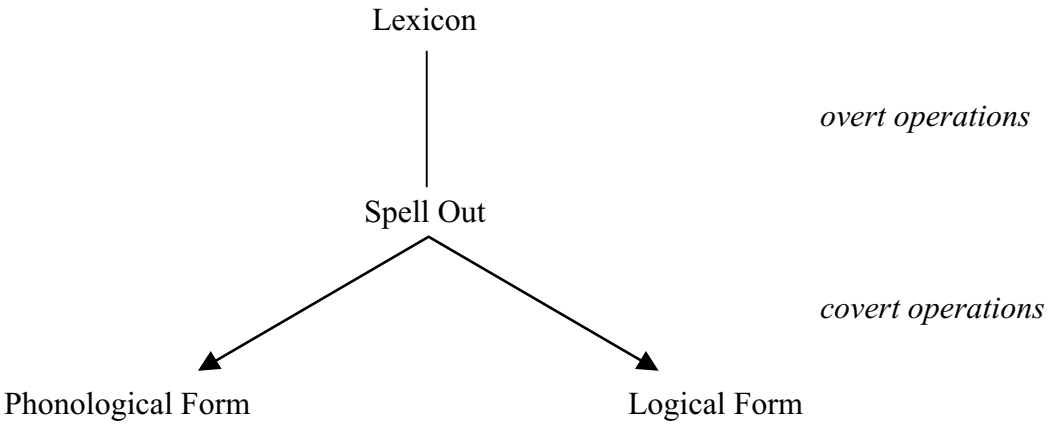
The central premise of the minimalist program is the desire to assume only "bare essentials" for a theory of grammar. Thus any assumption we make about the structure of the grammar, its interface components, the nature of displacement, licensing conditions, and so on must conform to "virtual conceptual necessity." This desideratum led Chomsky (1993) to consider abolishing superfluous levels of representation, reformulating all filters, conditions and principles of the computational system in terms of Bare Output Con-

[7] E.g. Hornstein (1995, 2000), Uriagereka (1998), Lasnik (1999), and the collections of papers in Epstein & Hornstein (1999b), Martin, Michaels & Uriagereka (2000).

ditions, and sketch an economy-driven approach to the computational system.[8] The resulting framework differs quite substantially from earlier formulations of the Principles and Parameters approach.

Regarding levels of representation, Chomsky provides arguments that the levels of D- and S-structure do not meet virtual conceptual necessity—both are (for the most part) theory-internal constructs, and other developments in the program allow for recasting empirical work in other terms, basically following economy principles and Bare Output Conditions.[9] The old T-model can then be replaced by the following model, representing a rough structure of the grammar:

(8)



The computational system C_{HL} performs operations on a pre-selected collection of lexical items, the Numeration (Chomsky 1993, 1995a) or Lexical Array LA (Chomsky 1998, 1999), which also include any functional items needed. The only operations we have are Merge and Move. Merge takes an item out of the Lexicon[10] and puts it together

[8] This program builds on Chomsky (1991), Chomsky & Lasnik (1993) and was further extended in Chomsky (1994, 1995a), or essentially, Chomsky (1995b).

[9] For a more appropriate exposition, see, for example, Epstein & Hornstein (1999a), Martin & Uriagereka (2000), Hornstein, Nunes & Grohmann (in progress).

[10] For our purposes, we can take "Lexicon" to be a cover term for what either the Numeration or the LA, or any combination thereof. Chomsky (1999) actually considers the

(continued…)

with another, and the same operation applies iteratively (on the thus constructed objects of the phrase marker).[11] Move takes an object in a phrase marker and displaces it. The theory of movement adopted here extends this view within the Copy Theory of Movement (Chomsky 1995a, Nunes 1995, 1999). We take the operations that are available to C_{HL} to be essentially Copy and Merge: we take an item (from the numeration or the phrase marker) and merge it with another, and we can copy and re-merge these up to Full Interpretation (see below).[12]

1.2.2 The Computational System

We call the iterative application of these operations the derivation (but recall the disclaimer in fn. 6). We apply Copy and Merge up to convergence. At a certain point the derivation is shipped to the interfaces, the Articulatory-Perceptual (AP) interface and the Conceptual-Intensional (CI) interfaces. This is a simplification. What we actually take to happen at this point (Spell Out) is sending the information to the interface levels which, in turn, are mapped onto the interfaces proper. Arguably, we know very little about the latter, and I will confine myself to a discussion of the syntactic computation. LF is the interface level feeding the CI and PF the AP interface.

---

latter: the Numeration is the original choice of (tokens of) lexical items from the LEXicon, which is then mapped into the LA prior to syntax proper ("narrow syntax," the derivational system sketched here); see also Uriagereka & Martin (1999).

