PROLIFIC PERIPHERIES:
A RADICAL VIEW FROM THE LEFT

by

Kleanthes K. Grohmann

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2000

Advisory Committee:

Professor Norbert Hornstein, Chair
Professor Werner Abraham
Professor David Lightfoot
Professor Paul Pietroski
Professor Georges Rey
Professor Ian Roberts
Professor Juan Uriagereka
Copyright 2000

by

Kleanthes K. Grohmann
ABSTRACT

Title of dissertation: PROLIFIC PERIPHERIES:
A RADICAL VIEW FROM THE LEFT
Kleanthes K. Grohmann, Doctor of Philosophy, 2000

Dissertation directed by: Professor Norbert Hornstein
Department of Linguistics

This dissertation concentrates on movement dependencies under minimalist assumptions and suggests that the opposite of moving too far (standard locality) exists also, namely moving too closely: anti-locality. A theory of anti-locality will be developed that incorporates a tripartition of clause structure into Prolific Domains, areas of the clause that share a common “super-feature” or context value. The lowest Prolific Domain is responsible for thematic relations (θ-domain), the next one for agreement relations (φ-domain), and the highest one for discourse relations (ω-domain). Under this approach, anti-locality is expressed in terms of the Condition on Domain Exclusivity which simply bans Domain-internal movement. This ban is motivated on grounds of PF-legibility, much as deleting lower copies is driven by PF-requirements.

However, there are constructions which contain two elements that can be argued to be related by movement, the lower one pronominal in nature. The proposal is that in these cases, the lower copy of the movement dependency cannot simply be deleted, as in other instances of movement, but must be spelled out to yield a drastic effect on the (PF)
output. This strategy is referred to as Copy Spell Out and a number of potentially relevant and at first glance unrelated phenomena will be investigated (in particular, reflexivization and certain types of left dislocation).

Considering another empirical domain, one particular aspect of a tripartite clause structure can be applied to variation found in the syntactic formation and semantic interpretation of multiple Wh-questions across languages. The main thesis of this discussion is that German seems to behave syntactically like Bulgarian, in that it obligatorily moves all Wh-phrases into the ω-domain, while at the same time it also resembles Italian in not actually being able to ask a well-formed multiple constituent question to begin with. Rather, the two Wh-phrases are always D-linked and hence topic-like, a hypothesis that fits well with the overall framework.

The framework laid out here challenges some technical assumptions in current theory, yet maintains a distinct minimalist spirit in its approach. Moreover, the theory developed is independent of particular choices in implementing the syntactic tools made available in the minimalist program.
DEDICATION

Vor sechs Jahren ging es los
Es war wie eine Revolution
Ziel und Richtung unbekannt
Das Alte wurde wieder bekannt

from Wehende Fahnen, Die Toten Hosen, 1983

In loving memory of Athena, Helmuth, Martin and Willi
ACKNOWLEDGMENTS

I remember clearly that sunny (!) day in Bangor, walking up Penrallt and reading Gert Wehelhuth’s beautiful acknowledgements in his *Syntactic Saturation* book (basically verbatim from his dissertation) That was the first time I read this little section and swore to myself right there that one day I would write acknowledgements for a doctoral dissertation myself.

It’s hard to believe but this day has come. I have dreamed about this moment for so long, I already had all the excitement one could wish from acknowledgements, and after a hectic month or two I’m brain dead anyway. In my mind, Gert’s still remains the best, and in the meantime I have had the pleasure of reading many others, some of them pretty good. I don’t have the energy to compare my own attempts and rather than qualifying and apologizing—something I’ve sworn to drop now that I am a doctor—I’ve decided to change the format a little. I’ve had the pleasure of writing a long acknowledgements section in my BA thesis, and then again in my generals paper, and I know that some people make fun of this fact to this day. I also know that many people are anxiously waiting to see how many pages I would fill this time.

I am incredibly grateful to so many people. If you have sat down and discussed my work with me over the years, I am grateful for your input. If you have sat down with me and talked about your own work, I am grateful for sharing it with me. If you have sat down and just chatted about linguistics, I am grateful for the exchange. If you just had a pint with me, or a coffee, shared a smoke or made me laugh or laughed with me or at me or if you did anything else pleasurable with me, for me or against me, I am grateful. I am grateful for my friends and my enemies, and above all my family, for encouragement, stimulation, support and love. You know who you are and you know that I love you.

Thank you everyone. Thank you world. Thank you Jonathan Walker and friends, thank you Alfred Dunhill for 865 good times, and thank you R.J. Reynolds for 24/7. Thank you punkrock, nostalgia and everything that comes with it! Oh Joy! Danke.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>1 Prolific Peripheries: By Way of Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Locality Conditions</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Movement and Locality</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2 Construal and Locality</td>
<td>7</td>
</tr>
<tr>
<td>1.1.3 Local Summary</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Minimalist Inquisitions</td>
<td>12</td>
</tr>
<tr>
<td>1.2.1 The Structure of the Grammar</td>
<td>12</td>
</tr>
<tr>
<td>1.2.2 The Computational System</td>
<td>14</td>
</tr>
<tr>
<td>1.2.3 Conditions on the Derivation</td>
<td>17</td>
</tr>
<tr>
<td>1.2.4 Some Terminological Conventions</td>
<td>20</td>
</tr>
<tr>
<td>1.3 Lay of the Land</td>
<td>21</td>
</tr>
<tr>
<td>1.3.1 Theory and Practice</td>
<td>21</td>
</tr>
<tr>
<td>1.3.2 Chapter 2: Anti-Locality</td>
<td>22</td>
</tr>
<tr>
<td>1.3.3 Chapter 3: Copy Spell Out</td>
<td>23</td>
</tr>
<tr>
<td>1.3.4 Chapter 4: Left Dislocation</td>
<td>25</td>
</tr>
<tr>
<td>1.3.5 Chapter 5: Wh-Questions</td>
<td>28</td>
</tr>
<tr>
<td>1.3.6 Chapter 6: Interfaces</td>
<td>31</td>
</tr>
<tr>
<td>2 Anti-Locality: Prolific Domains and Exclusivity</td>
<td>33</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>33</td>
</tr>
<tr>
<td>2.2 Anti-Locality</td>
<td>34</td>
</tr>
<tr>
<td>2.2.1 Anti-Locality in Theta Dependencies</td>
<td>35</td>
</tr>
<tr>
<td>2.2.2 Anti-Locality in Agreement Dependencies</td>
<td>42</td>
</tr>
<tr>
<td>2.2.3 Anti-Locality in Discourse Dependencies</td>
<td>46</td>
</tr>
<tr>
<td>2.2.4 Summary</td>
<td>53</td>
</tr>
<tr>
<td>2.3 Towards a Theory of Anti-Locality</td>
<td>53</td>
</tr>
<tr>
<td>2.3.1 Prolific Domains</td>
<td>54</td>
</tr>
<tr>
<td>2.3.2 Exclusivity</td>
<td>58</td>
</tr>
<tr>
<td>2.4 On the Internal Architecture of Prolific Domains</td>
<td>65</td>
</tr>
</tbody>
</table>
5.5 German = Italian: Categorical Absence of Multiple Wh-Questions? .......... 268
  5.5.1 Discourse-Restricted Quantification ........................................... 268
  5.5.2 Italian Wh-Questions ................................................................. 270
  5.5.3 Speculations: Discussion, Consequences and Predictions ............... 275
5.6 Conclusion ....................................................................................... 279
6 Picking up the Pieces: A Radical View from the Left .................................. 281
  6.1 Prolific Domains: The Framework ...................................................... 281
  6.2 Syntax at the Interfaces: A Sketch .................................................... 284
  6.3 Exclusivity, (Anti-)Locality and Cyclicity ......................................... 292
  6.4 Concluding Remarks ....................................................................... 300
Bibliography .......................................................................................... 301
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>236</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>239</td>
</tr>
</tbody>
</table>
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
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. 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.


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 seems [t_i to t_i like Mary].
b. John 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), 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. Who₁ does John like tᵲ?
    b. Who₁ does Bill believe [tᵲ (that) Peter said to Jane [tᵲ (that) John likes tᵲ]]?
    c. * Who₁ does Bill believe [whoᵲ Peter said to tᵳ [tᵲ (that) John likes tᵲ]]?

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, 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. * Who₁ does Bill believe [DP the claim [CP t₁ that [TP John likes t₂]]]? 
b. * Who₁ does John read [DP the book [CP (t₁) which Mary gave t₂]]? 
c. * Who₁ was [DP a friend of t₁] known to kiss Bill? 
d. * Who₁ was [CP t₁ that John likes t₂] 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 than who).

---

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).
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):

\[(4)\]
\begin{itemize}
  \item a. Can\(_t\) Bill \(_t\) have kissed Mary?
  \item b. * Have\(_t\) Bill can \(_t\) kissed Mary?
  \item c. * Kissed\(_t\) Bill can have \(_t\) Mary?
\end{itemize}

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 \textit{can} is closer to C, the locus of inversion, than \textit{have} or \textit{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):

\[(i)\]
\begin{itemize}
  \item * Who\(_t\) did Bill kiss Mary [(\(_t\) after John told him (\(_t\) that he liked \(_t\))]?
\end{itemize}
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$_{ij}$.  
  c. * He$_i$ thinks that Mary/himself$_{ij}$ 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\textsubscript{i} thinks that Mary likes himself\textsubscript{i}.  
   b. John\textsubscript{i} thinks that Mary likes him\textsubscript{i}.  
   c. Jane said that Mary really likes him\textsubscript{i}. John\textsubscript{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\textsubscript{i} hopes PRO\textsubscript{i,j} to kiss Mary.  
   b. * John\textsubscript{i} hopes Jane\textsubscript{j}/Bill\textsubscript{k}/him/himself\textsubscript{i} to kiss Mary.  
   c. John\textsubscript{i} hopes that Mary\textsubscript{j} wants PRO\textsubscript{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 elimi-
nation 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 reformu-
lating 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 them-
ematic 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, Boëcković (1994, 2000b), Lasnik (1995c), Nunes (1995), Boëcković & Takahash-
i (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 depend-
encies 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-
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. 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-
ditions, and sketch an economy-driven approach to the computational system. 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. The old T-model can then be replaced by the following model, representing a rough structure of the grammar:

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 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...
with another, and the same operation applies iteratively (on the thus constructed objects of the phrase marker). 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).\footnote{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).}

\subsection*{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.

\footnote{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.\footnote{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.}

Under a Bare Phrase Structure theory of $X'$-structure (see especially section 3.2), we can then take the entire process of deriving $\text{John kissed Mary}$ 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)  
\begin{enumerate}
  \item $\text{LA} = \{\text{John}_1, \text{likes}_1, \text{Mary}_2, \text{v}, \text{T}, \text{Agr}\}$
  \item $\text{Merge likes and Mary}$
  \item $\text{project likes}$
  \item $\text{Merge v}$
  \item $\text{project v}$
  \item $\text{Merge John}$
  \item $\text{project v}$
  \item $\text{Merge Agr}$
  \item $\text{project Agr}$
  \item $\text{Merge T}$
  \item $\text{Copy John}$
  \item $\text{project T}$
  \item $\text{Spell Out}$
  \item $\text{compute LF and PF}$
    \begin{enumerate}
      \item $\text{Copy Mary}$
      \item $\text{Merge Mary}$
    \end{enumerate}
  \item $\text{convergence}$
\end{enumerate}
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 $[\theta], [\phi], [\text{Case}]$. We take $[\theta]$ 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). $[\phi]$ 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). 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.

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 makeup 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. As positions in the phrase marker will become 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 $\theta$-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 $\theta$-domain, where thematic relations are expressed, (ii) the $\phi$-domain, where agreement properties are licensed, and (iii) the $\omega$-domain, where discourse information of the relevant elements is manifested.

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

An object $O$ in a phrase marker must have an exclusive Address Identification $AI$ per Prolific Domain $\Pi \Delta$, unless duplicity yields a drastic effect on the output.

i. An $AI$ of $O$ in a given $\Pi \Delta$ is an occurrence of $O$ in that $\Pi \Delta$ 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 $\theta$-domain. Extending the idea that thematic roles should be viewed as formal features, I consider reflexivization as a possible instance of movement within a $\Pi \Delta$ 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:

\[
(13) \quad \begin{align*}
\text{a.} & \quad [\text{O} \ldots \text{O} \text{➲ X} \ldots] \\
\text{b.} & \quad \# [\text{O} \ldots \text{O} \ldots]
\end{align*}
\]

(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:

\[
(14) \quad \begin{align*}
\text{a.} & \quad \text{John likes himself.} \\
\text{b.} & \quad \text{John introduces Mary to himself.} \\
\text{c.} & \quad \text{John put the cards on each other.}
\end{align*}
\]

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:
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 rela-
tions” 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 rela-
tions that really result from the first-order composition of these relations. I suggest under-
standing appropriate checking configurations defined entirely over these relations (Sister,
Immediate Contain, Identity, Contain and what I call Extended Sister, the composition of

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 topicaliza-
tion, 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 prop-
erties 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, Vater], den hat jeder, gern. \hspace{1cm} (CLD)  
   b. [Sein, Vater], den hat jeder\(_{ij}\), gern. \hspace{1cm} (HTLD)  

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 \(\omega\)-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)):

\[
(17) \begin{align*}
\text{a. } [\text{CP} \, \text{XP} \, \text{C}] & \quad [\text{TopP} \, \text{XP} \, \supset \, \text{RP} \, \text{Top} \, \ldots \, [ \ldots \, \text{XP} \, \ldots \, [ \ldots \, \text{XP} \, \ldots ] ] ] \\
\text{b. } [\text{CP} \, \text{YP} \, [\text{CP} \, \text{C}] & \quad [\text{TopP} \, \text{RP} \, \text{Top} \, \ldots \, [ \ldots \, \text{RP} \, \ldots \, [ \ldots \, \text{RP} \, \ldots ] ] ]
\end{align*}
\]

\[
(18) \quad [\text{TopP} \, \text{XP} \, \text{Top} \, \ldots \, [ \ldots \, \text{XP} \, \ldots \, [ \ldots \, \text{XP} \, \ldots ] ] ]
\]

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 \( \phi \)-domain. Implementing some version of the generation of clitics that base-generates them as argument-DPs inside the \( \theta \)-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:

\[(19)\]

\[\begin{align*}
\text{a. } & \left[ \text{CP AP CP BP CP XP } \overset{\text{TopP}}{\Rightarrow} \text{RP V-Top [TP ... RP}_{\text{AP}} \ldots \text{RP}_{\text{BP}}]\right] \\
\text{b. } & \left[ \text{TopP ZP Top } \text{TopP YP Top } \text{TopP XP Top } \text{Top ... } \text{CL}_{\text{XP-YP-ZP-V}} \ldots \right] 
\end{align*}\]

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. Bočković 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 Bočković’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 & Tomiooka 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, Bošković 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.
CHAPTER 2

ANTI-LOCALITY:
PROLIFIC DOMAINS AND EXCLUSIVITY

The previous chapter started off with a discussion of locality issues. The point was made that the licensing of two related positions, whether construed or derived, must be local. I also hinted at the suggestion that some configurations might be thought of as too local—or anti-local—and that some condition, one on anti-locality, might bar them. In this chapter I propose the theory of anti-locality in terms of Prolific Domains and the Condition on Domain Exclusivity. I explore both further in the subsequent chapters.

2.1 Introduction

This chapter deals with anti-locality. In section 2.2 I present the basic set of data that might lead us to consider something like “anti-locality” in the syntactic computation \( C_{HL} \). I split the discussion into three sections: one concerning thematic relations within the VP-complex, one concerning agreement properties found inside INFL, and one concerning discourse information expressed in COMP. I explore a formal tripartition of clause structure as a core property of grammar along these lines in section 2.3. I introduce the concept of Prolific Domain and consider empirical support for this division. We can
identify three Prolific Domains in the clause, namely the theta, the agreement, and the discourse domain (roughly corresponding to the layers dominated by vP, TP and CP, as discussed in section 2.2). The Condition on Domain Exclusivity (CDE) is the theoretical focus relevant to anti-locality, conceived as a condition that bars an object in a phrase marker from multiple occurrence in a simple Prolific Domain; also, it comes with a PF-driven repair strategy for spelling out copies, which can save violations of the CDE. The latter property will be the topic of the next two chapters. In this chapter I discuss the first part of the CDE and apply the condition on the data presented in the first section. Section 2.4 thus discusses the finer architecture of each Prolific Domains and integrates the purported anti-local movements. I conclude in section 2.5, paving the road to tackle the second part of the CDE in the next chapter.

2.2 Anti-Locality

Locality restrictions force dependencies to be licensed locally, whether we talk about movement, as in the case of raising, for example, or apparent construal involving two different, yet coreferent objects, as in instances of binding. Some of these conditions ask for a closer local relation than others (such as subject-to-subject raising vs. reflexivization). The following discussion suggests a converse understanding of locality, namely that such dependencies cannot be computed too closely. I briefly sketch particular movement steps that appear to be illicit and follow a potential train of thought that will allow us to draw some interesting conclusions and predictions in the next section. There I provide an understanding of Prolific Domains and a working definition of anti-locality. I then expound some technical issues and address potential complications in section 2.4.
2.2.1 Anti-Locality in Theta Dependencies

Recent work on thematic relations and Checking Theory suggests that the Theta Criterion may not hold, at least not the way it has been used since its conception (Freidin 1978, Chomsky 1981). In particular, it has been argued that Theta Theory as a whole is an unwanted remnant of D-structure in a model of grammar that aims to dispense with superfluous levels of representation, namely those that are not virtually conceptually necessary (cf. section 1.3). Alternatively, we could implement checking-theoretic notions to thematic relations and treat theta roles as formal features that need to be checked in the appropriate configurations (Bošković 1994, Lasnik 1995b, Bošković & Takahashi 1998, Hornstein 1998, Manzini & Roussou 1999, 2000). This implementation allows the operation (Direct) Merge to be, no less than Move, treated as greedy in the sense that merging any element into a phrase marker must check a feature to be licit.

One concrete case which could profit from such considerations is control. Illustrating with Hornstein’s (1999) analysis,\(^1\) (1c) is a sample derivation for a typical control structure, in contrast to the traditional (1b), involving PRO (regardless of whether PRO is also the thematic subject or simply inserted into SpecTP at the relevant time step):

\(^1\) Independently, Manzini & Roussou (1999, 2000) develop basically the same intuition as Hornstein does, though the details are slightly different. Note also that while subjected to criticism, the Hornstein-line has been pursued and extended to capture facts from other constructions, and across languages, in much recent work. Castillo, Drury & Grohmann (1999) aim to get rid of the EPP altogether, Aoshima (2000) applies the ideas to control structures in Japanese, Kiguchi (2000) manages to shed some light on the PRO-gate phenomenon, Matsuya (2000) discusses, among other things, gerunds in Japanese under a similar set of assumptions, Motomura (2000) derives the intricate properties of the Japanese reflexive zibun, Pires (2000a) proposes a novel structure for clausal gerunds, employing arguments from the role of control (Pires 2000b) and ties in the methodology of the Hornstein approach with inflected infinitives (Pires, to appear), Rodrigues (2000) identifies thematic null subjects in Brazilian Portuguese as the residue of movement, San Martin (2000) captures obviation effects in Basque which have long resisted an analysis (see also Hornstein & San Martin 2000).
(1)  

a. John hopes to kiss Mary.

b. \([TP \text{John}, T \[TP \text{John hopes } [TP \text{PRO, } T \[vP \text{(PRO) kiss Mary}]]]\]

c. \([TP \text{John } T \[TP (\text{John}) \text{ hopes } [TP (\text{John}) \text{ T } [vP \text{ John kiss Mary}]]]\]

Such a derivational approach to control moves one element from a \(\theta\)-position of the embedded clause to a corresponding \(\theta\)-position in the matrix. The difference between raising and control then boils down to whether the matrix subject bears one or two \(\theta\)-roles. A controller checks two, derived by movement from one to another \(\theta\)-position, all things being equal—such as additional conditions to rule out unwanted derivational operations, or regardless of whether the controller moves via an embedded SpecTP.\(^2\)

Leaving further discussion aside for the moment, the obvious question one might pose—if \(\theta\)-roles really are formal features and if movement into \(\theta\)-positions is allowed in certain instances—is why arguments do not seem to move from one \(\theta\)-position to another in simple constructions, i.e. non-control environments.

One striking property of the relevant operation in (1c) is that the two copies of the element bearing two \(\theta\)-roles are identical.\(^3\) By analogy, we could imagine other circumstances that involve identity of sorts to be derivable in this way. Looking back at early generative analyses, one might ask why Lees & Klima’s (1963) rule of reflexivization has not been developed in minimalist terms. We could thus consider an analysis that links local reflexives and their antecedents derivationally, where an argument could be merged into the phrase marker and move to yield reflexive interpretation. The lower copy then

---

\(^2\) The questionable intermediate position is indicated by the parentheses; see Castillo, Drury & Grohmann (1997, 1999), Epstein & Seely (1999), Grohmann, Drury & Castillo (2000), Boković (2000), Nasu (2000), and others, for theoretical and empirical arguments against this intermediate position. I will return to these issues later, especially in section 6.3 where I suggest that it is present in control, but not in raising structures.

\(^3\) Under the traditional approach (1b), this identity is encoded into PRO by coindexation with the controller.
deletes, as in other cases of movement, and the dependency created by movement involves one phonetically realized DP with two θ-roles.4

Recall from (1c) that at one point in the derivation, John moves to a θ-position, having saturated (or checked) one θ-role (θ-feature) already, and the lower copy is deleted. The final product John has two θ-roles, but only one copy gets pronounced. We could envision a very similar derivation for reflexivization. This does not turn out to be the case, though. Concretely, we do not say (2a) to express (2b),5 where the relevant elements are boldfaced, with a relevant derivational history roughly sketched in (3).6

(2)  
a.  * John likes.
b.  John likes himself.

(3)  
\# [\_p John \_v \[\_p \_v \_likes \(\_\_\_\_\_\) \\
[\_AG] \quad ([\_TH])]

An element with one θ-role such as THeme cannot move to the next higher position and check a further θ-feature, such as AGent. However, if the only lexical items in the lexical array LA were John and like, it is not clear what would rule out (3), as it conforms to locality (in terms of the MLC or any version thereof)—unless moving into θ-

4 Hornstein (2000) tackles precisely this question and argues for an analysis of reflexives very similar in spirit to the one I propose in the next chapter. Lidz & Idsardi (1998) suggest an alternative way to derive reflexivization rather than construe it. Both approaches will be discussed in chapter 3.

To link antecedent and anaphor derivationally, “covert reflexive movement” has been proposed, that moves the anaphor to its antecedent at LF, or at least adjoins it to the head governing the subject (cf. Lebeaux 1983, Chomsky 1986b, 1993). This approach is very different in technical and conceptual details from the one I explore in this work, or the ones proposed by Hornstein and Lidz & Idsardi.

5 I address some obvious exceptions, such as Mary dressed or John shaved and others, in section 3.4.3.

6 As a convention, I indicate all intended derivations that violate anti-locality with the hash mark ‘#’.
positions is ruled out by force or stipulation, or unless some independent constraint on $C_{HL}$ applies.\(^7\) As I will be going over similar cases in other areas of the clause, the second option has some natural appeal, and this derives anti-locality (section 2.3).

Neither do we, nor could we, employ (4a) to mean (4b), with a relevant derivation in (5), analogous to (3) above:

(4)  
  a.  *John introduced Mary (to).  
  b. John introduced Mary to himself.

(5)  # [\(vP\) John introduced-\(v\) [\(vP\) Mary introduced (to) John] ]  
  [AG] [TH] [GO]

Note that if movement into $\theta$-positions is accepted in general, there could not be a condition that forces the second $\theta$-role to be of the same type as the first. In other words, the reason why the steps indicated in (3) are bad cannot be that they involve two different $\theta$-roles, [AG] and [TH]. Hornstein provides cases which do involve two different $\theta$-roles, the most obvious being object control (Hornstein 1999). Consider a construction like (ia) with a rough representation sketched in (ib), the interesting copies boldfaced:

(i)  
  a. John persuaded Bill to leave.  
  b. [\(TP\) John T [\(vP\) John persuaded-\(v\) Bill [\(TP\) to-T [\(vP\) Bill leave-\(v\)]]]]  
  [AG] [TH] [AG]

Under a movement approach to such structures, Bill checks two $\theta$-roles in (i). This DP starts out as the subject of the embedded clause and ends up as the direct object of the matrix clause; it bears [TH] as the controller and [AG] as the controllee. (The highest copy of Bill also needs to check accusative case in the matrix clause, not indicated here; see Hornstein (1999: 83ff.) for discussion. I leave Case out of the discussion entirely, especially with respect to the embedded position; Grohmann, Drury & Castillo (2000) address this issue.)

Another example are certain instances of adjunct control, discussed by Hornstein (1998):

(ii)  
  a. John asked Bill to leave after being kissed (by Mary).  
  b. [\(TP\) John T [\(vP\) John asked-\(v\) Bill [\(TP\) to-T [\(vP\) Bill leave-\(v\)]]]]  
  [AG] [TH] [AG]

In (ii), John starts out as the thematic object of the passivized verb inside the adjunct, here unlabelled and unattached, and ends up as the subject of the matrix clause; it checks [TH] first, then [AG]. (Movement out of adjuncts is an instance of sidewards movement, which I will not discuss here; see Nunes (1995), Hornstein (2000).)
In this case, the movement to be ruled out is from the GOal to the agent position. The parenthesized preposition indicates that it is not so clear what the exact underlying structure of ditransitive constructions is. In all likelihood, the indirect object introduced by a preposition is part of a more complex structure (cf. den Dikken 1995a, Baker 1997) for reasons that will become apparent in due time.

On the same token, (6a) does not mean (6b), but is ungrammatical, just as the other relevant a-examples above:

(6) a. * John introduced Mary (to).
b. John introduced Mary (to) herself.

(7) # [vP John introduced-v [vP Mary introduced (to) Mary]]
   [AG] [TH] [GO]

The same applies to (8), where the intended anaphor is a reciprocal, which is similar enough to reflexives to consider here, as it involves two different elements (hence two different θ-roles) bearing more or less the same reference—not identical but close enough, so that it could be a possible candidate for a movement dependency:

(8) a. * John put the cards (on).
b. John put the cards on each other.

(9) # [vP John put-v [vP the cards put (on) the cards]]
   [AG] [TH] [GO]

---

8 I ignore double object constructions for the time being, which I address in the next chapter, when we turn to the analysis in the framework laid out below.
Again, apart from some form of the Theta Criterion it is not immediately obvious what would rule out either derivation sketched here. On formal syntactic grounds (such as locality conditions qua MLC), both should be fine.

We can generalize the attempted movement steps for the thematic layer as in (10), where movement of arguments within vP at some point in the derivation is ruled out (again, abstracting away from a more explicit structure of the goal position)—otherwise some constructions like those discussed above would be found.\(^9\) (The arrows in (10) indicate the intended move of the elements boldfaced above, the copy of the moved element is struck through, and the purported illicit movements are marked with a hash mark rather than a star, conforming to the notation employed throughout.)

\(^9\) At this time, I consider the structure of this part of the clause as in (10), for expository purposes only. In section 2.4 I address the structure in more detail, supporting the assumptions sketched here. The same goes for illustrations regarding the structure of INFL and COMP in the following two sub-sections.
Movement from theme to agent position, as illustrated in (10a), seems to be ruled out (corresponding to (2a) above); likewise, movement in (10b) appears to be forbidden, from goal to agent position (cf. (4a)), and so does (10c), movement from theme to goal position (see (6a) and (8a)).

As the data above suggest, the movements sketched in (10) really are illicit, whether under a reflexive reading or any other. Thus, at first sight, movement into θ-positions does not seem to be an option in simple clauses. It would be rather unfortunate if arguments could only move into θ-positions in one specific type of construction (i.e.
control), even if a generalization could be concocted that would make this fall out from something else. I am going to propose a generalization in terms of anti-locality that can capture both states of affairs in a principled way—allowing for θ-movement in control, but also in other environments—by banning movement within a specified domain.

2.2.2 Anti-Locality in Agreement Dependencies

A similarly paradigmatic case can be constructed for other illicit derivations, and in fact it has been evoked to partly motivate the Case Filter. Just as all arguments apparently must receive one and only one θ-role, according to the Theta Criterion, (at least) all noun phrases need to be Case-marked, and they can only be Case-marked once, that is they can receive only one and only one morphological Case.

If Case-marking is the reflex of a Spec-Head relationship, and if all (at least, structural) Case is checked outside vP,10 we could resort to anti-locality to account for the illicit structures in (11), from German (a language with overt Case-marking on DPs):

---

10 Within Checking Theory, the traditional distinction between structural and inherent Case is not so easy to maintain anymore. Moreover, if proposals such as Zwart’s (1991, 1993, 1994, 1997a) are on the right track, even some languages traditionally viewed as SOV are underlyingly SVO (such as Dutch or German; see also Kayne 1994 for more universal claims); this analysis forces overt movement of all argument DPs and could possibly be motivated by the need to get Case-marked (nominative, accusative, dative etc.). (For counterarguments against this approach, especially for German, see some of the references given in fn. 45.)

However, if this is the driving force behind A(rgument)-movement in these cases, it should also apply to PPs, which sit outside VP at Spell Out. Whether PPs can be argued to need Case (cf. Hornstein 1995) or whether some other feature needs to be evoked is outside the scope of the present discussion. For concreteness, I assume that all argument DPs and PPs in German must move out of their thematic position overtly and the need to check for φ-feature agreement is a plausible trigger, accompanied by Case-marking. (These issues will be discussed in more detail as we go along.)
    the.ACC father like his.NOM son
    intended meaning: ‘The father likes his son.’
    b. Der Vater mag seinen Sohn.
    the.NOM father like his.ACC son
    ‘The father likes his son.’
    (German)

We could envision (11a) to follow a derivation where the object stays in its θ-
position, while the subject moves and first picks up accusative, and then moves on to the
canonical subject position (receiving nominative case in addition, not indicated in
(11a)).\(^{11}\) The object cannot receive ACCusative case, as that position is already occupied
by a copy of the subject, and could come out as NOMinative, arguably the default case in
German.\(^{12}\) The relevant part of this hypothetical derivation might then look as follows:

(12)  # [\[TP den Vater [mag-\(\nu\)-AgrO]_T [\[AgrOP den Vater t\(\nu\)-AgrO [\(\nu\)P ...]]]
    [NOM]                                               [ACC]

On a par, in (13a) the indirect object would receive accusative before moving on
to its “correct” position where it could be marked dative. The direct object would have to
stay in situ, where it could be equipped with default nominative.

---

\(^{11}\) Note that this sentence is perfectly acceptable if den Vater ‘the father’ is the topicalized
object and sein Sohn ‘his son’ the subject. For obvious reasons, this is not the reading we
are concerned with here, though.

\(^{12}\) Alternatively it could move to a higher Agr-projection, call it AgrIOP (where indirect
objects check φ- and Case features), and check dative. The particular choice of derivation
does not affect the argument. I will employ AgrIOP in the following for illustration and
discuss it in section 2.4.2 below.
(13) a. * Der Vater gibt seinen Sohn der Wagen.
    the.NOM father gives his.ACC car the.NOM car
    intended meaning: ‘The father gives his son the car.’

b. Der Vater gibt seinem Sohn den Wagen.
    the.NOM father gives his.DAT son the.ACC car
    intended meaning: ‘The father gives his son the book.’ (German)

We could sketch the relevant aspects of this derivation as follows:

(14) # [TP der Vater T [AgrIO der Sohn AgrO [AgrOP seinen Sohn AgrO [vP ...]]]]
    [NOM]     [DAT]     [ACC]

In earlier approaches to the Principle and Parameters Theory, these illicit steps are ruled out by the Case Filter. In minimalism we could argue that were a DP to move to two Case positions, another DP could not check its Case and the derivation would crash. The latter approach cannot, however, easily account for why a given DP might not enter the derivation with two Case features, while another could come with default case. There are instances in which default case is clearly licensed, such as in hanging topic constructions.\(^{13}\) We would have to stipulate that two Case features cannot be part of the internal make-up of a DP. In either case, something additional has to be said, and combined with the illustrations from the previous sub-section we now have a couple of instances expressing anti-locality: some structural configurations are too close to be licensed.

This brief discussion suggests that a generalized derivational step such as (15a) is neither found nor theoretically desirable. Suspending further discussion for now, this constellation can be extended in ways sketched below, and possibly more. Hence, what we are looking for are ill-formed steps in the derivation of the sort illustrated here:

\(^{13}\) I return to hanging topics in chapter 4, in the larger context of left dislocation. The notion of default case shall not play any further role here. For an interesting proposal of its properties and licensing conditions in a minimalist framework, see Schütze (1999).
(15) a. 

```
(15) a.   #               F1P
        ARG               F1'  
                F1 ...  F2P  
          ARG             F2'
```
(15a) is the literal transposition of (10) into the next structural layer.\textsuperscript{14} This relates apparently forbidden multiple $\theta$-movement and equally apparently forbidden multiple Case-movement, basically subsuming the Theta Criterion and the Case Filter under one illicit configuration. I challenge both the Theta Criterion and the Case Filter in their original application on the one hand, and the prohibition of multiple $\theta$- or Case movements on the other. However, before going there, I want to stress the parallelism between the data covered in this sub-section and the last one, expressed in (15a). The hypothesis at this point would be then that the thematic layer and the agreement layer both constitute a specified domain within clause structure over which anti-locality could be defined.

For the sake of completeness, I illustrate other possible movements to be ruled out in this part of the clause. (15c) is immediately ruled out under the conditions on movement mentioned before: if movement is driven by feature checking and if feature checking is the result of Spec-Head configurations only, movement to an adjoined position is ruled out (see section 3.2 for extensive discussion). \textit{Ceteris paribus}, this should thus also apply to the structure (15b).

2.2.3 \textit{Anti-Locality in Discourse Dependencies}

Lastly, consider the left periphery of the clause. Among many others, Rizzi (1997) proposes a finer structure for COMP, an articulate CP-layer (or \textquotedblleft split CP\textquotedblright). Presumably, propositional operators are licensed in this part of the clause, and so are clear discourse-relevant elements such as topic or focus, or the part of the derivation that is responsible for typing a clause, the illocutionary force.

\textsuperscript{14} Note that only (15a) has been illustrated above. In (15a), \textquoteleft\textquoteleft FnP\textquoteright\textquoteright\ stands for any $\phi$-position (TP, AgrOP, AgrIOP) and \textquoteleft\textquoteleft ARG\textquoteright\textquoteright\ for any argument. (15b) illustrates a split subject position into AgrSP and TP, which I will come back to. \textquoteleft\textquoteleft ADV\textquoteright\textquoteright\ in (15c) stands for adverb and the labels of the projections do not matter much, also to be discussed later.
I discuss the properties and structure of COMP in more detail in section 2.4.3, but let us assume some structure of the form Rizzi suggests:

\[(16) \quad [\text{CP } C [\text{TopP Top [FocP Foc [TopP Top [FP F INFL]]]]}]\]

Pending more discussion, CP (Rizzi’s ForceP) expresses the clause type, Top(ic)P hosts syntactically displaced topics, Foc(us)P a focus-moved constituent, and FP other, discourse-relevant material. (Chapters 4 and 5 deal with COMP in considerable detail.)

Regarding two sometimes related positions at the high end of the COMP-spectrum, a move from FocP to CP could be envisioned under the following set of assumptions concerning Wh-questions: Wh-phrases, being inherently focused, target FocP, if they move at all (Brody 1990 and much subsequent work).\(^{15}\) CP serves as the locus of clause-typing in Rizzi’s terms. Within a framework assuming a simple, non-split CP, Cheng (1991) held CP responsible for hosting all Wh-elements at the relevant level of interpretation, i.e. LF, possibly also hosting a Wh-operator. Putting Cheng’s claim into a split CP perspective, the Wh-phrase could be argued to move to FocP and check a focus feature, and then move on to CP to type the clause interrogative.