[11] We assume, under Bare Phrase Structure (Chomsky 1994), that one of the merged objects projects, and that it is unambiguous which one. For an exposition, see especially Chomsky (1995b: 241-249).

[12] This is a gross simplification, but appropriate for the current presentation. Chomsky (1995a) originally proposed the relevant operations to be Copy, Merge and Delete, to which Nunes (1995) added Form Chain, among other minor modifications. It is not clear that we really want to consider a "chain" to be a real object in the sense often understood, and I dispense with further illustration (hence my choice of the term "dependency" to denote all relationships between an element and its traces or copies, or other elements). See Hornstein (2000) for discussion and references.

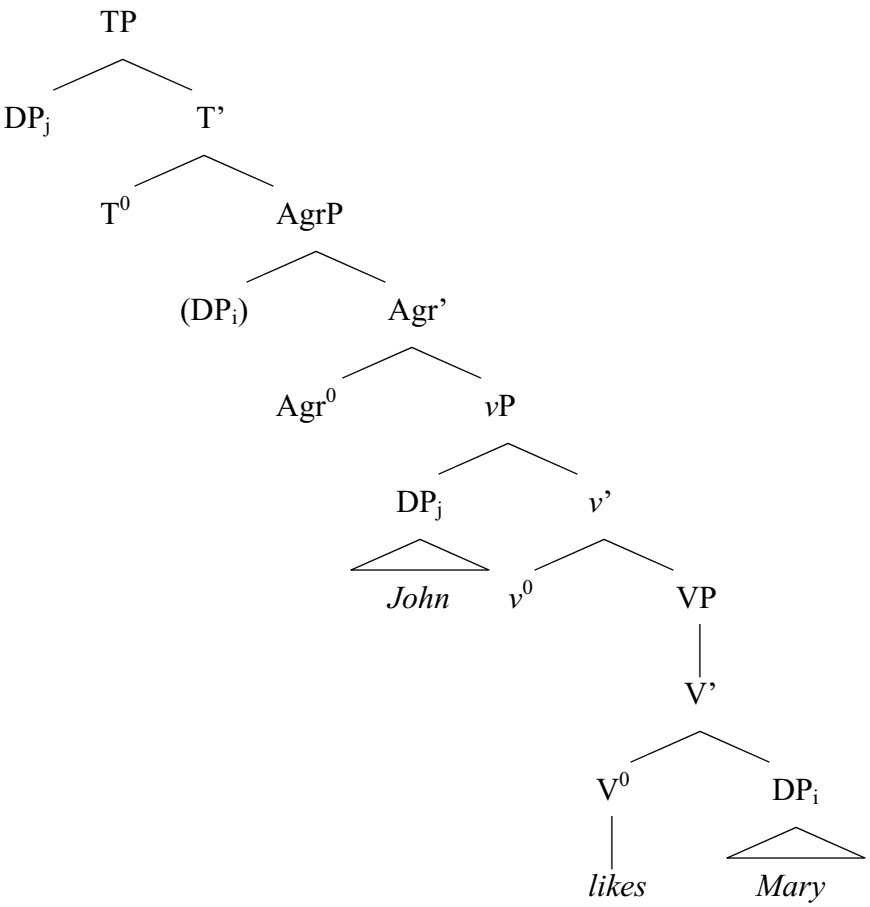
This is a rough guide to the structure of the computation sketched in (8), a modified T-model; I will return to many issues left open here in subsequent chapters.[13]

Under a Bare Phrase Structure theory of X'-structure (see especially section 3.2), we can then take the entire process of deriving *John kissed Mary* along the lines sketched in (9). This is still a simplified (and idealized) derivation, setting aside at least head movement. We can represent this visually in a more appealing way, as in (10), where different copies of one element are coindexed, and complex head formation is ignored (see chapter 2 for details on derivations). I go over both (9) and (10) in more detail below.

(9)
a. LA = {John$_1$, likes$_1$, Mary$_2$, v, T, Agr}
b. Merge *likes* and *Mary*
c. project *likes*
d. Merge *v*
e. project *v*
f. Merge *John*
g. project *v*
h. Merge Agr
i. project Agr
j. Merge T
k. Copy *John*
l. project T
m. Spell Out
n. compute LF and PF
   n1. Copy *Mary*
   n2. Merge *Mary*
o. convergence

[13] This model has recently been further refined, in particular with respect to the application of Spell Out. Uriagereka (1999b) proposes this operation to apply cyclically, just as any other formal operation. The result is a model of "Multiple Spell Out." Chomsky (1998, 1999) extends the minimalist framework in similar directions. I will discuss the issues in more detail once the current proposal is out in the open, in chapter 6. I suggest adopting some version of Multiple Spell Out, differing in its application slightly from Uriagereka's but keeping the spirit.