These assumptions are not adopted here. Rather, following much recent literature, I assume that Wh-phrases indeed target FocP but need not (hence must not) move further; neither do in-situ Wh-phrases undergo LF-movement. Clause-typing is done in CP, by (movement of) a Q-morpheme.\(^ {16}\) It is [Q] that targets to CP, and Wh-phrases may (Eng-

---

\(^ {15}\) I refer to movement in the overt syntax. Wh-in situ languages might not involve overt movement (Huang 1982) or maybe no movement at all (Aoun & Li 1993a, 1993b, Brody 1995). I address this issue in chapter 5.

lish) or may not (Chinese) move to FocP, or somewhere below (Serbo-Croatian). SpecCP is thus not a possible landing site.

Hence, only (18b) is an admissible representation for the relevant part of (17), disregarding for the time being where the Q-operator comes from:

(17) **Who** did Mary kiss?

(18) a. # [CP **who** C [FocP **who** … ]]  
[Wh] [Foc]

b. [CP Q C [FocP **who** … ]]  
[Q] [Wh]

In other words, if there is something to a (so far, intuitive) notion of anti-locality ruling out movement to a position too close, we might arrive at an explanation why structures such as (18a) should be ruled out. While other conditions, such as an economy condition that forces the least amount of movement (“Fewest Steps;” cf. Zwart 1996b, 1997b) or some version of Greed, might be able to derive the same result in this case, they could not do the work that a generalized theory of anti-locality, as I present in the next section, can: my observation that movement to a position too close is illicit holds for all relevant structures in the clause—theta domain, agreement domain and discourse domain.

Consider another structure that involves possibly multiple C-related positions. We could envision something like the following if we followed Rizzi (1997) to the letter and assume that FocP is couched in between two topic projections. Suppose that (19a) and (19b) are well-formed Italian sentences, and that *domani* ‘tomorrow’ is a topic in both cases, with *questo* ‘this’ being the focused constituent (indicated by capitals):
(19) a. Credo che a Gianni, QUESTO, domani, gli dovremmo dire. believe.1SG that to Gianni this.FOC tomorrow him should.1PL say ‘I believe that to Gianni, we should say THIS tomorrow.’

b. Credo che domani, a Gianni, QUESTO, gli dovremmo dire. (Italian; Rizzi 1997: 295-296)

If anti-locality can be shown to apply across the board, domani ‘tomorrow’ must move to TopP in one fell swoop from wherever it originates, and not pass through a lower TopP, as illustrated in (20):

(20) # … che [TopP domani [TopP a Gianni [FocP QUESTO [TopP domani gli [TP … ]]]]]

Naturally, the same kind of considerations already mentioned could rule out this movement (Fewest Steps, Greed etc.) or we could assume that no element may check the same feature twice (such as a [Top]), at least not within one clause (as to accommodate successive-cyclic movement). But to my knowledge, no single condition could capture all of the above, especially if we include the thematic dependencies discussed.

A third case of an anti-local dependency in the COMP region that comes to mind concerns so-called “Wh-topics,” Wh-phrases that also function as topics (cf. Tang 1988, Wu 1996, 1999, Grohmann 1998, 1999c, den Dikken & Giannakidou 2000; see also the discussion in chapter 5). A technical implementation of anti-locality would not only rule out a derivation that moves the Wh-phrase to both, a topic and a focus/Wh-position in theory, but also empirically. Thus, whatever the status of Wh-topics, it cannot involve movement to a topic and a Wh-position, as the following shows.

One instance of a Wh-topic is a fronted Wh-element in a Wh-in situ language such as Chinese (see chapter 5). Wu finds peculiar interpretive restrictions on the Wh-phrase, reminiscent of topics rather than (focused) Wh-elements: the fronted Wh-phrase is not quantificational, it does not allow pair-list interpretations, and so on. One possible
explanation could be that the Wh-phrase moves to both, TopP to check [Top] and FocP to check [Wh]. However, if anti-locality in the loose sense used so far applies to COMP as well, the derivational steps in (21b) for a construction like (21a) should be ruled out:

(21) a. Shenme Zhangsan mai-le? what Zhangsan buy-ASP
    ‘What did Zhangsan buy?’ (Chinese; Wu 1996: 173)

b. # [TopP shenme Top [FocP shenme Foc [TP Zhangsan mai-le shenme]]
    [Top]                     [Wh]

Regardless of the ordering of the topic and the Wh-position, we do not want this double movement above TP. As I will discuss in detail later, this objection receives independent support. As a Wh-in situ language, Chinese does not overtly front Wh-elements to license an interrogative (Huang 1982). Thus, it should never front a Wh-phrase to check [Wh], at least not in the overt syntax (any version of Economy will do). As such, we can accept Wu’s analysis and assume that shenme ‘what’ just moves to TopP.

A similar case can be constructed for English. As the paradigm in (22) suggests, a Wh-phrase in English that possibly occupies a structural topic position—regardless of whether it moved to TopP first or last—is ruled out (cf. Lasnik & Saito 1984, 1992, Rizzi 1996). Given that the English Focus head is verbal (cf. (22b), where never presumably sits in SpecFocP) and that the Topic head is not (the topic Mickey Mouse has moved to SpecTopP in (22a)), (22c) cannot involve the Wh-phrase in FocP (compare with (22d), where the auxiliary has raised to Foc). On the other hand, Wh-elements in English move to FocP overtly, so we could not say that who(m) in (22c) sits in SpecTopP at Spell Out.
(22)  

a. **Mickey Mouse**, the Americans will elect for president.  
b. **Never** will the Americans elect Mickey Mouse for president.  
c. * **Who(m)**, the Americans will elect for president?  
d. **Who** will the Americans elect for president?  

The relevant derivational step of (22c) cannot look like either possibility in (23), setting aside possible movement of *will* (which should not rescue either structure):

(23)  

a. # … [TopP *whom* Top [FocP *whom* Foc [TP the Americans will elect…]]]  
b. # … [FocP *whom* Foc [TopP *whom* Top [TP the Americans will elect…]]]

Again, one might want to resort to other conditions. Criteria come to mind: the one element satisfying the Topic Criterion cannot at the same time satisfy the Wh-Criterion (Rizzi 1996), for example. Or we could blame it on the semantics and say that a topic cannot be a focus at the same time. However, there are good reasons to avoid representational conditions such as criteria, for the same reason we do not want to keep filters of sorts, if there is anything to substantive economy: they are grammatical formatives that are not virtually conceptually necessary. Likewise, as we will see in chapter 5, there are good reasons to believe that Wh-elements can behave topic-like and still be quantificational. In short, there is currently no uniform explanation for all the illicit movements noted in this section; an anti-locality condition defined over Prolific Domains offers one.

On analogy with the illicit moves depicted in (10) and (15) above, (24) contains a sample of movements that would violate anti-locality in the C-layer, given the structure sketched in (16) above which I address in section 2.4.3 and subsequent chapters. (24a) illustrates Wh-movement from FocP to CP, initially considered for (17), (24b) sketches topicalization of one element involving two separate topic positions within one clause, relevant for (19), and (24c) demonstrates the illicit movement of a Wh-phrase from TopP to FocP (cf. (21a) or (22c)).
(24) a. #
   \[\text{CP}\]
   \[\text{WH}\]
   \[\text{C'}\]
   \[\text{C} \ldots\]
   \[\text{FocP}\]
   \[\text{WH} \ldots\]

b. #
   \[\text{TopP}\]
   \[\text{TOPIC}\]
   \[\text{Top'}\]
   \[\text{Top} \ldots\]
   \[\text{TopP}\]
   \[\text{TOPIC} \ldots\]

c. #
   \[\text{FocP}\]
   \[\text{WH}\]
   \[\text{Foc'}\]
   \[\text{Foc} \ldots\]
   \[\text{TopP}\]
   \[\text{WH} \ldots\]
2.2.4 Summary

To summarize this section, I introduced some basic data that, I argue, demonstrate anti-locality: some movement steps are illicit because the potential dependency formed would contain two positions that are too close to each other, or too local (anti-local). In the thematic layer, the proto-typical case involves a movement dependency that would include two θ-roles. The agreement layer cannot involve dependencies that contain two Case-markings (or two sets of φ-features). Likewise, illicit dependencies in the discourse layer involve any interaction between topics, focused elements or Wh-phrases and checking of more than one of these features. The choice of the three-way split above is no accident: I now take these layers and consider them in a bigger picture of anti-locality.

2.3 Towards a Theory of Anti-Locality

It is time to lay out some specifics with regard to the concept of anti-locality. This is the road map to anti-locality: I propose to split the clause into three Prolific Domains, each ranging over a specific context. I first discuss what a Prolific Domain is and what it does (section 2.3.1); chapter 6 deals with more discussion beyond the obvious. I then define anti-locality in terms of the Condition on Domain Exclusivity (section 2.3.2); I first demonstrate Exclusivity empirically in section 3.4, continue in chapter 4, further assume its presence in chapter 5, and finally address some predictions and consequences in chapter 6. Finally, I turn to some aspects of a finer architecture for each Prolific Domain, that is the internal structure of each layer, or domain, relevant for anti-locality evaluation, hence prolific (section 2.4).
2.3.1 Prolific Domains

In specifying what a Prolific Domain is, I rely on the earliest conceptions of clause structure in generative grammar. For the longest time (see Chomsky 1986a in terms of X'-Theory), a sentence has been taken to have a structure like (25):

\[(25) \quad [ \text{COMP} [ \text{INFL} [ \text{VP} ]] ]\]

One way of motivating the three parts of the clause noted in (25) is functionally. The verb phrase (or other appropriate equivalents) contains the predicate of the clause and the minimally predicated element(s). That is to say, in a typical transitive clause we would find the predicate (e.g. verb) and its arguments (e.g. agent, theme) merged here. INFL(ection), S, IP, TP or any other notational variant, marks the clause for at least agreement and tense. In this part of the clause, we can typically observe subject-verb agreement (found even in English), alongside nominative case-marking on the subject, accusative for the object, other types of verbal agreement, and so on; this part also seems to license most additional information regarding the more articulate properties of tense and aspect, including licensing of the relevant adverbial material. COMP(lementizer) or CP is the locus of A'-elements (at least fronted Wh-phrases, certain types of topic and focus, and so on); it is also the domain that expresses the clause type and that contains propositional operators (both of which I will return to in detail).

---

17 By convention, I capitalize Prolific Domain to designate its technical meaning, and set ‘Domain’ apart from other uses of ‘domain’ (as relevant for binding, as a descriptive cover term etc.).

18 By “functionally” I mean the three-way split expounded here—theta, agreement and discourse properties. While it may be an interesting question why exactly these (and why, if I am correct, only these) three layers should bear such relevance for computation on the clausal level, but I do not pretend to have anything insightful to say. I would rather pick up the pieces from the apparent fact about language that information is structured this way and the functions of the (three) domains of the clause are the way they are.
We can summarize the intuition behind Prolific Domains as follows, where the Greek letter corresponding to the properties and \( \Delta \) for Domain are conventions:\(^{19}\)

(26) **The concept of Prolific Domains (\( \Pi\Delta \))**

i. **\( \theta \)-domain**: the part of the derivation where theta relations are created

ii. **\( \phi \)-domain**: the part of the derivation where agreement properties are licensed

iii. **\( \omega \)-domain**: the part of the derivation where discourse information is established

The layer responsible for selection requirements ("theta relations") encompasses the lexical verbal head \( V \), the functional verbal head \( v \), and their respective complement and specifier positions plus adjoined material, or whatever of these is present in a given construction. Given that up to the point of completion of \( vP \) (by merging the external argument into Spec\( vP \), glossing over possible adverbial modification or other adjuncts)—but not beyond—each XP is limited to one occurrence, this layer pertains to the theta domain, or \( \theta \)-domain. (As a shorthand, I indicate this part as \( \theta\Delta \), where relevant.)

Beyond the \( \theta \)-domain, a layer of functional projections emerges that licenses verbal morphology and agreement, from aspectual, negative and certain modal properties up to tense (see especially Cinque 1999 for elaborate discussion); \( TP \), thus, is the projection that ultimately dominates the articulated INFL—whether we take \( T \) to be "real" or a cover term (possibly in the sense of Chomsky 1998, 1999).\(^ {20}\) The projections employed

---

\(^{19}\) See also Rizzi (1997: 281). In a recent paper, Platzack (in press) comes to a similar conclusion about the composition of the clause, independently developed. Space does not allow a detailed comparison between Platzack’s system and the present one, but some hints will be given in chapter 6 (see Grohmann 2000a).

\(^{20}\) I shall remain agnostic as to whether an agreement projection for the subject, such as \( \text{AgrSP} \), should be part of the inventory (see Pollock 1989, Belletti 1990, Chomsky 1991, Zwart 1993), or whether its purpose could be cashed out otherwise. The presence of \( \text{AgrSP} \) seems to be supported by recent work on pronominal properties of Italian dialects, in particular, further suggesting a more intricate split into separate projections for person, number and gender, each licensing a different kind of clitic (see Manzini & Savoia 2000, Poletto 2000 or Tortora 1999, to appear, for example).

*(continued...)*
in (15) above seem to serve the main purpose to license verbal morphology (overtly or covertly).\textsuperscript{21} This licensing is arguably achieved by checking, through successive-cyclic head movement. Moreover, the subject is licensed in SpecTP in a Spec-Head configuration; dispensing with separate Agr-projections, all properties pertaining to Case, agreement and φ-features would thus ideally be checked somewhere in this domain. Let us say that (argument) XPs must move into this Prolific Domain for purposes of agreement or φ-licensing. Call this part the φ-domain (or φΔ for short).

Running through further projections, and following the standard notion of successive head-movement, we reach the C-layer next. Initially conceived of as COMP or S’, empirical evidence suggests a more intricate internal structure, such as the finer articulation of CP proposed by Rizzi (1997) briefly introduced in the previous sub-section. I assume such an articulate structure, making available positions that express discourse properties (such as point-of-view, topic, focus or illocutionary force). On analogy with the other two Prolific Domains, this part represents the ω-domain (ωΔ).\textsuperscript{22}

As will perspire from the remainder, AgrSP could easily be integrated into the present framework—as long as movement to SpecAgrSP does not pass through SpecTP. I thus employ TP as the designated subject position, which might be just a collective cover term. This, and other issues, cannot be addressed in this work. Once the framework is on the table, and supported with the coverage offered in the following chapters, a number of predictions and consequences for a host of relevant phenomena will be apparent, but it has to be left for future research to work out the details of those phenomena that have been argued to involve movement within a Prolific Domain (participle agreement discussed by Kayne 1989, Belletti 1990 or quantifier scope à la Beghelli & Stowell 1997, for example). Chapter 5 contains a hint at a possible solution in terms of “syncretic” categories (Giorgi & Pianesi 1997, Zubizarreta 1998) in the context of types of scrambling.


\textsuperscript{22} The choice of the terms θ- and φ-domain is obvious; ω as a label for the C-layer is invented, not so much as to confuse but to be uniform. Moreover, as the C-layer is the highest part of the clause, capping it off, the last letter of the Greek alphabet might be an appropriate choice. There is a metaphorical mnemonic for ω which might be useful, too, derived from the Greek word ὄρμιοτητα ‘ripeness, maturity, full growth’. (As such, one may pronounce “ω-domain” as ‘omega-domain’ or ‘orimo-domain’.)
There is an intuitive appeal behind the idea that clauses divide functionally along the lines noted. We can find further support for the idea that clauses are arranged in a three-tiered set of Prolific Domains within verbal morphology. We can split the occurrences of verbal elements into the same three parts: the lowest part, without any inflection (purely creating thematic relations with arguments), the middle part, where inflection is marked on the verb (licensing agreement relations), and the highest part, where morphological marking is basically absent (establishing at most discourse information).

Natural language possesses a large number of verbal inflection for elements in the $\phi$-domain, such as tense, aspect, voice or agreement (see especially Cinque 1999). Interestingly, though, there are no unique morphological markers defined over the $\theta$-domain (such as agent-, goal- or theme-inflection) or the $\omega$-domain (such as Wh-, topic- or focus-inflection). Some morphemes may show up on the verb, of course, but for different reasons, namely as the result of verb raising to a C-head which presumably is driven by forces other than the need to pick up agreement or license arguments. Some languages encode reflexive markers on the verb as well, but there are good reasons not to think of these morphemes in the same way as we think of the tense-aspect-agreement complex.

Let us take (27) to be an adequate definition of a Prolific Domain:\textsuperscript{23,24,25}

\begin{footnotesize}
\begin{enumerate}
\item Except, perhaps, mood (e.g. subjunctive), which arguably is construed with the C-layer rather than the T-layer (cf. Rivero 1994). This concept is directly linked to other properties, however, which we can abstract away from for now. It does not really matter whether this state of affairs indeed holds. The reasoning is not meant to constitute an argument for anything, but simply another indication of the present three-way split.
\item I will flesh out the concept of Prolific Domains in chapter 6, where I discuss how it fits in with recent ideas on interfaces, such as ‘phases’ (Chomsky 1998, 1999), ‘multiple interfaces’ (Platzack, in press), and potentially ‘opportunistic’ interaction between the PF and LF interfaces and $C_{HL}$ (Boeckx 1999a). To put it crudely, I suggest that the two interface components (or levels of representation), PF and LF, are mapped cyclically: each time a Prolific Domain is established, it gets spelled out, that is, shipped to the interfaces.
\item Anticipating the following discussion, I still hold on to a notion of successive-cyclic head movement. The attentive reader might be puzzled as to how this could fit in with the instances of anti-locality demonstrated so far, basically ruling out too local movement steps —under the reasonable assumption that head movement is very local. I will show (continued...)
\end{enumerate}
\end{footnotesize}
(27)  *Prolific Domain* (working hypothesis)

A Prolific Domain $ΠΔ$ is a contextually defined part of the computational system

i. which provides the interfaces with the information relevant to the context and

ii. which consists of internal structure, interacting with derivational operations.

I take the mnemonics $\theta$, $\phi$ and $\omega$ to be “context values,” part of the information contained within each head (cf. “extended projections” à la Grimshaw 1990 or Ernst, to appear). Thus, $V$ and $v$, for example, come equipped with the context value $|θ|$, $T$ or $\text{Agr}$ with the context $|ϕ|$, and $\text{Top}$ or $C$ with the context $|ω|$. This convention identifies a number of heads, and by extension projections, as part of the same context. This information is relevant in two aspects: first, it groups various projections into a single Prolific Domain; second, each Prolific Domain contextually identified in this way ships the information relevant for the specific context to the interfaces. I will expound this idea in chapter 6. For now, let us take (27) at face value when we talk about Prolific Domains.

### 2.3.2 Exclusivity

With the notion of Prolific Domains in place, we can account for the data discussed in section 2.2 in terms of an anti-locality condition on movement. The instances of illicit movement, which I ascribed to anti-locality throughout the discussion, all take place within a single Prolific Domain.

As a first pass then we could understand anti-locality as a ban on movement within a single Prolific Domain. This seems to be too restrictive, though. It would imply that head movement does not exist. The standard implementation of head movement is a successive-cyclic operation, subject to very strict locality conditions, expressed by the
Head Movement Constraint (Travis 1984). As such, we assume that every head moves to
the next available head and cannot skip it (see, for example, Chomsky 1995a, Lasnik
1999, and the references cited there).

If we said that anti-locality bans all Domain-internal movement, we would expect
a relevant structure of the type of (29a) for a simple sentence like (28) from German,
which has overt verb movement (boldface). The standard analysis, however, is the com-
plex head depicted in (29b), the result of successive-cyclic movement via all intermediate
head positions:

(28) Der Anna hat der Peter einen Kuß gegeben.
the.DAT Anna has the.NOM Peter a.ACC kiss given
‘Anna, Peter gave a kiss.’ (German)

(29) a. #
   [TopP der Anna hat-Top [TP der Peter hat-T [AgrIOP der Anna AgrIO
   [AgrOP einen Kuß AgrO [vP der Peter v [VP hat-V [VP einen Kuß
geggeben-V der Anna]]]]]]]

b.  
   Top
      /
     /    
     /
   T   Top
      /
     /  
   AgrIO  T
      /
     /  
   AgrO  AgrIO
      /
     /  
   v  AgrIO
      /
     /  
  V v
    /
hat
In other words, in order to rule in (29b) we must allow verb movement within a Prolific Domain. If so, anti-locality could be a ban on *phrasal movement* within a single Prolific Domain (Grohmann 1999a, 2000a, 2000c). Call this “XP-uniqueness,” a condition that a given XP must be unique within a single Prolific Domain, or the occurrence of copies of one object in a phrase marker must be exclusive in each Prolific Domain.

The main problem with this definition is its apparent *ad hoc* character. Why are only XPs subject to such an anti-locality condition? I suppose this could be motivated, especially in the light of recent doubts on the nature of head movement, if not its place in the grammatical component altogether (see Chomsky 1998, 1999, Boeckx & Stjepanović, in press). XP movimiento could thus be considered a syntactic operation, while $X^0$-movement would be something else, perhaps a (strictly) phonologically driven displacement operation taking place in the PF component.

In an ideal world, however, we have a theory with as few stipulated conditions as possible. Intuitively, though, a theory which derives most of its conditions—either from other, independently needed or motivated components of the theory or by virtual conceptual necessity—should be chosen over an alternative that might yield the same empirical coverage, but needs to stipulate additional components. This is a basic application of Ockam’s Razor to choosing a particular theoretical framework; see also Martin & Uriagereka (2000) for some relevant discussion on a theory of grammar.

I take this as my guiding principle to leave behind the idea of “XP-uniqueness” in a formal way (i.e. built into the definition) and replace it with a principle very similar in spirit, but easier to motivate or derive. I propose the Condition on Domain Exclusivity (CDE or Exclusivity for short) to capture anti-locality defined over Prolific Domains.\(^{26}\)

\(^{26}\) (30) is different from an earlier formulation of the CDE purely as XP-uniqueness (Grohmann 2000a: 63).
(30) **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.

The object $O$ refers to any structure within a Prolific Domain, such as DP, PP, VP, $v'$ and so on (and their sub-structures) within the $θ$-domain, for example. The notion of $AI$ will be discussed in section 6.1; let us take it to mean here that a licensed occurrence of $O$ signals its presence to the interfaces and this presence must be unique (or exclusive). The “drastic effect on the output” is understood in terms of a different PF-matrix of $O$. A multiple occurrence of $O$ is more than one *non-distinct* copy of $O$. In other words, if a multiple occurrence of $O$ involves two *phonetically distinct* copies of $O$, the CDE is not violated. This translates into the following: if moving $O$ within a $ΠΔ$ is followed not by deletion of the copy of $O$, as in standard cases of movement (which would violate the CDE), but by *spelling out* of the copy, with a different PF-matrix, the CDE can be upheld (by (30ii)).

Empirical coverage of this part of the CDE will be the topic of the second part of chapter 3, the main focus of chapter 4, a background assumption in chapter 5, and addressed further in chapter 6. I consider the CDE in more detail next.

---

27 In chapter 6 I will suggest that Prolific Domains can be conceptualized as shipping their internal information off to the interfaces, the PF and LF components. LF could not care less how many copies there are, as long as each copy fulfills a different function (i.e. checks a different feature). PF, however, cares about too much information, particularly worries about repetition. This is the standard argument for the obligatory deletion of copies (Chomsky 1995a), reformulated by Nunes (1995) in terms of linearizability of strings. An LCA-driven need for deletion fits in nicely: if the PF-matrix of $O$ looks different, I consider later, both copies should not only be interpretable at PF, but also linearizable. A distinct PF realization of two copies may thus serve as an escape hatch (see chapters 3 and chapter 4 for empirical demonstration).

I take an occurrence of $O$ to be defined over $O$’s mother (“Immediate Contain,” Chomsky 1999) rather than its sister (“Sister,” Chomsky 1998). I will address the issue in considerable detail in the first part of chapter 3 and discuss consequences in the second part of chapter 3, returning to those in chapter 6.
Let us start with terminology. Recall from section 1.3 that in Copy Theory, the complex operation Move takes an object O, Copies it and Merges it at some later point. This operation is one of substitution. The copy, usually the lower one, is then deleted, presumably for phonological reasons: as Nunes (1995, 1999) shows, failure to delete one copy leads to a failure to linearize the two copies; the lower copy is deleted because by definition it has more unchecked features than the higher copy. However, it is not erased—in many cases we want it to be there at the interpretive component (LF). And as many cases of reconstruction indicate, the displaced element often has access to the original position, and in principle, either copy could be interpreted. This said, we can understand “occurrence” as a copy, whether active or deleted, within a relational domain defined in terms of motherhood (see fn. 27). Non-distinctness is understood in terms of Select: if a lexical item LI is selected for the LA once and copied in the course of the derivation, the two (or more) copies are non-distinct; two selections of one LI are distinct copies. The CDE in (30), then, rules out two copies of the same element within one Prolific Domain.

In light of this, consider head movement once again. The standard view of head movement is adjunction. A head X moves to another head Y by adjoining to Y, yielding \([Y X-Y]\). Rather than (31b), (31c) or any other alternative, a complex head formed by one movement of X to Y looks like (31a):\(^{28}\)

\[
\begin{align*}
(31) \quad a. \quad [_{YP} \ ZP \ [X^0-Y^0] \ Y^{0\max} \ [_{XP} \ ZP \ X^\theta]] \\
b. \quad \# \quad [_{YP} \ ZP \ X^0 \ Y^0 \ [_{XP} \ ZP \ X^\phi]] \\
c. \quad \# \quad [_{YP} \ ZP \ X^0 \ [_{XP} \ ZP \ X^\phi]]
\end{align*}
\]

\(^{28}\) Following conventions of Bare Phrase Structure, a complex head is \(X^{0\max}\), and \(X^0\) denotes the intermediate projection of the head which another element adjoins to. We can mostly refrain from specifying heads.
Following Chomsky (1994, 1995a), head movement adjoins by forming a two-
segment category (adjunction, (31a)), rather than a one-segment category (never allowed, (31b)) or a new category (substitution, (31c)). Let me graphically magnify (31a).

(32) \[
\begin{align*}
\text{YP} \\
\text{ZP} & \rightarrow \text{Y'} \\
\text{Y}_{\text{\text{0max}}} & \rightarrow \text{XP} \\
\text{X}^{\text{0}} & \rightarrow \text{Y}^{\text{0}} \rightarrow \text{ZP} \\
\text{X'} & \rightarrow \text{X}^{\text{a}} \rightarrow \text{...} \\
\end{align*}
\]

To illustrate the abstract representations with natural language, consider German again, where the lexical verb arguably moves quite high. The correct derivation of the relevant parts of a simple sentence like Peter schläft ‘Peter sleeps’ is (33), where I use the simplified notational convention of t(race) to indicate former head positions:

(33) \[
[[[\text{schläft-}\text{v}_{\text{0}}]-\text{v}_{\text{0max}}]-\text{t}_{\text{0}}]-\text{t}_{\text{0max}} \quad [\text{v}_{\text{p}} \text{Peter} \text{t}_{\text{0}} [\text{v}_{\text{p}} \text{t}_{\text{0}}]]
\]

Returning to the definition of Exclusivity (the CDE in (30)), the two copies of Peter in (33) are non-distinct: there is only one LI Peter selected for the LA. Moreover, both copies look roughly like \([\text{DP Peter}]\) and have the same morphological make-up and share the same PF-matrix; yet they differ with respect to their feature set-up: the lower copy has unchecked \(\phi\)-features, possibly Case also, which the upper copy has checked. This is the motivation to delete the lower copy, where deletion is taken to be necessary for linearization (see Nunes 1995, 1999). As they occur in two different Prolific Domains (the \(\theta\)- and the \(\phi\)-domain, respectively), this movement is well-formed. The lower copy gets de-
leted and the two occurrences of Peter in the phrase marker (not in one Prolific Domain) can be linearized. The three heads involved here (V, v, T), are distinct, however. The original V is [schläft-V₀], v consists of [[[schläft-V₀]-v₀]-v₀max] and T presumably has the complex structure [[[schläft-V₀]-v₀max]-T₀]-T₀max. Head movement arguably involves concatenation of morphemes, and we can safely assume that even functional heads such as v bear a (possibly, zero) morpheme (often expressed in causative morphology, for example). As a result, the newly created complex head V-v-T is distinct from its original; whether the different morphological forms imply a different PF-matrix, or whether it is simply enough to distinguish two copies shall not be our concern. ²⁹ We can thus see that the simple head V and the complex head V-v occur in the same Prolific Domain, but are distinct; hence this movement is permitted, and also falls under (ii) of the CDE. ³⁰

In sum, anti-locality boils down to a ban on movement of a maximal projection from one position to another within the same Prolific Domain, as it results in two non-distinct copies. If either copy is deleted (arguably a “drastic effect on the output” in (30)), the two cannot be linearized (see also next chapter). Spelling out the lower copy is a repair strategy, corresponding to (30ii), which I will discuss later. Movement of a head does not play a role, as this operation creates a distinct copy, also satisfying (30ii).

²⁹ Alternatively, head movement could be viewed as changing the PF-matrix: assume that V is actually schlafen ‘sleep’ (the infinitive) and it checks the relevant inflection in higher heads (such as third person singular –s in T). Even if no inflection is “picked up,” as in v, we could assume that the PF realizations of V and V-v are different. This is reminiscent of an Affix Hopping approach to head movement, originally introduced by Chomsky (1957) and recently revamped under minimalist assumptions by Lasnik (1995a). This might not be the preferred implementation of distinctness of heads, possibly being incompatible with a lexicalist approach (see Potsdam 1997, Roberts 1998 for criticism and Lasnik, in press for a reply).

Regardless of whether we assume that head movement is always an instance of incorporation, which quite clearly has a morphological effect, or something else entirely, the assumption that head movement changes the internal make-up relevant for distinctness does not seem to be too controversial. I am grateful to Anna Roussou for discussing this point with me.

³⁰ Movement of v to T has no bearing on the issue for obvious reasons (as it involves two different Prolific Domains, the θ- and the φ-domain).
2.4 On the Internal Architecture of Prolific Domains

The concepts introduced above will be supported empirically, explored conceptually and refined technically throughout the remainder of this work. First, though, let us look at the fine structure of each Prolific Domain, by way of introducing the present assumptions on clause structure and derivation in a minimalist framework.

2.4.1 The θ-Domain

Many efforts have been made into figuring out the exact structure of the θ-domain. For example, Larson (1988, 1990) proposes a reiteration of VP in terms of VP-shells, Hale & Keyser (1993) implement a functional verb phrase on top of the lexical one, where the functional verb is often manifested as a light verb, and so on (Baker 1988, Marantz 1993, den Dikken 1995a, Koizumi 1995, Pesetsky 1995, Anagnostopoulou 1999). I adopt a version of the light verb approach and take the structure of the thematic layer to consist of VP (hosting the lexical verb), usually dominated by vP (light verb).

I assume, without further discussion, the Predicate-Internal Subject Hypothesis (PISH): the subject is base-generated inside the predicate. In the case of a transitive verbal predicate, the thematic subject (‘agent’) position is Spec\(v\)P.

Neither assumption seems unreasonable. An issue of much debate, however, are the positions of direct and indirect object (or ‘theme’ and ‘goal’), both with respect to each other and regarding the question whether the latter in particular is a simple constituent of VP (Spec or Comp) or part of a more articulate sub-structure (such as a small

---

31 See e.g. Fukui (1986), Kuroda (1988), Koopman & Sportiche (1991), for arguments. Recent presentations of the theoretical motivations and empirical evidence in favour of the PISH can be found in McCloskey (1997), for example, further reviewed by Hornstein, Nunes & Grohmann (in progress).
clause or other predication structure). For reasons that will become apparent shortly, I take a stance on the former question and assume that in ditransitive constructions the direct object is merged into SpecVP and the indirect object is merged into the complement position of the verb (henceforth referred to as CompVP\(^{32}\)).

For the main part, I remain agnostic as to whether CompVP is more articulated (see, in particular, den Dikken 1995a); it may very well be, but as far as I can see, nothing much of the discussion in this chapter, or any other, hinges upon this assumption. I also refrain from discussing whether the direct object is always in SpecVP or whether its original merging site is CompVP in simple transitives.\(^{33}\) This might not be implausible. Baker (1997) advocates the former, and Baker & Stewart (in progress) offer more empirical support; Pietroski (2000, in progress) supports this view from a semantic perspective. If this turns out to be fruitful, an additional projection might be called for, predicated of V. I do not believe that the final outcome has any effect on what follows, however.

I already indicated TH over GO in the structures employed in section 2.2.1 above. For the time being let us then assume that (34a) is the correct, if simplified, underlying structure for ditransitive structures and (34b) for transitive ones (if only for convenience):

\[
\begin{align*}
(34) & \quad a. \quad [_{vP} \text{AGent} V - v [_{V P} \text{THeme Verb GOal}]] \\
& \quad b. \quad [_{vP} \text{AGent} V - v [_{V P} \text{Verb THeme}]]
\end{align*}
\]

I do, however, assume that Chomsky (1991, 1993) was right in further splitting Belletti’s (1990) expanded version of Pollock’s (1989) original “split INFL” structure.

\(^{32}\) On analogy with the convention CompXP, I refer to adjoined positions as AdjXP.

\(^{33}\) The same goes for the subject of unaccusatives, for example, which has been argued to originate in the “object position” (whatever that turns out to be; cf. Belletti & Rizzi 1988). As far as I can see, the answer to this question is also mostly irrelevant for the remainder of the dissertation—as long as this view does not force us to assume additional movement of the subject, such as from CompVP to Spec\(v\)P.
Basically, theta relations and $\phi$-features must be checked in a licensed checking configuration, and it is plausible that each functional head only bears one relevant feature for checking purposes (as Pollock originally proposed). In other words, theta and agreement relations must be dissociated (see also Haeberli 1995, Manzini & Roussou 1999, 2000). I also suggest that A-movement is not Case-driven, but agreement-driven, where Case-marking is a reflex of the $\phi$-checking relation (see also Ormazabal 2000).

I do not assume that multiple specifiers are, or even can be, part of the structures made available by the grammar. Thus, accusative cannot be marked on an outer specifier of vP, as assumed in recent years, but must be related to a separate head. Given that any nominal direct object can raise from its $\theta$- to its $\phi$-position, I take it that the latter is outside the $\theta$-domain.\textsuperscript{34} I will spend considerable efforts on laying out my empirical motivations and theoretical justifications in the first part of chapter 3.

I thus take the structure of the $\theta$-domain roughly to look like (35)—also assumed in previous structural representations above—, where I indicate the argument positions with AAgent/SUbject, THeme/Direct Object, and GOal/Indirect Object, abbreviations I will use throughout. Without any further ado, I will take this to be the correct, or for us relevant, configuration for any structural analysis in the remainder of this dissertation.\textsuperscript{35}

\textsuperscript{34} Naturally, the dissociation of $\theta$- and $\phi$-features also follows from the previous two sections on Prolific Domains and Exclusivity. To raise this argument would be tautological and unconvincing, however: after all, my goal is to derive Prolific Domains and their properties, rather than stipulate them and base immediate consequences for clause structure and derivation on the correctness of my proposal.

Note further, though, that recent work on $\theta$-roles in terms ofaspectual properties might enable us to reduce them to formal features as desired, thus setting them clearly apart from $\phi$-features (cf. Borer 1994, Tenny 1994, Salles 1997, Arad 1998).

\textsuperscript{35} Abstracting away from transitives, where I follow standard convention of merging the theme into CompVP position, as mentioned above. In addition, the “subject” of some types of verbs is generated inside VP, rather than vP, such as unaccusatives or psych-verbs (cf. Belletti & Rizzi 1988). I return to some considerations of the internal structure of the goal position at the appropriate time.
There is something common to the elements that appear in this part of the clause: they all involve licensing of thematic properties, and all theta relations are only licensed in this part of the clause. Extending Theta Theory to a quasi-minimalist framework, the standard line of reasoning (Hale & Keyser 1993) is that the verb assigns the thematic roles corresponding to TH and GO to its internal arguments through a licensing configuration such as (35) and AG to its external argument through movement to \( \nu \).