(10)



Let us consider (9), and by extension (10), in more detail as a way of introducing further properties of the system. The principle of Full Interpretation is the main assessment for any derivation to converge: all operations must be motivated and yield an output legible at all interfaces. One way of conceptualizing this motivation is in terms of formal features, Checking Theory. Each item in the LA has its set of features—in (9), the feature bundle for *John* contains at least [θ], [ϕ], [Case].[14] We take [θ] to be its thematic information which is licensed by *v* (making *John* the "agent" of the sentence; see chapter 2 for a more detailed exposition, and some of the issues involved). [ϕ] is a collection of phi-features, pertaining to [person], [number] and [gender], giving us "3rd person singular masculine" in this case. These, like [Case] (here, nominative) are checked against finite T

[14] As a convention, I display formal features inside brackets throughout this work.

(see Chomsky 1981). *Mary* bears similar features, but its θ-features are licensed by V (making it the "patient") and [ϕ], i.e. "3rd person singular feminine," and [Case], which is accusative, by Agr.[15] Setting further discussion aside, we commonly take the process of feature checking to be overt, i.e. prior to Spell Out, or covert, in the LF component.

To return to (9), the LA consists of all lexical and functional material present for the computation. The indices indicate number of tokens picked. I will return to this at a later point (particularly in sections 2.3.2 and 3.2.4), but it is simply a mechanism to tell different copies apart. Thus, the two instances of *John* in the present derivation are copies of the same token ($John_1$ in the LA); as we will see later, they are non-distinct copies (Chomsky 1995a, Nunes 1995).[16] What I did not indicate is the formation of the heads. We take these to successive-cyclically move up the structure Again, we will amend some of these assumptions, but for the time being, we can take (9) and (10) to be accurate depictions of a (very) minimalist approach to the computation of human language.

1.2.3 Conditions on the Derivation

The above sketch was kept short deliberately. This work is not as much concerned with the technical details of minimalist implementation, but rather presupposes that something along the lines sketched in section 1.2.1 is on the right track. As indicated already, I will not only lay out some necessary details in the course of the presentation (particularly in chapters 2 and 3), but I will also refine others, hence a discussion more detailed than sketched in section 1.2.2 would only lead to a cumbersome presentation.

---

[15] For the function of Agr and other issues related to checking the formal features of (in)direct objects, see section 3.2. For the dissociation of [θ] on the one and [ϕ] and [Case] on the other, see chapters 2 and 3 in general, where I also lay out technical differences between the licensing of Case and agreement.

[16] The same applies for *Mary*, but note that in English, the object arguably moves at LF (indicated by parentheses in (10)).

There are some guiding principles of one version of minimalism which shall play a role throughout, however, pretty much in the form they have been presented in recent work. I focus on three broad issues: economy, locality and checking.[17]

Economy comes in, at least, two guises, what Epstein & Hornstein (1999a: xi) call *methodological economy* and *linguistic* or *substantive economy* (cf. Hornstein 2000: 4-7). The first is the familiar methodology of evaluating theories, Occam's Razor, namely that simplicity and parsimony call for as few primitive relations and additional assumptions as possible—the motto is: the less, the merrier. In the course of this dissertation, I often appeal to this guiding principle, which might not be the correct way to proceed, but all other things being equal, it is one way of teasing apart different paths that reach the same door.

The notion of substantive economy embodies more specific aspects of the computational system. Locality conditions are one core property of syntactic dependencies. Under a substantive, least effort view of economy, we could formulate these conditions in terms of short, rather than long, moves (in the form of the Shortest Move Principle or the Minimal Link Condition). Likewise, a derivation with fewer applications of Copy and Merge is more economical than one with more (in terms of the Fewest Steps Condition or the Shortest Derivation Condition). In addition, these operations only apply when they need to (regulated by the notions of Least Effort or Greed), and any symbol used in a grammatical representation must have its purpose (subject to Full Interpretation). Moreover, the combination of Copy and Merge may rearrange the elements of the computation, or the objects in a phrase marker, but it cannot add new elements, i.e. those that are not part of the LA, in the course of the derivation (ruled out by the Inclusiveness Condition). I take all these conditions to be part of the computational system, and they can arguably classified as virtually conceptually necessary. I suppose the majority, if not all, of these principles are ingredients of any minimalist theory, however formulated in their details.

[17] For further literature, see e.g. Takahashi (1994), Collins (1997), Kitahara (1997, 1999).

With the introduction of formal features, in a Checking Theory, one might also appeal to moving elements in their smallest form, such as features rather than categories, if at all possible. This point raises questions as to the directionality of displacement. In the original framework of Chomsky (1993), displacement was the result of Move, strictly taken the requirement of an element to move to another position. This requirement comes in form of a formal feature F that the element X must check against a matching feature F contained in some higher element Y. If X and Y are in an appropriate licensing configuration, F can be checked (and erased or deleted).[18] In this system, it is a property of the moving element to Move. Turning attention from mover to movee, an alternative formulation of displacement replaces Move by Attract (Chomsky 1995a): the need to check comes from some feature (on Y), attracting a lower element (X) to move into a checking configuration. The empirical differences are very subtle. I do not believe that it matters one way or another for the framework I am going to present, but in the interest of taking sides, a conception of Copy and Merge from the mover's point of view seems to be no less plausible than the other way around. For recent arguments, see Boeckx & Grohmann (2000), Hornstein (2000), among others.[19]

This wraps up our concise introduction to the larger theoretical framework and intuitions guiding it adopted here. In the next five chapters, I will add my own salt and consider (dis)advantages, predictions and consequences of either.

---

[18] I leave out many details, mainly for the reason that they will not play a major role in this work. Formal features come in two strengths, "weak" and 'strong." The latter force overt movement, that is checking or erasure before Spell Out, while the former can wait until LF. Features are further distinguished into +Interpretable and –Interpretable, and as a rule of thumb, all –Interpretable features are invisible at LF. Chomsky also makes use of different types of features, classifying in particular D- and V-features. Again, in the interest of clarity, I shall abstain from a deeper discussion, especially as the finer make-up of formal features and its consequences will not be crucial for us.

[19] Chomsky (1998, 1999) introduces a third alternative, one that takes formal features more seriously than simple triggers for displacement. In this framework, Agree may check and license features across a longer distance than strictly local checking configurations. I will return to a brief discussion in chapter 6.

1.2.4 Some Terminological Conventions

The above mentioned "appropriate checking configuration" is usually taken to include three types of relational configurations in an X'-structure: between a head and its complement (Head-Comp), a head and another head (Head-Head) and a head and its specifier (Spec-Head), where the last one is the canonical configuration for XP-movement. Chomsky (1993) proposes an explicit evaluation of a Checking Domain. In section 3.2.4, I will modify this definition in terms of "Natural Relations," for reasons that will become clear along the way. One important aspect of the theory endorsed here makes a distinction between specifier and adjunct on a number of grounds, in particular it aims to differentiate the two on purely structural grounds.[20] As positions in the phrase marker will be come increasingly important, I adopt these notational conventions:

(11) i. SpecXP identifies the (unique) specifier Spec of a maximal phrase XP
ii. CompVP denotes the (unique) complement Comp of XP
iii. AdjXP stands for the (not necessarily unique) adjunct Adj to XP

I propose one dimension of distinction between specifiers and adjuncts to be relevant for the computation: an element can only be adjoined to XP by base-generation, while a specifier can be formed by movement also. To tease the two notions apart, Merge qua base-generation will often be referred to as Direct Merge (see e.g. Epstein et al. 1998). I will thus not specially label Merge by movement (which could be referred to, on a par, as Re-Merge). When I use "Merge," the difference is not important in that context, and when it is I call one operation Direct Merge.

[20] I often use the term "adjunct" in its purely structural sense, as an object in the phrase marker that is adjoined to another. As will become clear, the distinction between being adjoined to XP (a maximal projection, to be refined in section 3.2.4) vs. X^0 (roughly, a head) is important: the latter is the result of movement, the former is not.

1.3 Lay of the Land

This dissertation explores whether locality conditions should only be formulated in terms of an upper bound on the distance between two positions in a dependency. We will consider hypothetical cases which seem to be ruled out because the movement would not cover a certain amount of distance. I call this lower bound anti-locality and propose a framework that captures this addition to standard cases of locality. Parts of this work are purely conceptual in nature. Other parts are predominantly empirical, and if nothing else, the following analyses constitute alternative approaches to the phenomena under investigation. Ideally, though, the theoretical apparatus introduced will complement the analyses proposed for each of these phenomena and vice versa. One aim of this dissertation is to group together the at first glance very different phenomena investigated here in a natural way, and, as such, present a unified picture of clause structure and syntactic computation.

1.3.1 Theory and Practice

Let me lay out a road map to this dissertation. The content can be divided into two major components, one conceptual-theoretical and one more empirical-theoretical.

The theoretical framework I propose splits the clause into three Prolific Domains, parts of the derivation with particular contribution to the interfaces. I discuss theoretical conditions that are part of the theory: the Condition on Domain Exclusivity, the notion of Copy Spell Out, and a modification of X'-relations allowing us to not only distinguish specifiers from adjuncts, but also maintain uniqueness for specifiers, among other things. I also address some consequences that could follow and predictions the theory makes, in particular with a radical derivational approach to dependency formation (as opposed to a rules of construal, not necessarily opposed to a representational view, as noted in fn. 6).