As mentioned in passing, in a framework that aims to dispense with superfluous levels of representation (by force of “virtual conceptual necessity”), this account has the touch, smell and feel of D-structure, as argued by Hornstein (2000), following work by Bošković (1994), Lasnik (1995b), Bošković & Takahashi (1998), Hornstein (1998) and others. Pursuing the premise that all operations are subject to Greed, we would want every operation to be followed by feature checking or another appropriate licensing operation, unless there are good reasons to go another route. We take all instances of Merge to be costly in the sense that they need to do something. It is thus conceivable that initial

---

\(^{36}\) The current work is not a study of argument structure. As such I am not particularly concerned with how exactly arguments are licensed and what theta properties really express. Maybe the \( \theta \)-roles are “real” in the sense of Baker (1988), for example. Maybe they are “proto-roles” (Dowty 1991) or even something else entirely (e.g. Grimshaw 1979; see also fn. 34). I use ‘agent’, ‘theme’ and ‘goal’ purely descriptively.
merger of all arguments checks a “θ-feature” [θ] (and its specified values [AG], [TH] and
[GO]). (See also the references in fn. 34 for plausible justification of this move.)

As mentioned in the initial discussion of (1), this allows for a derivational account
of control. Whether this is correct shall not be our concern;\textsuperscript{37} I adopt, however, that mov-
ing into θ-positions is not ruled out \textit{a priori}, as, say, under Theta Theory. I take θ-to-θ-
movement (or even movement from a non-θ-position) desirable in some circumstances
and as such possible—provided that other conditions are met. I will return to such condi-
tions throughout; some circumstances that allow for θ-movement will be discussed in
section 3.4, in the context of the CDE and a derivational approach to local reflexivization.

In sum, we have established the thematic layer as the first Prolific Domain of the
clause. The CDE rules out vP-internal movement in the majority of cases; I propose a
way around this, following the second part of the definition of the CDE, in chapter 3. On
the assumption that multiple specifiers are not part of \(C_{HL}\), which I also discuss in chapter
3, there are no movements inside vP other than for θ-checking that one could imagine.

\textbf{2.4.2 The φ-Domain}

The next clausal domain is characterized by licensing agreement or φ-properties
—instantiated by formal features pertaining to at least Case and φ-features (arguments) as
well as inflection for agreement, tense and aspect (verbal morphology), and it hosts the
majority of adverbs (which I am not going to discuss in detail).

\textsuperscript{37} In fact, the θ-role-qua-formal-feature approach has been criticized on a number of
grounds which I cannot review here. The fairest rejection in my view, however, is that of
Uriagereka (1999a) who basically says that if all signs point towards D-structure, then we
have to swallow the bullet and assume D-structure. In the interest of a smooth progres-
sion, I will not discuss the need of D-structure and related questions, but simply follow
the line of research that explores θ-roles as features, and see what it buys us (see also fn.
1). Needless to say, I argue that it buys us a lot. This said, though, if it turns out that this
is the wrong move, the thesis of my proposal will not collapse, unless it also turns out
that everything else I assume is moot.
The desired dissociation between $\theta$- and $\phi$-checking mentioned above allows for one—in this framework, crucial—assumption: specifiers are unique. Just as arguments are merged into the thematic layer to saturate thematic information, they need to move into a separate position to check agreement. The approach to grammar in terms of Prolific Domains, which I will conceptualize in chapter 6, further suggests that this is the road to take. Following Chomsky (1993), I assume that Case is marked in a Spec-Head relation with (finite) T for nominative and (V in) Agr for accusative case. As such, all arguments must raise into this part of the clause at some point (overtly or covertly).\(^{38,39}\)

One stipulated, yet not implausible, assumption regards the manner of feature checking. This not only concerns, what we may call, “directionality” or whether features can be checked or valued without displacement, but especially how many features a particular head may bear (or at least check) in one particular configuration.\(^{40}\) The thrust of Pollock’s proposal was that a separate head for tense and a separate head for agreement features (as well as aspect, negation etc.) could account particularly well for cross-linguistic variation regarding adverb placement and verb movement (cf. Emonds 1976).

Cinque (1999) extends this reasoning and proposes as many such functional heads as there are morphological properties that need formal licensing.\(^{41}\)

---

\(^{38}\) This is a stipulated consequence of the following discussion, but an immediate fall-out from subsequent chapters (and the relevant discussion above).

\(^{39}\) Contra Chomsky (1995a), who suggests an outer SpecvP as the locus of accusative case checking (the by now “standard” minimalist treatment) or Koizumi (1993, 1995), who posits one AgrP contained within each of the different VP-shells where Case is checked (see also Lasnik 1995b, 1995c, 1999).

\(^{40}\) On directionality issues, see Chomsky (1995a) for discussion, supporting Attract over Move (but see Hornstein 2000 for the opposite view); valuing features long-distance is a property of the operation Agree (Chomsky 1998, 1999). (Bošković 1999, among others, explicitly argues for multiple feature checking.)

\(^{41}\) For classic work on “split INFL,” see Pollock (1989), Belletti (1990), Laka (1990), Chomsky (1991). Many more projections have been proposed over the years—beyond TP, AgrP, NegP (see Cinque 1999 for succinct overview). Some of these proposals can be easily integrated into the present framework, others cannot.
projections and positions for negation, he distinguishes a number of different types of tense, aspect, modality, mood etc. Interestingly, most of these projections host material relevant for verbal morphology; apart from AgrP, Cinque does not specifically assume that arguments move into the numerous available specifier slots, although he leaves this option open for quantified expressions, along the lines of Beghelli & Stowell (1997). The majority of specifiers are needed to license adverbs, which, following Kayne’s (1994) work on antisymmetry in syntax, are supposedly structurally indistinguishable from specifiers and can only be licensed as a unique left branch per projection. The proposal I make can plausibly do without such a large array of functional projections, but I follow the spirit of restricting feature checking.

I will return to adverbs (and the specifier-adjunction debate) in section 3.3. At this point it suffices to say that I treat specifiers distinctly from adjuncts in at least that the former are unique and the latter cannot be derived by movement. The myriad of possible functional projections in the φ-domain is not relevant for our purposes. What is relevant is that all arguments must at some point in the derivation raise to the specifier position of a head which bears the relevant φ-features. I further assume (with Pollock) that each head may check at most one feature. One immediate consequence is that we need two separate projections for each argument: one for its θ-properties and one for its φ-relations.

Alexiadou (1997), Laenzlinger (1998) express similar ideas. See Ernst (to appear) for an alternative theory of adverbs, one that arguably fits in more cleanly with my proposal.

Adjuncts in this context refer to adjoined material such as adverbs, not to the operation underlying head movement. The first part of chapter 3 aims to derive three core properties of phrase structure that I assume. First, specifiers and adjuncts are different (only the former check features). Second, specifiers are unique, adjuncts are not (by virtue of natural structural relations which I will discuss). Third, only specifiers can be created by movement or base-generation (substitution), while adjuncts can only ever be base-generated (as the result of Direct Merge). Note that this will buy us the desired dissociation of θ- and φ-features (see also fn. 34). As we will see especially in chapter 4, these assumptions about specifiers and adjuncts are indeed empirically driven and as such natural enough to derive the consequences I desire to derive.
Let us take the $\phi$-domain of the clause to minimally look as follows (where the ordering relation between the two Agr-projections shall not concern us):\(^4\)

\[(36)\]

\[
\text{TP} \quad \text{Spec} \quad T' \quad \text{T} \quad \text{AgrIOP} \quad \text{Spec} \quad \text{AgrIO'} \quad \text{AgrIO} \quad \text{AgrOP} \quad \text{Spec} \quad \text{AgrO'} \quad \text{AgrO} \quad \nu P
\]

We could say that raising of the verbal complex $V$-$\nu$ endows the Agr-heads to check Case on their specifiers and T to check Case on its specifier, and this is indeed roughly the line of Chomsky (1993). However, as all nominal arguments come with two different sets of features that need to be licensed, $\phi$-features and Case, the one-checking-per-projection constraint would be jeopardized. I thus take the driving force behind checking in the $\phi$-domain to be $\phi$-features. All arguments raise to check $\phi$-feature agreement, which may or may not lead to further agreement reflexes (such as agreement with

\(^4\)It has been argued that Agr has intrinsic $\phi$-properties, and there should be an AgrS-projection also (cf. fn. 20). But it has also been argued that SpecTP is not a possible landing site and nominative is assigned in AgrSP on grounds of economy (Chomsky 1991, 1993), and it has been argued that Agr has no intrinsic properties at all, and in the interest of Full Interpretation should be dispensed with entirely (Chomsky 1995a). I take all this confusion to indicate that this part of the clause is rather messy and ill-understood, and as such the suggestions I put on the table might not be completely off the mark. If AgrS exists as a separate head, (or even more than one head, considering clitics), hosting the subject, so be it—as long as this does not incur multiple XP-movement within the $\phi$-domain.
the verb in various positions). But this is not relevant for the issue of anti-locality. What is relevant is the way arguments get Case-marked.

I can think of two reasons why Case should not be the trigger for movement, nor even be checked the same way other features are checked. I suggest that Case-marking is a reflex of the agreement relation between a specifier and its head in this domain (to some extent following Chomsky 1999).

First, it has often been observed that we might want to raise PP-arguments out of their base-generated θ-position (e.g. Bowers 1993, Zwart 1994, Hornstein 1995). If Case were the trigger for such movement, we would either have to postulate that PPs check Case or stipulate that some elements may move and not check Case. Zwart (1994) provides a number of arguments that the West Germanic languages (Dutch, German and their respective dialects) should be head-initial throughout, and as a consequence underlyingly SVO. I cannot review the evidence here, but it is obvious that under such a view all DP and PP-arguments must leave the θ-domain.

As is well known, German (and Dutch) have SVO order in matrix clauses, where the finite verbal element is in second position, and SOV in embedded clauses, where it is last, or at least in a low position.46

45 See Zwart (1993, 1994, 1997a) for extensive discussion. This approach, and its consequences for the grammar of West Germanic, has been disputed by a number of scholars, including (not exhaustively) Gärtner & Steinbach (1994), Schwartz & Vikner (1996), Büring & Hartmann (1996), Müller (1996), Abraham (1997), Haider (1997), Wurmbrand (1998), Abraham & Molnárfi (2000). But it has also been endorsed and applied (see e.g. Haeberli 1995, Haegeman 1996, Alexiadou 1997, Hinterhölzl 1998, and, as relevant for the framework developed here, my own work as cited throughout, such as Grohmann 1995a, 1995b, 1996a, 1996b, 1997c). Some of these authors refute the empirical objections raised in the critical works cited. Theoretically, the approach is at least as sound as assuming that some projections are right-branching, the subject always moves to COMP, and clausal complements obligatorily extrapose.

46 There are dialectal differences whether the finite element has to be last or whether it may precede non-finite elements, an issue that will not concern us here.
To put it crudely, the Zwartian approach takes the arguments to be roughly in the same positions in both types of clauses, while in one type the inflected verbal element moves and in the other it stays put. Relevant for us is the former conjecture. I thus take movement of the PP in (37) to be driven by the need to check $\phi$-features, regardless of whether this agreement relation is then followed by (abstract) Case-marking.47

The second reason to assume $[\phi]$ rather than $[\mathrm{Case}]$ to drive argument movement into the $\phi$-domain concerns a recent proposal in Chomsky (1998, 1999). Basically, Chomsky suggests that some checking relations do not need to be followed by movement (via Agree), but if they are, Case-marking may occur. We could thus treat Case-marking as a “reflex” of sorts, a by-product of an agreement relationship.48,49

---

47 This presupposes a rather liberal view of $\phi$-features, as I will expound in chapter 6. Intuitively, though, it seems more reasonable to endow PPs with $\phi$- rather than Case features. Moreover, many indirect objects expressed as PPs in some languages (such as English) are expressed through Case-marked DPs in others.

48 Or we could think of Case as marking an “address” in the sense used in recent work on dynamic syntax (Uriagereka 1997, 1998; see also Martin & Uriagereka 1999, Uriagereka 1999b, Uriagereka & Martin 1999, or Raposo & Uriagereka 2000). As an “address,” Case indicates to the interfaces that a given element has a specific place in the clause. This is reminiscent of Manzini’s (1992) use of the notion, borrowed from Vergnaud (1985), that Case-assignment is a device to mark arguments.

49 I propose a technical implementation of this reflex in terms of “feature mediation” in chapter 3 under which the highest segment of a complex head is able to check the feature on its specifier. Whether “(formal) feature” is indeed the appropriate term to use in this instance is another issue.
In sum, either way we go, we can maintain a form of the one-feature-per-checking approach. A natural question would, of course, be why Agr (AgrIO, AgrO) is endowed with \( \phi \)-features necessary for objects, while subjects check their \( \phi \)-features on T.

Again, there are at least two possible answers: T is actually not T, but AgrS. This way, there is one head for each argument, an Agr-head, endowed with \( \phi \)-features and able to mark Case (through verb movement). On the other hand, T is obviously a category we would want to keep: tense-marking seems to be a pervasive property of language. And if we assumeaspectual, negation and modal projections, to mention but a few, we would be hard pressed to scrap T. In this case, we could work with both (as Pollock 1989 did, or Chomsky 1991, 1993; see also Cinque 1999). This raises questions as to whether the subject can pass SpecTP without stopping (given the CDE, it cannot target both), and others. Some of the issues will be addressed in chapter 6, others will have to be left open.

### 2.4.3 The \( \omega \)-Domain

COMP is the traditional position of moved Wh-expressions and propositional operators, later expanded to host topics, certain types of focus, and other items. In our terms, the \( \omega \)-domain is endowed with the context variable “discourse.” Elements that are licensed here all play a role for interpretation, topic and focus being prime examples. COMP also types the clause. As I will come back in more detail to all these properties in due time (throughout chapters 4 and 5), I will forego a more detailed presentation at this point (see also the fairly elaborate discussion in section 2.2.3 above).

Following much recent research of the left periphery of the clause, I take COMP, the prolific discourse domain, to look like (38), slightly modified from Rizzi (1997: 297).
Following Rizzi and much related work, I take the highest clausal projection, CP (his ForceP), to type the clause and in most cases host the complementizer. Focused phrases and Wh-interrogatives typically target Foc(us)P, topicalized constituents Top(ic)P. “FP” is a functional projection yet to be defined more clearly, but as a first stab it is a hybrid position hosting clitics in some Romance dialects (Uriagereka 1995a) and possibly in West Germanic, too (Grohmann 2000d). We will see more of FP in chapter 5.

While an articulate structure of CP has been pursued for quite some time (at least since Cardinaletti & Roberts 1991), the inventory of projections is still debated, and so is the hierarchical ordering. I will make a certain amount of idealization regarding the structure of the ω-domain, but not unreasonably, and certainly in line with current assump-
tions (the same goes for the other Prolific Domains). I will address all of the projections mentioned here, especially in chapters 4 and 5, but this part of the clause will turn out to be more intricate than the structure in (38) suggests.

2.4.4 Summary

This concludes our discussion on some pertinent assumptions on feature checking, clause structure and the internal composition of a tripartite system of Prolific Domains. We have seen theoretical and empirical motivation to split the clause into the three domains I suggested, the θ-domain (for thematic relations; vP, VP), the φ-domain (for agreement properties; TP, AgrIOP, AgrOP, plus assorted projections such as NegP, AspP, ModP etc.), and the ω-domain (for discourse information; e.g. CP, TopP, FocP, FP). Naturally, a number of questions arise at this point, some of which I will address in due time. We now have the basic ingredients of a Domain-driven framework.

2.5 Conclusion

Summing up the main proposal, the concept of Prolific Domain follows from the observation that some projections are more closely related to one another than others; I defined this relationship in terms of context (namely, θ for theta, φ for agreement and ω for discourse). An understanding of anti-locality in terms of the Condition on Domain Exclusivity allows us to ban Domain-internal XP-movement, yet rules in X₀-movement, a desired result in light of the above discussion. We will modify the ban on movement of a maximal projection within a single Prolific Domain in chapters 3 and 4 by introducing the concept of Copy Spell Out, a natural extension from the definition of the CDE, which
basically allows Domain-internal movement, as long as it involves distinct PF-matrixes for multiple copies of the same element within a given Prolific Domain.

To visualize, movement of the sort illustrated in (39a), generalized over the entire clause, is illicit, while (39b) is fine; head movement, as shown in (39c), is legitimate:

\[(39)\]

\[
\begin{align*}
\text{a. } & \quad \# \ [\ldots \ YP \ X^0 \ \ldots \ [\ldots \ \& \ YP \ \\ldots]]
\text{b. } & \quad [\ldots \ YP \ X^0 \ \ldots \ [\ldots \ \& \ YP \ \ldots]]
\text{c. } & \quad [\ldots \ Y^0 \cdot X^0 \ \ldots \ [\ldots \ \& \ Y^0 \ \ldots]]
\end{align*}
\]

The conceptual basis of the CDE and Prolific Domains is the target of chapter 6. Everything else should be clarified in the next three chapters. I will start with a discussion of the status of adjuncts in Bare Phrase Structure, addressing uniqueness of specifiers, generation of adjuncts, and X'-relations. I also present first empirical doubts on multiple specifiers employed in much current work, which I will extend in chapters 4 and 5. The second part of the next chapter deals with the second part of the definition of Exclusivity; an apparent caveat built into the conception of Exclusivity, but, as I will argue, a natural one. Moreover, it has empirical consequences, tying in nicely with recent research.
CHAPTER 3

COPY SPELL OUT:
X’-THEORY, PROLIFIC DOMAINS AND LOCAL ANAPHORICITY

Now that the proposal of a Domain-driven approach to syntactic computation is on the table, we can attend to empirical issues. How do the assumptions introduced in the last chapter fit with natural language, how does the Condition on Domain Exclusivity I proposed work out, and is there really motivation for a repair strategy that renders an otherwise illicit movement well-formed? In two parts I address these issues in this chapter: the first part contains a discussion of X’-structure and the specifier vs. adjunct distinction, and the second part illuminates the repair strategy in terms of Copy Spell Out, proposed to save Exclusivity, for local anaphors.