The empirical parts provide evidence in favour of this tripartition and the theory behind it. Particular phenomena I investigate are reflexivization, left dislocation and Wh-questions. I employ the rescue operation of Exclusivity in the form of Copy Spell Out as regards the derivational introduction of local anaphors as well as certain resumptive elements in left dislocation constructions. Exclusivity will play a major role in teasing apart various properties of Wh-phrases across languages, paying particular attention to multiple Wh-questions and issues arising from movement, additional discourse information etc.

The conceptual discussion, mainly towards the end of the work, ties in Prolific Domains with larger issues, such as language design, the structure of the grammar, and the interaction between the computational system C_{HL} and the LF and PF interfaces. I argue for a cyclic application of the operation Spell Out on the one hand and a derivational feeding of the interfaces on the other, here understood not as levels of representation, but rather as interface components.

1.3.2 Chapter 2: Anti-Locality

In chapter 2, I present what we might call anti-locality, the idea that while movement must be local, it cannot be too local. Basically, I sketch a few derivations which one might expect, given that movement steps tend to be subject to rather strict locality conditions. These pertain to movement from one θ-position to another, from one Case-position to another, and so on. As it turns out, constructions related to these derivations do not seem to exist in natural language, and we might wonder why. I suggest that each of these steps would constitute movement within a "Prolific Domain." There are three such Prolific Domains per clause: (i) the θ-domain, where thematic relations are expressed, (ii) the ϕ-domain, where agreement properties are licensed, and (iii) the ω-domain, where discourse information of the relevant elements is manifested.

Part of the demonstration of anti-locality is the Condition on Domain Exclusivity:

(12) Condition on Domain Exclusivity (CDE)
An object O in a phrase marker must have an exclusive Address Identification AI per Prolific Domain ΠΔ, unless duplicity yields a drastic effect on the output.
i. An AI of O in a given ΠΔ is an occurrence of O in that ΠΔ at LF.
ii. A drastic effect on the output is a different realization of O at PF.

What (12) expresses is that movement within a Prolific Domain is prohibited unless the lower copy is spelled out. This spelled out copy must have a different PF-matrix from the original, and I suggest that languages make a default form available in the guise of pronominal elements. (12) is also compatible with a dynamic conception of the computational system, much in the spirit of Epstein et al. (1998) or, especially, Uriagereka (1999b), in which PF and LF are interpretive "components," rather than "levels," of representation, which are fed cyclically. This issue will be addressed further in chapter 6.

I then lay out initial considerations of the form and function of each Prolific Domain, couched in a discussion of clause structure in general, and of additional proposals regarding phrase structure. I suggest viewing each Prolific Domain as an area of the clause (hence domain) which consists of several projections, as much work over the past decade or so has shown (hence prolific). I sketch a derivational understanding of these without stipulations, justifying each additional assumption that I introduce.

1.3.3 Chapter 3: Copy Spell Out

Chapter 3 contains an empirical testing case for the θ-domain. Extending the idea that thematic roles should be viewed as formal features, I consider reflexivization as a possible instance of movement within a ΠΔ that implements the caveat made available by the CDE. As such we can think of reflexives as the residue of movement, building on and extending recent proposals by Lidz & Idsardi (1998), Hornstein (2000), Kayne (2000). Much of the analyses in the chapters to follow aim to extend this analysis to all three

ΠΔ's which we may sketch in a general format as follows, where ΠΔ stands for any of the θ-, ϕ- or ω-domains (θΔ, ϕΔ and ωΔ, respectively), O for the object moved, and X for the element which is introduced by the repair strategy Copy Spell Out encoded in (ii) of the CDE. In other words, derivational steps like (13b) are excluded, in favour of (13a), with the circled arrow indicating Copy Spell Out:

(13) a. $[_{\Pi\Delta}$ O … ⊖ ⊃ X …]]
b. # $[_{\Pi\Delta}$ O … ⊖ …]]

(13) is the derivation explored for local anaphors—reflexives and reciprocals alike—, corresponding to sentences like the following, where O is the antecedent, X the anaphor and ⊃ the symbol representing Copy Spell Out:

(14) a. John likes himself.
b. John introduces Mary to himself.
c. John put the cards on each other.

I also address the role of inherent reflexives (*John shaved, Mary dressed* etc.), predicates with implicit arguments (such as *the fish ate*) and double object constructions (the alternation of *John introduced himself to Mary* vs. **John introduced Mary herself*). One goal is to identify the nature of X found in anaphoric relations, but also others (see chapter 4).

In the first part of this chapter I discuss the properties of X'-structure, how the relations between the elements that make up the phrase marker can be expressed. Drawing from conceptions of Bare Phrase Structure, I set out to present, illustrate and deduce the following desiderata:

(15) *Specifiers and adjuncts are formally different objects in the phrase marker*
   i. adjunction to XP must be the result of base-generation (Direct Merge)
   ii. specifiers are base-generated or result from Move (Copy plus Merge)
   iii. specifiers enter a checking relation with a head and must be unique
   iv. adjunction cannot check features with a head and need not be unique

In the course of the discussion, I present empirical arguments against multiple specifiers, conceptual problems with the notion of "X'-invisibility," and the "natural relations" that come with a set-theoretic conception of X'-structure. The proposal follows Chomsky (1998, 1999) who points out that the operation Merge yields two relations for free: Sisterhood and Immediate Contain. But unlike Chomsky, I work with the three relations that really result from the first-order composition of these relations. I suggest understanding appropriate checking configurations defined entirely over these relations (Sister, Immediate Contain, Identity, Contain and what I call Extended Sister, the composition of Sister and Immediate Contain; cf. Chomsky 1998: 31).

1.3.4 Chapter 4: Left Dislocation

Chapter 4 deals almost exclusively with left dislocation constructions (LD). By way of introducing topicalization, I address what might be called a variant of topicalization, a topic coreferent with a resumptive pronoun. I concentrate on the Germanic variety that has been identified as Contrastive Left Dislocation (CLD) and then work towards a more general classification of LD-constructions across languages, including Hanging Topic Left Dislocation (HTLD) and Clitic Left Dislocation (CLLD).

HTLD is a type of construction we also find in English, the only LD-variety in this language. It best resembles topicalization, as the HT seems to share all relevant properties with topics. Yet there are some differences which we will explore and then contrast with CLD in German, a language that has both varieties HTLD and CLD.

I concentrate on the debate of movement vs. base-generation, providing evidence that in CLD the CLDed constituent must have moved from a lower position. The relevant domains of evidence are the usual diagnostics of movement: the CLDed XP, but not the HT, matches in Case with the resumptive pronoun RP (in form of a demonstrative), it is sensitive to weak crossover and Condition A, it does not show Condition C effects, the CLDed XP may be an idiom chunk, it obeys islands, and it is restricted to root contexts. Interestingly, a sentence may have at most one CLDed element but allows for multiple HTs, in which case the HTs precede the CLDed XP.

I build on the contrast of the following, presenting a number of arguments that suggest movement of the CLDed XP, but base-generation of the HT:

(16) a. [Seinen$_i$ Vater], den hat jeder$_i$ gern. *(CLD)*
b. [Sein$_i$ Vater], den hat jeder$_{*i/j}$ gern. *(HTLD)*
his.ACC/NOM father RP.ACC has everyone dear
'His$_i$ father, everyone$_i$ likes (*him).'

The general structure of an LD-construction is [LDed XP – RP – V…], so it is safe to assume that the RP is in topic position (given that German is a Verb Second language). This makes XP and RP both part of the same Prolific Domain, namely the ω-domain. XP and RP are not only coreferent, but presumably identical. As such, there is a clear resemblance to reflexives, and I propose a movement analysis of CLD conforming to the CDE in terms of Copy Spell Out. The alternative, and wide-spread standard, approach involves construal of identity in the form of null operators followed by predication.

One aim of the current work is to present an alternative to such devices. To this end I assign a similar derivational procedure to CLD as to local reflexives: the LDed XP is the originally merged argument, undergoes topicalization and subsequent fronting; the last step is illegitimate, unless the copy in topic position spells out (as a pronominal). HT

constructions, on the other hand, involve a base-generated HT and an RP which is the original argument undergoing topicalization. Moreover, following the assumption that specifiers are unique and follow adjuncts which need not be unique, the following picture emerges, where (17a) represents the relevant, abstract derivational history for CLD and (17b) for HTLD (compared with topicalization, cf. (18)):

(17) a. $[_{CP}$ XP C $[_{TopP}$ ~~XP~~ ⊃ RP Top … $[_{\phi\Delta}$ … ~~XP~~ … $[_{\theta\Delta}$ … ~~XP~~ …]]]]
b. $[_{CP}$ YP $[_{CP}$ C $[_{TopP}$ RP Top … $[_{\phi\Delta}$ … ~~RP~~ … $[_{\theta\Delta}$ … ~~RP~~ …]]]]]

(18) $[_{TopP}$ XP Top … $[_{\phi\Delta}$ … ~~XP~~ … $[_{\theta\Delta}$ … ~~XP~~ …]]]]

Next I turn to another variation of LD, CLLD as found across Romance languages, Arabic, and others; I concentrate on CLLD in Greek. Pretty much the same diagnostics have been used to argue for movement of the CLLDed XP. Now that we have the tools to identify the LDed element as the mover and the resumptive as a derivationally introduced element, an obvious extension would be to see whether and how this approach could be extended. I suggest a similar derivational history as for CLD, but with a twist: Copy Spell Out occurs inside the ϕ-domain. Implementing some version of the generation of clitics that base-generates them as argument-DPs inside the θ-domain and moves it to an agreement head, I suggest that the DP undergoes subsequent movement to the specifier of that phrase, arguably too local to be licit. The rescue comes in form of Copy Spell Out in the Agr-head, inserting the clitic. This analysis allows for a straightforward treatment of some differences between CLLD and CLD regarding LDable constituents. This approach allows a straightforward explanation of the properties common to CLD and CLLD, while it also accounts for the differences.