3.1 Introduction

This chapter deals with two main issues: distinguishing (relational and technical) properties of specifiers and adjuncts in a Bare Phrase Structure framework (X’-Theory), and expounding the Condition on Domain Exclusivity; both parts contain adequate theoretical exposition and empirical illustration. Section 3.2 can be viewed as an elaborate intermezzo between the discussion of Prolific Domains and Exclusivity in the previous
chapter on the one hand and an analysis of reflexivization as a first demonstration of the repair strategy Copy Spell Out applied to the θ-domain on the other. I extend specific aspects of X’-structure relevant to the discussion of specifiers vs. adjuncts. I argue that unique specifiers can be derived from the system and that adjunction, possibly multiple, results from Direct Merge only. To help bring the previous chapter and the second part of the present chapter closer together, I summarize the main proposal in section 3.3 before tackling one particular case study. The empirical focus in section 3.4 lies in a first illustration of the proposed constraints and conditions within the θ-domain, reanalysing local reflexivization in terms of the repair strategy Copy Spell Out. Section 3.5 concludes this chapter, sending us off to consider a radical view of the left periphery of the clause.

3.2 (Bare) Phrase Structure: An Empirical and Theoretical Intermezzo

For the most part, the CDE also precludes multiple specifiers: if accusative were checked in (an outer) SpecvP, for example, the cut-off point for the θ-domain would have to be postulated for “right above the first specifier of vP” or something like that. Similar examples can be found for the other Prolific Domains also. In addition, multiple specifiers are not predicted if each head can only check one feature. I address conceptual, empirical and theoretical arguments against multiple specifiers and related issues next. After laying out my motivations to critically consider the issue, I present a set of data that casts serious doubt over the justifications offered to replace Agr with v as the accusative case-marker. Having conceptual and empirical back-up, I then tackle the theoretical validity of specifiers, and ways to distinguish unique specifiers from (multiple) adjuncts. I introduce a version of Bare Phrase Structure that does so, yet keeps the spirit of defining structural identification over relational rather than categorial properties.
3.2.1 *Adjuncts vs. Specifiers*

In Bare Phrase Structure Theory—where apparently only minimal and maximal projections (roughly, a non-projecting element and a fully projected phrase) count for interpretation, and intermediate levels of projection (whether we call this recursive projection X’ or not fully projected XP) can be formally ignored—we face the problem of integrating one major property of the original X’-Theory, namely the distinction between specifiers and adjuncts. Intuitively, these entities are different: adjuncts are purely “optional” (not necessarily in the technical sense), while specifiers are obligatory, much like complements, a fact that should ideally be cashed out in terms of Checking Theory.

In this section, I express why one might want to maintain a structural difference between specifier and adjunct, which basically boils down to the different types of licensing for each: specifiers check matching features with a head, while adjuncts check a feature on themselves. I also propose theoretical and empirical arguments against multiple specifiers in general and in favour of (possibly, multiple) adjunction as the result of base-generation only. The empirical evidence from accusative Case-marking in Dutch boils down to the presence of a position which has purely grammatical function; in other words, whether this position is assumed to be an outer SpecvP or a unique SpecAgrOP does not really matter—what matters is that the original motivation to dispense with AgrP is lost: if AgrP only has grammatical function, while vP has other intrinsic properties (interpretable θ-features), then something should be amiss if we encounter evidence suggesting that there are instances when vP would play no role other than marking accusative case. Finally, I suggest a technical implementation in terms of “most natural” (meaning basic) relations which, moreover, come free with the operation Merge: sisterhood and immediate containment. I lay out how (the composition of) these two relations can be used to force feature checking to take place only between a head and a complement, an adjoined head or a specifier; moreover, it further enforces unique specifiers.
In a nutshell, I modify relational definitions of $X'$-structure, as desired in Bare Phrase Structure Theory (cf. Muysken 1982, Freidin 1992, Chomsky 1994, 1995a). This modification—independently needed, if we take Chomsky’s (1998) suggestions of basic relations that Merge yields for “free” and the composition of these seriously—has at least two relevant consequences: (i) multiple specifiers are undesirable, do not buy us much empirically, and can be banned from the grammar, and (ii) specifiers and adjuncts are distinct relational objects, an assumption that can also be built into a definition of $X'$-structure, has a high empirical pay off, yet loses strict asymmetry (pace Kayne 1994).

The proposal of this section is very strong and makes a number of immediate predictions. In particular, beyond the obvious (that specifiers, but not adjuncts, are unique and that adjunction to XP must be base-generated), it predicts that adjuncts always c-command specifiers. In other words, a left-branch specifier and its head cannot be separated by a left-adjoined adverb. Chapter 4 deals with one phenomenon which receives not only a straightforward analysis conforming to the CDE, but also supports this prediction. However, given that this chapter lays out the main ingredients of my theoretical proposal, I feel that the right place to discuss specifiers and adjuncts is at this point, following the introduction of Prolific Domains and Exclusivity, and preceding the empirical discussion.

3.2.2 $X'$-Structure Considerations

I start off with first considerations why specifiers and adjuncts could be handled differently. I introduce the basic terminology and concepts used in the remainder. To start with the former, I employ “adjunct” in the sense of a phrasal modifier, usually taken to adjoin to $X'$ (adverbs, in the proto-typical case). The $X'$ is recursive, and for reasons that become clear soon, I refer to this level as XP, reserving $X'$ for the node immediately dominating the head and its complement. None of the claims about specifiers vs. adjuncts pertain to head movement, conceived of standardly, and here as well, as adjunction.
The original formulation of the EPP in Chomsky (1981, 1982) said that sentences must have subjects, or in formal terms: SpecTP (to use current labels) must be filled. It says nothing about the necessity of each sentence having temporal modification expressed by a (possibly null) adjunct of time adjoined to TP. Likewise, with the rise of the PISH, it became obligatory that a transitive verb have its specifier filled with the external argument; yet it did not postulate obligatory manner specification, to name one typical VP-adjunct. And neither does an interrogative C-head, which in many languages needs to have a Wh-specifier, also demand an adjunct of sorts, expressing interrogativity.

(Basic) X’-Theory (Jackendoff 1977, Stowell 1981, Chomsky 1986a) had an elegant way of dealing with specifiers and adjuncts: a specifier is sister to X’ and daughter of XP, while an adjunct is sister to and daughter of X’.

1 A concept which Chomsky (1999) returns to (cf. Lasnik, to appear). That is, after years of trying to motivate the EPP by feature checking, analogously to, say, Case checking, the current direction suggests a formal implementation of the original idea, that basically forces certain specifier positions to be filled. Whether this is the right approach, or whether other alternatives exist shall not concern us here. On the latter, see, for example, the analysis in terms of predicate raising of Moro (1989, 1997), picked up by Zwart (1992), den Dikken (1995b), or a more conservative derivational approach investigated by Epstein & Seely (1999), Castillo, Drury & Grohmann (1997, 1999), Boeckx (2000a), Grohmann, Drury & Castillo (2000).

2 Chomsky (1999: 2) expresses similar relations as primitives: the relations Sister and Immediately Contain come for free. Later on, he also admits that “[t]he conceptual and empirical arguments for X’ invisibility are slight” (p. 32). I will return to the significance of both in section 3.2.4.
These relations are illustrated in (1). (Adjuncts are parenthesized to indicate their optionality; directionality of projection is not a concern here.) Naturally, this dichotomy did a lot of empirical work, such as define the type of constituent that could be substituted by one (N’) or so (V’), account for why adjuncts on the same side of a head may be reordered, but not with respect to complements, and help with other constituency tests.

With the rise of functional projections, especially in the nominal domain (Abney 1987), the role of X’ became less clear. However, the empirical facts could plausibly be reformulated in terms of a not fully projected XP-level, replacing X’. The structural modifications we have to make are obvious (e.g. former SpecNP becomes D⁰); I suggest (2) as the relevant X’-theoretic object. I do not want to exclude right-adjunction, but will not discuss it further. I will thus defend the following structure:

---

As will become clear, I am very sympathetic to the theory of adjunction recently proposed by Ernst (1998, to appear) in terms of Weight Theory—even though space does not permit a more detailed presentation—and as such leave the option of right-adjunction open, if not endorsed. Under this theory, adjuncts are not licensed by checking a formal feature but by certain “LF” properties (regulated by a set of rules for the composition of events, propositions, times and predicates) and “PF” properties (for directionality of adjunction and the weight of adjuncts). In a nutshell, Ernst’s theory takes the impact of Full Interpretation seriously in that he considers the alternative to feature checking: if a property of the grammar can only satisfy Full Interpretation by feature checking, the elements in question must enter a checking relationship. But if some other property does not need to check formal features to fulfill Full Interpretation, no checking is required. Adjuncts arguably express properties that do not require formal checking, yet their licensing is driven by the same principles that we employ for all other licensing configurations.
If intermediate levels now do not count anymore (whether X’ or XP is taken to be recursive), the distinction between adjuncts and specifiers falling out from (2) is lost and might suggest one of several consequences: maybe specifiers and adjuncts are the same object and there is no distinction, or maybe both/either can have multiple occurrence in a given phrase, or maybe, if both are the same, their occurrence is unique, and so on.

Many maybe’s, and each one has been considered: Kayne (1994) argues for collapsing the two notions, understanding each left branch as a unique occurrence of an adjoined element, pushed further by Cinque (1999) from a cross-linguistic perspective and a vast inventory of functional projections (basically, one per left branch); the driving force determining phrase structure is the Linear Correspondence Axiom (LCA) which requires the collapse of the two notions. Lasnik & Saito (1992) argue against adjunction (expressing all relations in terms of substitution). Chomsky (1994, 1995a) “proposes” multiple specifiers (drawing from ideas by Koizumi 1994, Ura 1994; see also Ura 1996, Mulders 1996, Richards 1997, 1999 and many others), whose order is determined by certain conditions and adjuncts and specifiers may co-occur in a given phrase (but see fn. 16 be-

\[ \text{(2) } \]
\[
\begin{array}{c}
\text{XP} \\
\text{(Adj)}
\end{array}
\]
\[
\begin{array}{c}
\text{XP} \\
\text{(Adj)}
\end{array}
\]
\[
\begin{array}{c}
\text{Spec} \\
\text{X’}
\end{array}
\]
\[
\begin{array}{c}
\text{X} \\
\text{Comp}
\end{array}
\]

\[ \text{If intermediate levels now do not count anymore (whether X’ or XP is taken to be recursive), the distinction between adjuncts and specifiers falling out from (2) is lost and might suggest one of several consequences: maybe specifiers and adjuncts are the same object and there is no distinction, or maybe both/either can have multiple occurrence in a given phrase, or maybe, if both are the same, their occurrence is unique, and so on.}
\]

\[ \text{Many maybe’s, and each one has been considered: Kayne (1994) argues for collapsing the two notions, understanding each left branch as a unique occurrence of an adjoined element, pushed further by Cinque (1999) from a cross-linguistic perspective and a vast inventory of functional projections (basically, one per left branch); the driving force determining phrase structure is the Linear Correspondence Axiom (LCA) which requires the collapse of the two notions. Lasnik & Saito (1992) argue against adjunction (expressing all relations in terms of substitution). Chomsky (1994, 1995a) “proposes” multiple specifiers (drawing from ideas by Koizumi 1994, Ura 1994; see also Ura 1996, Mulders 1996, Richards 1997, 1999 and many others), whose order is determined by certain conditions and adjuncts and specifiers may co-occur in a given phrase (but see fn. 16 be-}

\[ \text{Actually, they do not express their proposal quite in these terms. So as to not put words into their mouths, Lasnik & Saito (1992: 87, ex. (81b)) state the condition that \text{“[a]djunction creates a separate maximal projection.” This is potentially very interesting for the present framework, if the condition that all adjunction is necessarily base-generated comes off too strong. This modification could easily be integrated.} \]
Within Bare Phrase Structure, Ernst (to appear) proposes a system that distinguishes specifiers from adjuncts not by force, but in that the latter do not need to check features to satisfy Full Interpretation. As such, the two are different, but no appeal to uniqueness is made, nor needed.

To be honest, I am puzzled by the stipulation—and arguably it is nothing more, as we will see shortly—that intermediate projection levels should be not interpretable at LF and hence do not count for the computation. If it really is a stipulation, then replacing the stipulation that multiple specifiers are allowed (and intermediate XP-levels invisible to the computation) with another stipulation that bans multiple specifiers from the system should fare equally well. We will see which stipulation does the better work. In the worst case scenario then I assume that multiple specifiers are not part of C_{HL} by fiat. But I believe we do not have to go that route. I return to this matter in chapter 6 as part of a more general discussion of the framework of Prolific Domains, Exclusivity and the interfaces.

### 3.2.3 On the Interpretability of Agr

First I will present an empirical argument casting doubt on a hardliner’s view that the accusative case checking position must always have other intrinsic properties, then I will address technical issues. In the framework of Chomsky (1995a), multiple specifiers (henceforth, multi-Specs) were employed in three empirical domains of the clause: multi-Specs of vP, of TP, and of CP. The first relate the external argument (agent θ-role) and the object (accusative case) to the same head, v—these days adopted in most minimalist scenarios—, while the latter has been applied to multiple fronted Wh-phrases (as in Bulgarian or Romanian; see especially Richards 1997, Bošković 1999, Pesetsky 2000);

---

5 See also Nunes & Thompson (1998) for a technical exposé on specifiers and adjunction in the “classic” Bare Phrase Structure of the Minimalist Program (drawing on Chomsky 1994, 1995a).
multi-SpecTP has been argued to be invoked at least in “Multiple Subject Constructions,” which are found in Japanese, Korean or Hebrew (Ura 1996, Doron & Heycock 1999), but have also been proposed for Icelandic in the form of Transitive Expletive Constructions (see Chomsky 1995a, drawing from e.g. Bobaljik & Jonas 1996).

I will not address the possibility of multi-SpecTP in this section; firstly, there are reasons to believe that one of the two “subjects” in these constructions is much more adjunct-like, if not even a topic proper, and secondly, I will address at least part of the issue later. I will not discuss multi-SpecCP at this point either; see chapter 5 for a discussion of multiple Wh-fronting. In the following presentation I concentrate on multi-Specs of vP, going over an empirical counter-argument from Dutch—one that reopening the discussion of the place of Agr in C<sub>HL</sub>—, originally due to Zwart (2000).<sup>6,7</sup>

### 3.2.3.1 Setting the stage

Dutch is a West Germanic language with SVO order in subject-initial matrix clauses (and obligatory Verb Second) and SOV in embedded ones. I assume Zwart’s (1993) analysis: all projections are head-initial and all nominal arguments move out of their θ-position in overt syntax to check φ-features and get Case-marked. The ensuing debate concerns accusative case. If V is responsible for checking it from some higher functional projection, how far does V have to raise? There are (at least) two options: V raises to ν, which it does anyway, and assigns Case to an outer specifier of vP. Altern-

---

<sup>6</sup> All data in this section are from Dutch and taken straight from Zwart (forthcoming), as is most of the argumentation and mode of presentation, often near-verbatim.

<sup>7</sup> The following is not so much a justification of Agr—or an “explanation” why we might want to have it and why it appears where it appears—, as it is a critique of the one and only argument against Agr: a purely grammatical element that has no intrinsic, semantic (+Interpretable) properties. I follow common assumptions in taking V to be a θ-assignor (internal argument) and the element that marks Case (accusative) when combined with a functional head (namely, Agr rather than ν); ν assigns the external θ-role. (For expository reasons, I sometimes switch between “Agr” and “AgrO,” wherever it is not critical to finer distinguish agreement.)
tively, the V-v complex raises to a higher functional projection, call it AgrOP, and as-
signs Case to Spec of AgrOP.

Chomsky’s (1995a) main objection to Agr is conceptual: it does not carry any in-
trinsic semantic properties. All Agr comes equipped with are φ-features, which are –
Interpretable, hence irrelevant for the interpretive component. The sole purpose of Agr is
to allow V to assign Case. It thus has a purely grammatical function, as opposed to any
other lexical or functional head. By dropping the assumption of a unique checking rel-
ation per projection, it seems more economical to evoke an additional specifier of an inde-
dependently needed head to check accusative. This head is v whose “intrinsic” property is
thematic. Movement of the object to an outer Spec of vP involves one projection less and
does not run into trouble with the interpretation of Agr. Zwart presents data that cha-
llenge the central presupposition of this line of reasoning. He argues that some construc-
tions involve a head that is responsible for nothing but accusative marking—just like Agr.

Consider a simple transitive sentence in Dutch. The object in (3) appears to the
left of the sentential adverb. Depending on whether we adopt multiple functional projec-
tions (AgrP) or multiple specifiers (of vP), we face the following two options:

\[
\begin{align*}
(3) & \quad \text{a. } [\text{CP dat } [\text{TP ik } [\text{AgrOP Jan } \left[ \text{VP zag Jan} \right]]]]] \\
& \quad \text{b. } [\text{CP dat } [\text{TP ik } \left[ \text{vP Jan } \left[ \text{VP zag Jan} \right] \right]]] \\
& \quad \quad \quad \text{that I saw Jan yesterday.} \\
& \quad \quad \quad \quad \text{‘…. that I saw Jan yesterday.’ (Dutch)}
\end{align*}
\]

\footnote{Note that this type of adverb is unlikely to appear as low as vP. As we will see pre-
seently, a temporal adverb such as gisteren ‘yesterday’ comes in handy, though, to tease
apart different interpretations, forcing the one we will be interested in. For the purpose of
illustration, namely to argue against multi-SpecvP, it does not hurt to adjoin gisteren
‘yesterday’ to vP. Manner adverbs (such as duidelijk ‘clearly’, snel ‘quickly’), which are
more likely to be v/VP-related, would do here, too, but lead to unwanted ambiguities.
Pending a more articulate theory of adverbial positions and insertion into the derivation—
in this or any other framework—might help settle the current argument also, but I leave
this issue for the future and concentrate on minimal assumptions.}
Bearing the availability of these two structures in mind, let us see what happens in Exceptional Case-Marking (ECM) contexts, i.e. contexts which typically involve accusative case-marking of the embedded subject by the matrix predicate.