The discussion on LD-constructions empirically supports the structural desiderata in (15). I show that contrary to popular belief, German allows more than one LDed con-

stituent. Crucially, however, the CLDed XP (derived by movement) is unique and any HTs (base-generated) must precede the CLDed XP. Moreover, as the CLLDed XP is identified to sit in a recursive topic projection, stacking of these constituents is predicted and just as easily accommodated. The resulting structures look as follows, simplified and in abstract representation:

(19) a. $[_{CP}$ **AP** $[_{CP}$ **BP** $[_{CP}$ **XP** C $[_{TopP}$ ~~**XP**~~ ⊃ **RP** **V**-Top $[_{TP}$ … **RP**$_{AP}$ … **RP**$_{BP}$]]]]]

b. $[_{TopP}$ **ZP** Top $[_{TopP}$ **YP** Top $[_{TopP}$ **XP** Top … $[_{TP}$ … **CL**$_{XP\text{-}YP\text{-}ZP}$-V …]]]]

These representations correspond to well-formed constructions in German ((19a), where AP, BP are adjoined HTs, XP is a derived CLDed element) and Greek ((19b), where XP, YP, ZP are all derived CLLDed constituents).

1.3.5 Chapter 5: Wh-Questions

In chapter 5, a comprehensive chapter on issues of (multiple) Wh-questions, I concentrate on questions with two Wh-phrases. The main claim is that German fronts both Wh-elements in such instances. Motivation comes from discourse restrictions underlying the felicity of asking a multiple Wh-question, empirical evidence comes from apparent "intervention" effects and typological evidence comes from cross-linguistic (non-)availability of single-pair readings.

Regarding the interpretation of multiple Wh-questions, Hagstrom (1998) and Bošković (1998b) provide a syntactic and semantic analysis for the observation that languages differ with respect to the availability of single-pair (SP) and pair-list (PL) readings. A general tendency is that Wh-in situ languages such as Japanese, Chinese, Hindi or one variant of French multiple Wh-question formation allow the SP reading in instances in which it is not available to singular Wh-moving languages such as English, the other

French variant and many other languages. The latter, however, allow a single-pair reading when the Wh-phrases are D-linked—but only if the originally higher Wh-phrase is fronted. Then there are multiple Wh-fronting languages such as Bulgarian or Romanian that never allow SP interpretation in multiple Wh-questions. Boškovič restricts Hagstrom's understanding of an abstract question-morpheme [Q]—the element that derives interrogative force in all languages, thus dispensing with LF-movement of Wh-elements (cf. Baker 1970)—as follows. [Q] may be merged in two positions: either low, together with the lower Wh-phrase, which results in PL readings, or in a high position, above both Wh-phrases, potentially yielding SP interpretation; the details follow from a decompositional semantics of Wh-questions explored by Hagstrom, and their mapping to syntactic structure. The relevant case is the latter. If [Q] is merged high and a Wh-phrase moves over it, a relativized minimality violation occurs and destroys the SP reading. As a result, Serbocroatian, which allows SP interpretation, is like a Wh-in situ language in that it does not front either Wh-phrase over [Q], while German, behaving like Bulgarian, moves both Wh-phrases into the C-domain. The result of this extension of Boškovič's proposal is a clean tripartition among languages into Wh-in situ, singular Wh-movement and multiple Wh-movement languages.

For German I suggest that the lower Wh-phrase also moves into the C-domain. I identify this position as the specifier of FP. The claim that German is like Bulgarian in that it moves all Wh-phrases into an articulate COMP (our ω-domain) makes a number of immediate predictions. One consequence of this structural identification is that only topicalizable material may intervene between the two Wh-phrases. This prediction is borne out in the intervention effects studied by Beck (1996a, 1996b). Rather than reintroducing the concept of barriers into the framework, these data can be captured as follows. Quantificational elements that may appear in between two Wh-phrases must be monotone increasing; decreasing ones are out. We have independent evidence that the latter cannot be topicalized, and if a (recursive) TopP is the only projection between FocP, the locus of

the higher Wh-phrase, and FP, the locus of the lower one, the so-called "Beck effect" is predicted and the data discussed by her (and more) are accounted for. The same rationale has recently been applied to similar effects in Japanese and Korean from a semantic point of view (Lee & Tomioka 2000). A number of relevant cases in the left periphery of the clause will be studied under a new light.

But German also shows a discourse restriction on the well-formedness of multiple Wh-phrases in a question, distinguishing it from English, for example. In German, a multiple Wh-question is only felicitous if potential values for each Wh-phrase (taken to be a set of individuals) has been introduced in the discourse. This contextual condition, called here "Discourse-Restricted Quantification" (Grohmann 1998), can be seen as a special type of D(iscourse)-linking (cf. Pesetsky 1987, Comorovski 1996). In a sense, then, all Wh-phrases in a multiple Wh-question must be D-linked.

This odd conclusion actually ties in very nicely with an observation of Italian, due to Rizzi (1978) and further explored by Calabrese (1984, 1992): multiple constituent questions in Italian (dialects) are not acceptable. Calabrese connects this property of Italian with another, namely that information focus is unique. Given that information focus is the property of an appropriate answer to a Wh-question, identifying the set of individuals inquired, a unique focus position implies only unary Wh-questions. I extend this account to German and show that here, too, information focus in unique. The result of this chapter is then that German behaves like Bulgarian in that both languages obligatorily move all Wh-phrases into the ω-domain. But it also behaves like Italian in that multiple Wh-questions do not really exist. What allows German the formation of these questions is the independent fact that as a scrambling language, German word order is driven largely by discourse factors.

While specific aspects of the proposal are tentative in nature and deserve more study, I will explore some of the consequences within the current framework in this chapter as well as chapter 6.

1.3.6 Chapter 6: Interfaces

Chapter 6 first reviews the proposals and results gathered in the previous chapters. I then address conceptual issues pertaining to the theory of Prolific Domains and Exclusivity, putting it into perspective with much recent work. This part contrasts the notion of Prolific Domain with that of a phase (Chomsky 1998, 1999), suggesting that while similar in spirit, the two are different enough to motivate the present framework, and, all other things being equal, both are compatible (see also Uriagereka & Martin 1999). I also discuss the notion of Address Identification from the definition of the CDE in (1) and suggest that each Prolific Domain ships the information up to that point in the derivation into Spell Out—to the PF and LF interfaces (cf. Uriagereka 1999b). An address is understood as a marker in the derivation relevant for subsequent computation (see also Vergnaud 1985, Manzini 1992, Uriagereka 1997). I will explore an implementation of the notion "Multiple Spell Out" (Uriagereka 1999b), the cyclic application of the operation Spell Out (see also Chomsky 1998, 1999). In this larger discussion, I also compare the current tripartition with a similar proposal made recently by Platzack (in press).

This view allows for an interesting twist on successive-cyclic movement, a desired operation, but hard to motivate in a strict minimalist program for reasons of economy. I suggest that movement within a Prolific Domain is not only banned (unless repaired by Copy Spell Out), but movement across domains is restricted in the following sense. In the interest of Full Interpretation, clause-internal movement must always target the next highest Prolific Domain, but movement across clausal boundaries has to target the corresponding domain in the next higher clause. Apart from motivating this condition, I also illustrate some welcome consequences. A derivational approach to many construal relations endorsed in recent work can be maintained (Nunes 1995, Boeckx 2000b, 2000c, Hornstein 2000 and others) and the Extended Projection Principle can really be dispensed with (cf. Castillo, Drury & Grohmann 1999, Epstein & Seely 1999, Boškovič 2000).

I am perfectly aware that the framework presented here leaves many questions and issues unaddressed. However, it also allows us to ask new questions, especially with respect to "anti-locality"—that it makes sense to consider not only the upper but also the lower bound of locality conditions. This work, then, tackles primarily this issue by laying out a concrete framework, including as many necessary considerations as possible, which is still kept rough enough to allow working out some of the issues in different ways. In this sense, an approach more committed to the general issue derivation vs. representation, for example, would define some of the concepts presented here in slightly different terms. This, in turn, will have consequences for the many phenomena that will not be discussed here, despite their relevance. In order not to bias the empirical outcome of such analyses, and in order to present the theory in a hopefully coherent way, I focus on a rather narrow set of phenomena that illustrate the ideas. Another aspect is one of similar generality: while I frequently cite literature that interprets the minimalist framework in a particular way, nothing hinges on the outcome of any of those modifications. Thus, whether control is really best analysed in terms of movement or whether null operators do really not exist plays no role for the impact of the theory proposed here. Rather, the tripartition of the clause into Prolific Domains and the lower bound locality condition of Exclusivity are genuine explorations of a generative approach, couched in a minimalist line of inquiry, and should be considered as such. The conceptual motivation I offer at times to go one way or another may come in various guises; the theoretical detail and empirical coverage I suggest are a whole different story.