Perception verbs like see can take infinitival complements. Such configurations employ ECM, even in Dutch. It is clearly the matrix verb that assigns accusative case to the subject of the embedded clause. In (4), Jan can be replaced by hem ‘him’, but not by hij ‘he’. Consider the following, where the embedded verb is italicized:

\[
(4) \quad \ldots \text{dat ik Jan gisteren zag winnen.}
\]
\[
\quad \text{that I Jan yesterday saw win}
\]
\[
\quad \text{‘… that yesterday I saw Jan win.’ (Dutch)}
\]

The adverb indicates that the embedded subject raises into the matrix clause, somehow checking case with the matrix verb zag, while thematically relating to the embedded verb winnen. We can thus assign (4) the following structures, parallel to (3):

\[
(5) \quad \begin{align*}
\text{a. } & [\text{CP dat [TP ik [AgrOP Jan [vP gisteren [vP ik [VP zag [vP Jan [VP winnen]]]]]]]]] \\
\text{b. } & [\text{CP dat [TP ik [vP Jan [vP gisteren [vP ik [VP zag [vP Jan [VP winnen]]]]]]]}
\end{align*}
\]

Recall that under a Spec-Head licensing configuration for Case checking, the matrix accusative case position is most likely the one where it is checked (unlike the INFL position of the embedded clause under some contortion of the government definition in earlier frameworks). We assume that the ECMed subject targets the matrix SpecAgrOP or SpecvP, respectively.\(^9\) So far either Case checking theory can handle the data. It gets more complicated if we add an additional argument to the embedded verb. This object also raises into the matrix clause, beyond the adverb (marked in boldface and italics):

Again, both hypotheses might deal with such cases, by simply adding (another) AgrOP or an additional vP-Spec in the matrix clause. Take (7) to be the options for (6):

(7) a. \[ \text{[CP dat [TP ik [AgrOP Jan [AgrOP de race [vP gisteren [vP ik [vP zag [vP Jan [vP winnen de race]]]]]]]]] \]

b. \[ \text{[CP dat [TP ik [vP Jan [vP de race [vP gisteren [vP ik [vP zag [vP Jan [vP winnen de race]]]]]]]]} \]

Examples like (4) or (6) show that movement into the matrix clause is allowed in ECM instances, and the fact that perception verbs like see are transitive suggests that they indeed have a vP, possibly licensing accusative case (under the multi-Spec approach). If we want to hold on to (some form of) Burzio’s Generalization\(^{10}\) — and to my knowledge, nobody in the “multi-Spec camp” has ever denied it —, we now expect that such movement is only allowed in the ECM-environment of transitive verbs; intransitive verbs (un-accusatives) do not have vP (Chomsky 1995b: 315f.). Again, this is not an unusual assumption which I (or rather, Zwart) adopt; and to my knowledge, multi-Spec supporters have not yet proposed to allow for a vP with these verbs also.

If we could now find environments that involve \(\phi\)-checking (and Case assignment) of an embedded DP in a matrix clause whose verb does not project vP, the multi-Spec approach would face a serious problem: it would look as if the matrix Case position is purely functional and does not involve intrinsic interpretive properties. Moreover, if this movement were the only strategy available, the problem would grow even bigger.

---

\(^{10}\) Burzio (1986) observed that (i) a verb which does not take an external argument does not assign accusative case and (ii) a verb that does not assign accusative case does not take an external argument.
3.2.3.2 Raising into an uninterpretable specifier?

Let us consider movement into the matrix environment of a raising verb.

(8) … dat Jan de race gisteren scheen te zullen winnen.
that Jan the race yesterday seemed to will win
‘… that yesterday Jan seemed to be going to win the race.’  (Dutch)

Two remarks are in order. First, the addition of an auxiliary ensures that the adverb is construed with the matrix clause, the event of seeming, rather than the embedded clause and the event of winning. The adverb refers to a past event, while this auxiliary indicates a future event. This is indicated in the translation; it thus transpires that in English, too, the adverb must be construed with the matrix clause. Second, the Dutch equivalent of seem is also a raising verb, which can be shown with the same diagnostics as for English (it does not have an external argument, it cannot assign Case to an internal argument etc.). As such, it presumably lacks vP and should not be able to license Case.

Alas, it does: the object can be replaced with a pronoun, and this pronoun is marked accusative.\(^\text{11}\) It thus looks like the head that licenses case for the embedded object must bear a purely grammatical function—something AgrO used to do, but not v, which is independently motivated as a θ-role assignor. However, v is not otherwise motivated in (8). If SpecvP is involved, something else must be said. (9) indicates the dilemma ‘?’:

(9) \[ [\text{CP dat [TP Jan [\text{de race [vP gisteren [vP scheen [TP Jan te zullen [vP Jan [vP winnen de race]]]]]]]]]] \]

\(^{11}\) Actually, pronominalized de race ‘the race’ is neuter and does not show a Case distinction. However, if we replace the embedded predicate with something that takes a masculine (de man ‘the man’) or feminine (de vrouw ‘the woman’) argument, we get the distinction between hem ‘him’ or haar ‘her’ vs. *hij ‘he’ or *zij ‘she’. Moreover, this cannot be the “default Case” in Dutch or some other quirk, as we then would expect nominative, as transpires from left dislocation constructions (see chapter 4).
Traditional raising predicates are not the only contexts in which we find sentential complements of a “deficient” predicate of sorts, where by “deficient” I have in mind that does not necessarily make available a full-blown v/VP-structure.

3.2.3.3 Passivization

We know that a passivized verb loses its property of assigning accusative case to its internal argument, the object; moreover, the external argument is at best optional and, if it shows up, it comes as a by-phrase. In other words, we could conjecture that passivized structures lack a vP. Zwart does, so let us see what it could do for us.\(^\text{12}\) The Dutch passives work in the relevant respects just like their English counterparts. Consider (10):

\[(10) \quad \text{… dat Jan \textit{het boek} niet werd geacht te hebben gelesen.} \]

\[\text{that Jan the book not was considered to have read.} \]

‘… that Jan was not considered to have read the book.’ (Dutch)

First, the object of the matrix passivized verb is in the overt subject position, marked nominative; second, the embedded object has raised into the matrix clause where it gets accusative. The diagnostics are familiar: replacing the object by a pronoun shows this clearly (cf. fn. 11), it could not have moved to a position below (such as somewhere

\(^{12}\) Arguably, matters are more complicated. Absence of vP in passives would follow from Burzio’s Generalization if passive verbs do not \(\theta\)-mark their subjects (Chomsky 1981): they do not bear an external \(\theta\)-role and do not assign accusative. But it does not need to. Baker, Johnson & Roberts’s proposal (1989), for example, could be taken to mean that passive verbs actually do have vP to which the by-phrase is adjoined; \(v\) absorbs the \(\theta\)-role of the verb (cf. also Roberts 1986). Presence of vP would distinguish passive from middle verbs (but see Zwart 1997c, 1998; see also Abraham 1995 for discussion of middles in German, Dutch and English). I cannot pursue the issue further (but see Tsimpli 1990 for treating Greek middles like passives, for example). Note, though, that if passive verbs have vP, and SpecvP is not filled, one could envision a derivation proposed by Radford (1997), in which the complement of V moves to SpecTP via SpecvP. This would violate the CDE (Exclusivity), so I dismiss this alternative. Relevant for the present point is that it could be argued, and in fact has been, that passive structures lack the additional structure with the same intrinsic, interpretive properties regular transitive verbs have.
inside the embedded clause), and sentential negation indicates the matrix position, just as the adverb in the cases above, as the translation suggests.

Not leaving out the disclaimer in fn. 12 completely, one could argue that this accusative position is purely grammatical, just as one might expect from Agr. Unless passive verbs really have a functional $v$ for thematic, or any other “intrinsic, interpretive” reasons, it is not clear how this position would be different from Agr.

3.2.3.4 Transitive Expletive Constructions

A third case to consider in this respect are Transitive Expletive Constructions in which the expletive is a grammatical place holder in subject position, not thematically selected; the matrix verb in (11) is a raising verb which neither takes an external argument nor projects a $vP$, yet the matrix expletive, the lexical subject (‘associate’) selected by the embedded transitive verb, and its object all sit obligatorily in the matrix clause:

(11) … dat er iemand het huis gisteren scheen te zullen kopen.

that there someone the house yesterday seemed to will buy
‘… that someone seemed yesterday to be going to buy the house.’ (Dutch)

The argument is clear: the embedded subject receives nominative in the matrix clause and the embedded object accusative, where the adverb is the standard diagnostic for position. The matrix verb does not have a $vP$ where accusative could be checked, so the only possible analysis is that the object has moved to a separate functional projection responsible for accusative case checking (such as AgrOP).

3.2.3.5 Some concepts and consequences

The above discussion has shown that some verbs can assign accusative to the object of a complement clause, even if the Case-marking verb lacks a thematic $vP$. What
makes the predicates above (i.e. perception and raising verbs) interesting is that they are restructuring verbs.\textsuperscript{13} Without ascribing to a particular analysis of restructuring (see Wurmbrand 1998 for detailed discussion and references), it seems as if one property is Case-marking of an embedded verb in the matrix clause, and in the matrix clause only.

Zwart (forthcoming) formalizes this generalization roughly as follows: if a verb has an external argument, or if it is a restructuring verb and the verb in its complement domain has an external argument, it can license accusative case to an object in its functional domain. He takes it as a given that a verb without an explicit external θ-role also lacks \( v \)P—it might be an elegant technical implementation, but it is not the only one.

One could suppose that in these cases \( v \) is actually present (such as on the raising verb above), but this \( v \) does not play any role other than marking accusative on the embedded object. This, in turn, could be empirically supported with Zwart’s generalization. We could then propose that accusative case is always marked in Spec\( v \)P and if \( v \) does not bear a θ-function, the restructuring default kicks in.

This would be an unfortunate move, however. Chomsky (1995a) wants to get rid of Agr for one main reason: Agr only bears \( \phi \)-features, these are –Interpretable and will be deleted after checking, hence Agr should be invisible at LF. Rather than dealing with invisible entities at the interpretive interface, accusative could be checked by an element with intrinsic interpretable features, such as \( v \). If, however, some structures have a \( v \) which lacks such thematic properties, it would have the same purely grammatical function as Agr, be by definition invisible at LF, and face the same conceptual problem.

Zwart provides three arguments against assuming \( v \)P in the contexts above. The first one is similar to the point just made. At least with raising verbs \( v \) is not semantically

\textsuperscript{13} Apart from raising and ECM-verbs, Dutch has two more restructuring verb classes, causatives and some control verbs. These are irrelevant for the present discussion, however, as both assign an external θ-role (and are thus on a par with ECM-verbs).
motivated, so merging unmotivated v with VP would be an instance of a global operation, a ‘look ahead’ mechanism, to license some relation further along the road.

The second argument concerns the specific circumstances. Consider (12):

(12) … dat Jan niet scheen te sterven/worden gearresteerd/dansen
that Jan not seemed to die/be arrested/dance
‘… that Jan did not seem to die/be arrested/dance.’  

(Dutch)

According to Zwart’s generalization, the matrix accusative position should only be active in the context of a transitive embedded verb. If the embedded verb is intransitive, it should be inactive. (12) shows that the generalization is not ad hoc but empirically grounded: the subject of an unaccusative, passive or unergative verb raises to the matrix subject position (and receives nominative)—it does not raise to the object position to get accusative. If the latter were to be identified as the outer Spec of vP, something else needs to be said why as to it cannot be an appropriate licensing position in these contexts.

Lastly, the entire restructuring complex is subject to Burzio’s Generalization. In (13a), the embedded object moves into the matrix object (accusative) position; in (13b), it moves into the matrix subject (nominative) position:

(13) a. … dat Jan de race niet scheen te winnen
that Jan the race not seemed to win
‘… that Jan didn’t seem to win the race.’

b. … dat de race (door Jan) niet scheen te worden gewonnen.
that the race (by Jan) not seemed to be won
‘… that the race didn’t seem to be won (by Jan).’  

(Dutch)

The difference between the two sentences is that the embedded verb of (13b) is passivized, hence does not assign an external θ-role, and by Burzio’s Generalization, no accusative should be assigned. This suggests further that it is transitivity of the embedded
predicate that is relevant in restructuring contexts, which could then be captured, if the potential for licensing accusative comes from the embedded v, with v being responsible for projecting Agr (which in restructuring contexts takes place in the matrix clause).

In sum, we now have an empirical reason to doubt the conceptual condemnation of Agr: it does not really seem to be the case that the (accusative) Case-assignor always has intrinsic +Interpretable features. The constructions above indicate that were v able to mark accusative on its specifier, it would do so without any interpretable properties in some circumstances. These circumstances involve restructuring contexts. Thus, there is no conceptual reason anymore to dismiss Agr as a possible functional head that only plays a grammatical role. This does not imply that Agr is a better choice as accusative-marker than v, but it levels the difference. Furthermore, there might be empirical pay-off of an Agr-based Case-marking (and φ-checking) framework.

The argument Chomsky raises against Agr is even more puzzling in the context that Chametzky (2000: 149) notes: “In the tradition as represented by Speas [(1990)] and Abney [(1987)], the semantic interpretability of [f]unctional [c]ategories is always understood to be dependent on some [l]exical [c]ategory or other: it is ‘parasitic’ or ‘second-order’.”14 Next we will see another proclaimed instance of “invisibility for interpretation,” namely of intermediate projection levels. The arguments for that have recently been dropped; maybe Agr can be resurrected from fallen grace on the same grounds.

14 Anna Roussou (p.c.) points out that there is a straightforward line of reasoning to consider Agr endowed with +Interpretable features (cf. fn. 43 of chapter 2 and especially Roberts & Roussou 1999). If our proposal regarding argument movement in terms of φ-feature (rather than Case) checking is on the right track, we would expect Agr to play an active role in the realization of φ-properties. We could push this a little bit further and think of Agr as the PF-licenser for arguments: in case its specifier is a DP, that spells out as the full nominal; in other instances, it might spell out as a clitic. I return to such a view of clitics at the end of chapter 4 (see also fn. 20 of chapter 2).
3.2.4 Phrase Structure and natural Relations: Specifiers vs. Adjuncts

Alongside the claim that phrases may not have more than one specifier, I also want to push the idea that adjunction, though theoretically unlimited in number, is restricted to base-generation; that is to say, movement qua Copy plus Merge cannot adjoin a phrase YP to some projection XP, it can only be merged to X’ and form SpecXP. As mentioned earlier, head adjunction is different from phrasal adjunction in this respect, and for good reasons, as we will see. I provide empirical justification for the theoretical discussion that follows. This section fits in well, however, with the discussion of clause structure, Prolific Domains and Exclusivity above, preparing us to leave the trail of assumptions and speculations very shortly.

3.2.4.1 X’-structure and “invisibility”

Recall the structural relations I suggest, extended from (2):\(^{15}\)

\[
\text{(14)} \quad \begin{array}{c}
\text{XP} \\
\text{AP} \quad \text{XP} \\
\text{BP} \quad \text{XP} \\
\text{ZP} \quad \text{X'} \\
\text{X}^{0\text{max}} \quad \text{YP} \\
\text{Y}^{0} \quad \text{X}^{0}
\end{array}
\]

\(^{15}\) By convention, I label the first projection of a head X.X’ and any subsequent, recursive levels of projection XP. While it might be confusing in light of the better known “X-bar recursion,” I believe it is a more appropriate notation in the current context and I hope it does not confuse the reader too much.
YP is the complement of X ($X^{0\text{max}}$) and ZP its unique specifier; Y ($Y^{0}$) has adjoined to X, and AP and BP are adjoined to (or are, in the sense used throughout, adjuncts of) XP. Structural relations must be defined as to allow Y, YP and ZP to enter into a checking relation with X, and only these. Likewise, we need to derive that AP and BP can only be base-generated in the positions indicated, that is, all adjuncts must be the result of Direct Merge only; if they could be derived, we would expect them to enter into a checking relation with the head (movement being driven by Greed); see chapter 4 for empirical support. This expectation just does not seem to fit with current assumptions. Moreover, we cannot uphold this expectation if the sketch of X'-relations I present here is on the right track. These are the desiderata. Let us go and derive them.

In the original formulation of Bare Phrase Structure (Chomsky 1994, 1995a), the relevant projection (intermediate, not fully projected X’ or, as used here, XP) was different from the element originally merged to ($X^{0}$) and the final projected phrase (XP or $X^{\text{max}}$) only in being neither minimal nor maximal. As such it was stipulated to be invisible to interpretation, as only $X^{\text{min}}$ (the terminal element) and $X^{\text{max}}$ (the fully projected phrase) are interpretable objects—apparently following from bare output conditions; see Chomsky (1995b: 242f.). Building on Muysken (1982), minimal and maximal projections are identified by relational properties of categories only (i.e. Spec, Comp and Adj, depending on the relation between these elements and the Head or its projections; see also Freidin 1992). These relations, and only these, basically yield (2) or (14)—without, though, giving the X’ any interpretive status. This allowed the original minimalist framework (and its extensions in Chomsky 1998, 1999) to rule in multi-Specs.\textsuperscript{16} It distinguished between adjunction (of heads or adjuncts) and substitution (specifiers) in that the former creates a two-segment category, whereas the latter forms a new category.

\textsuperscript{16} This conclusion might be wrong, at least the way I put it. As Chomsky (1999: 39, note 66) puts it, “[i]t is sometimes supposed that [multi-Spec] is a stipulation, but that is to mistake history for logic.”
Consider first the proclaimed “invisibility” of intermediate, not fully projected elements, here taken to be a unique X’ and all XPs dominated by the highest, fully projected XP. If we could remove the stipulation that these elements are not invisible, we could easily enforce unique specifiers by stipulating, in turn, that aspecifier must merge with X’, and that there is only one X’ per projection. I will try to do more than stipulate, but if all else fails, this position might not be the most unreasonable one to hold.

There are two arguments for X’ invisibility, one conceptual and one empirical. The conceptual argument is that X’ is not interpreted at LF. In support, Chomsky (1995b: 382, note 24) cites the works of Fukui (1986), Speas (1986), Oishi (1990), and Freidin (1992). He also tries to justify the oddness of an element being “present but invisible” (op. cit.) from a derivational perspective (acknowledging Sam Epstein) by virtue of these nodes being “fossils” (Chomsky’s quotation marks). By this he means that they were visible at some point, namely prior to the operation that turned them invisible. The empirical argument is a partial recreation of the LCA of Kayne (1994), without the need to adopt other ingredients of Kayne’s program (which are basically incompatible with Checking Theory).

In a recent paper, Chomsky (1999: 32) notes that “[t]he conceptual and empirical arguments for X’ [here, XP] invisibility are slight.” The lack of LF-interpretation of X’, or our XP-level, he continues, “is questionable and in fact rejected in standard approaches.” Regarding a minimalist implementation of the LCA, he remarks that “that result, if desired, could just as well be achieved by defining ‘asymmetric c-command’ to exclude (X’, YP),” adding “a stipulation, but not more so than X’ invisibility.”

We could then simply stipulate that the first projection of a head is X’, subsequent projections are XP, and only X’ accepts a specifier. But let us try to derive these results, or at least motivate them with coherent assumptions and conditions, much in line with the minimalist spirit of Bare Phrase Structure.
3.2.4.2 *X’*-structure and natural relations

It is my understanding that standard Bare Phrase Structure (Chomsky 1994, following Muysken 1982), and virtually all its applications, assume X’ to be invisible for interpretation, hence not relevant for expressing structural relations—despite Chomsky’s (1999) nonchalant remarks. The reason I am pounding on the latter is the following. If phrase structure should be expressed in terms of “relational properties of categories, not properties inherent to them” (Chomsky 1995b: 242), recourse to invisibility of some objects in the phrase marker need not be an issue at all—especially not, if labels do not exist in the first place (Collins 1999). We can define these objects, and as such the structure of a projection, with natural relations. As it happens, Chomsky suggests something very similar himself, and I am going to explore these relations.

As one of the conditions of “good design” of language\(^\text{17}\) Chomsky (1998: 27) lists “[r]elations that enter into CHL either (i) are imposed by legibility conditions, or (ii) fall out in some natural way from the computational process.” Regarding (ii), he suggests that Merge yields two relations for free, *Sister* and *Immediately Contain* (p. 31, also Chomsky 1999).\(^\text{18}\) Let us assume, maybe not innocently, that this is so. If these two relations come for free, they are arguably the most natural relations to express phrase structure (under the guidance of Muysken’s suggestion). Chomsky suggests that by merging the objects \(\alpha\) and \(\beta\), forming the new object \(K(\alpha, \beta)\), we can understand Sister to hold of \((\alpha, \beta)\) and Immediately Contain of \((K, \alpha)\), \((K, \beta)\) and \((K, K)\), if Immediately Contain is reflexive.

If Sister and Immediately Contain are the most natural relations—most natural because they are the only direct relational result of merging two objects and as such come

\(^{17}\) One aspect of such conditions is the above mentioned guiding principle, formulated as “less machinery is better than more” by Chomsky (1998: 27, fn. 61).

\(^{18}\) Note that this is already a non-trivial premise, which I am nevertheless willing to accept. One might pursue another route, replacing sisterhood (and related dominance) relations with a primitive notion of c-command. See e.g. Frank, Hagstrom & Vijay-Shankar (1999), Frank (forthcoming) for exposition.
for free (as I interpret Chomsky)—, it might be natural to assume that they play an elementary role in defining certain relations. One such relation is structural, among the objects of a given phrase. Another relation to be addressed is operational, say, to establish relevant checking configurations. Suppose this is indeed so, and suppose that one way of extending these most natural relations is by applying “the elementary operation of composition of relations,” as Chomsky (1998: 31) does. Such an application, “in all possible ways” (op. cit.), yields the new relations (15i) and (15ii). There is also a third relation—however not the one Chomsky suggests. I call this relation Extended Sister:

\[(15) \begin{align*}
    \text{i. } \text{Contain: transitive closure of (immediate-contain)} \\
    \text{ii. } \text{Identity: (sister(sister))} \\
    \text{iii. } \text{Extended Sister: (sister(immediately-contain))}
\end{align*} \]

If we take Sister and Immediate Contain to be primitives, and the first application of composition to be the next “most natural” relations—regardless of why this is or should be so (see Uriagereka 1999c)—, we should try and explore how far we can push these five relations to define the most local configurations in C\(_{HL}\). The area I want to concentrate on is an appropriate checking configuration in these terms.

### 3.2.4.3 \(X^\prime\)-structure and checking configurations

Features are arguably checked in very local relationships, and all evidence so far suggests that we want to include Head-Comp, Head-Head and Spec-Head configurations to be admissible, but no other (Chomsky 1995a). Chomsky (1993) defines a checking

\[\text{That (15iii) should be the third relation was also pointed out by Uriagereka (1999c). Chomsky’s original “third relation” was e-command, hence more or less trivially derived: (sister(contain)). This is not the case, however, as Contain itself is not a primitive but a derived relation. As Uriagereka points out, Chomsky’s third relation does not have the strictly local character Extended Sister has. The next sub-section deals with the possible relevance of this state of affairs to our discussion.}\]
domain which derives the desired results. However, it also allows more than one specifier to be within the checking domain of a relevant head as well as adjuncts (which would simply not check a feature by stipulation). In the framework I propose neither one is desired, not even acceptable. So let us consider a way of replacing Chomsky’s checking domain with an alternative way of capturing the checking configurations endorsed here in purely relational terms (analogous to the checking domain).

Given the natural relations discussed above, we could define feature checking with a Checking Condition along the following lines:

(16) **Checking Condition**
A head H endowed with feature F can enter into a checking relation with an object O in the phrase marker with matching F under a Natural Relation.

(17) **Natural Relation**
Let a Natural Relation be
i. any of the primitive relations provided by Merge and
ii. any relation resulting from the first-order composition of primitive relations.

We are thus dealing with five Natural Relations: Sister, Immediately Contain, Contain, Identity, and Extended Sister. Returning to (14), here in yet another modified

---

20 The checking domain is defined as an “elsewhere’ set” (Chomsky 1993: 12): it is the minimal residue of a domain of a head. The “domain” is evaluated over “the set of nodes contained in [the least full-category maximal projection] dominating the [head]” (p. 11), and the “minimal residue” is a “minimal subset” of the domain minus the “complement domain.” In essence, this allows a head to enter a checking relation with anything adjoined to it, with its complement, with its specifier(s)—formally not different from adjuncts—and also with the highest XP adjoined to specifiers (or adjuncts). We will modify these relations somewhat.

21 Capitalization of natural relations identifies this as a technical term in the sense outlined above—and subject to the usual disclaimer: if these are primitives, why so, why only these, why only the first-order composition of relations and so on. For the sake of discussion, let us assume that the first-order application of composition is the most basic one. This is a good place to thank John Drury for provoking discussions.
form to be as explicit as possible, we can now ensure that the three desired configurations, and only those (for better or worse), are permissible checking configurations:

(18) \[ \begin{array}{c}
\text{XP} \\
\text{AP} \\
\text{WP} \\
\text{X'} \\
\text{X}^{0_{\text{max}}} \\
\text{Y}^{0_{\text{max}}} \\
\text{Z}^{0} \\
\text{Y}^{0} \\
\end{array} \]

According to (16) \( X, Y \) and \( Z \) are heads which bear features that require checking, under the familiar umbrella of Full Interpretation, Greed and Economy, i.e. if no features are checked, movement should not occur. \( YP \) is in CompXP and hence must enter a checking relation with \( X \), if Merge is costly. \( WP \) is in SpecXP and must also enter into a checking relation with \( X \), for either one of the above reasons. (\( ZP \) is irrelevant for our discussion.) If \( AP \) is a specifier, it too must check on \( X \); if it is an adjunct (as assumed here), it need not, for reasons we get back to momentarily. Let us run through the desired and undesired checking relations and the predictions of applying the Checking Condition.

The most straightforward is presumably Head-Comp, as the Natural Relation Sister is an immediate fall-out from the application Merge. Given that \( X \) and \( YP \) above should enter into a checking relation, if we take (at least one understanding of) “selection” to be expressed this way, and that \( X \) is Sister to \( YP \), this is the first desired result: Head-Comp checking is licensed by Sister.
Once X and YP are licensed, we can move (the complex head) Y to X, an instance of head-to-head adjunction. There are two possibilities: \( Y^{0\text{max}} \) and \( X^0 \) enter into a checking relation (Sister) or \( Y^{0\text{max}} \) and \( X^{0\text{max}} \) do (Immediate Contain). Assume the latter, for reasons I address shortly. Head-Head checking is thus licensed by Immediate Contain.

Merging WP with X’, the label of the object \( (X^{0\text{max}}, \text{YP}) \), should ideally result in licit Spec-Head licensing. It does: WP is Sister to X’ which, in turn, Immediately Contains \( X^{0\text{max}} \), and \( X^{0\text{max}} \) is the Extended Sister of WP, one of the three results of the composition of our two primitive relations. Spec-Head checking is now legitimized by Extended Sister.

If we then merge AP and XP—which is the label of \( (WP, X') \), or more precisely, XP is the ordered set \( \{(X, (WP, (X, (X, YP)))\}) \), regardless of the label of the intermediate level of projection—, we should be able to find a Natural Relation between AP and X if the two are to enter into a checking relation. Alas, we do not find such a relation. The composition \( (\text{sister(}AP)) \) returns XP, \( (\text{immediate-contain(}AP)) \) gives nothing relevant (AP’s internal structure), \( (\text{contain(}AP)) \) is equally uninteresting, \( (\text{identity(}AP)) \) yields AP, and \( (\text{extended-sister(}AP)) \) churns out X’, not a head. We thus take AP to be unable to check a feature—at least not with the head of the projection.

In other words, licensing of AP is of a different nature than licensing of WP. If WP is indeed a specifier, we recreate the Spec-Head configuration and exclude multiple specifiers trivially: no element merged to any position above \( (WP, X') \) can enter into a Natural Relation with X and hence cannot check off a feature with X. It follows that AP is an adjunct, and that adjuncts do not enter checking configurations with heads.

The one non-obvious step above concerns Head-Head configurations. I suggested that Immediate Contain is the relevant configuration. Let us see how and why.

First, consider Y and X above. If the relationship were Sister, rather than Immediate Contain, we would expect that Z and Y are also licensed by Sister. Looking at the structure above, however, we see that \( Z^0 \) is in Sister relation with \( Y^0 \), whereas \( Y^{0\text{max}} \) is
Sister to $X^0$. We could imagine that $Y^0$ and $Y^{0\text{max}}$ have identical features, perhaps even duplicate, or that in $[Z^0-Y^0]-Y^{0\text{max}}$ the feature checked between $Z^0$ and $Y^0$ is activated on $Y^{0\text{max}}$ by some other mechanism, thus removing the need to multiply features. Neither option is impossible, and I opt for the latter, calling the mechanism “mediation.”

The conceptual advantage of Immediate Contain as the relevant relation is that it concerns the part of the two-segment head that projects, which could thus be conceived of as being more “active” in the derivational process. This then allows a higher segment of a complex head to enter into a checking relation with anything outside that head. A simple head may check a feature with its Spec as described above. But what if WP needs to be in a checking configuration with $Y^0$? In this case, $X^{0\text{max}}$ acts as the intermediary, by virtue of Containing $Y^0$. As we do not want to stack Natural Relations, I suggest that the relevant property of $Y^0$ can be mediated to $X^{0\text{max}}$ and as such be marked on WP.

The obvious case in hand is, of course, Case-marking. Suppose $X = \text{Agr}$, $Y = \nu$, and $Z = V$, with WP = DP. X checks $\phi$-features on WP (Extended Sister). Apart from head-internal licensing, Z can mark Case on WP via the intermediary X (Contain). Convoluted this may sound, it buys us another earlier assumption: Case-marking is not formal feature checking per se, but rather a reflex of a local configuration.

This subtle difference can now be expressed more formally.

(19) **Feature Mediation**

In a structure $[\alpha_n [\alpha_m \ldots \alpha_2 \alpha_1 \alpha_n ]]$, where $\alpha$ is a head and $\alpha_n$ Contains all $\alpha$, $\alpha_1$, $\alpha_2$ … $\alpha_{m-1}$, $\alpha_m$ mediate a feature F to $\alpha_n$, and $\alpha_n$ becomes intermediary to mark F as a reflex of a licensed checking configuration.

Feature mediation could thus be viewed as an escape hatch for the one-feature-per-projection checking restriction. This makes potentially a number of predictions, in particular that this strategy should not be on a par with feature-checking, i.e. it should not
be thought of as an alternative checking operation (Occam’s Razor). Case-marking conceived as a reflex of a feature is one such instance. We will see more cases in chapter 5.22

We can summarize the relevant relations and arising checking configurations:

(20) Checking Configurations
i. Head-Comp:  Sister
ii. Head-Head:  Immediate Contain
iii. Spec-Head:  Extended Sister
iv. F-mediation:  Contain
v. XP-adjunction:  Identity

We went through all relations except for (20v). As suggested, licensing of adjuncts if different from specifiers, and one way to express this difference is to deny them an appropriate checking configuration. On the other hand, as not every adjunct can just adjoin anywhere without restrictions, we want some control operation. This could be checking on itself. By Identity, then, an adjunct checks a feature on itself.23

As for the undesired configurations, we can ignore the fact that X’ Immediately Contains X or that XP Contains X, as both are projections of X and checking is not needed. Likewise, we can glance over the fact that XP Immediately Contains WP, as neither is a head and as such does not need check features either. As far as I can see, we can ignore all other hypothetical relations also, on the same grounds.

It might be the case that this modification of the original discussion of Chomsky (1994, viz. 1995b: 241-249) does not amount to much more than fancy words expressing

---

22 As can be tested trivially, Feature mediation can only ever have an effect on a complex head and its specifier, i.e. specifiers are the only elements that can be marked qua reflex.

23 Norbert Hornstein (p.c.) reports Jairo Nunes to have suggested that an adjunct checks a feature on itself. I cannot delve into any deeper discussion. The point is whether we want adjunction to be licensed in terms of checking or not (cf. Ernst, to appear), Identity, as a Natural Relation, might be expected to do something, and this could be one option. Again, this is not a necessary conclusion to draw, but a possible consequence.
that “a specifier is the object merged with X’ which immediately dominates X, an adjunct is an object merged with an intermediate projection of X which does not immediately dominate X.” If X’ (or XP) is not invisible, this might be the most straightforward way to implement our desiderata. I hope to have given my stipulations and assumptions, replacing some standard relational notions, a little bit of validity, enough to keep us going.

3.2.4.4 X’-structure and licensing conditions

The core property of distinction between specifiers and adjuncts in traditional terms is that the latter are purely optional: nothing hinges on insertion of an adjunct—if it is not part of the LA, it does not show up, yet the derivation will converge, all other things being equal. But if the object fails to raise to check φ-features and receive Case, or if the Wh-element in English does not undergo movement to check off [Wh], the derivation crashes. On the other hand, not all adjuncts are licensed equally well in all positions. This has cast some doubt on the above reasoning (see Cinque 1999 for a radical implementation of obligatory adverbial positions, or Alexiadou 1997, Laenzlinger 1998). However, if adjuncts and specifiers are formally different, such as along the lines sketched above, we could imagine that the licensing conditions of both are different, too.

If we take obligatoriness to be reflected in the syntax, we might want to tie it to feature checking. All displacement takes place to check features in the appropriate licensing configuration, and by definition, a head is involved. As we have seen above, adjuncts are not in a relation to heads that would permit exchange of feature values. But it is not unreasonable to assume that they need to be licensed somehow, and in that case we would still want to express this syntactically. I thus take either suggestion from above as a viable option, either in terms of Identity, where adjuncts check a feature on themselves.

A number of licensing conditions have been evoked over the years, and even if we assume that the licensing is not in a Spec-Head agreement (pace Kayne 1994, Cinque
1999), we still have a wide range of proposals to consider, and I will leave the discussion at that. See, for example, Jackendoff (1972), Ernst (1984) and relevant references cited for early proposals, and Laenzlinger (1998), Svenonius (2000) for more recent versions. Ernst (to appear) appeals to Full Interpretation as relieving adjuncts from the necessity of entering any checking relation. If this goes through, Identity can simply be considered the saturation relation for Full Interpretation, no formal checking is required.

3.3 Interim Summary

To recap the main points of the main ingredients of my proposal so far, I propose that clause structure is made up of (up to) three Prolific Domains: the $\theta$-domain, licensing theta relations; the $\phi$-domain, licensing agreement properties; the $\omega$-domain, licensing discourse information. Each Domain consists of internal structure, hence is Prolific, and all projections within a Prolific Domain $\Pi\Delta$ are identified by virtue of a context variable (namely, $|\theta|$, $|\phi|$ and $|\omega|$, respectively). These are part of the internal make-up of lexical and functional heads, and as such percolate up to their maximal projection (possibly in the sense of an “Extended Projection;” cf. Grimshaw 1990, Ernst, to appear).

I take the following structure to be an accurate (if partial) representation of a full clause, where I highlight the $\omega$-domain, anticipating chapters 4 and 5:
As understood so far, the CDE limits occurrence of an object in the phrase marker to one per \( \Pi \Delta \). Occurrence is understood in terms of copies which, in turn, are evaluated over distinctness (Chomsky 1995a, Nunes 1995). As such, the copy of a maximal phrase is non-distinct from the original, while a copy of a head is distinct. This could also follow from the assumption that an XP is copied and re-merged qua substitution, whereas an \( X^0 \) adjoins to a higher head and forms a complex head. The object in the phrase marker of the original head is whatever it looked like, say \([X^0]\), while the object of a more complex head looks like \([X^0-Y^0-Y^{0\text{max}}]\). An XP, say a DP, looks like \([_{\text{DP}} D [_{\text{NP}} N (\text{XP})]]\) in both cases, prior to copying and after moving. (See more elaborate discussion in section 3.4.1).

Taking the relevant parts from (21), (22) presents two admissible derivations, while the permutations in (23) and (24) all violate the CDE and are hence ruled out:
In (22a), both arguments are merged into their theta positions and receive their respective $\theta$-roles (where it the original position of the DO does not matter, as long as it is SpecVP in ditransitive constructions). The SU raises to SpecTP to check its $\phi$-features and receive Case, and the DO moves on to SpecAgrP to do so (either overtly or covertly, hence the parentheses). We also have a temporal adverb which, for simplicity, I assume to sit in AdjTP. (I ignore verb movement and abbreviate it with a coindexed trace.)

Alternative derivations, such as illustrated in (23), are all illicit. In (23a), the DO raises to an outer specifier of vP, which is in the same $\Pi\Delta$ as its copy, and thus violates the CDE. In (23b), the adverb is in SpecTopP; while this is not excluded per se (the corresponding German sentence would presumably have a structure very much like this), it is if it did not move from an adjoined position in its licensing domain. I take the licensing domain of temporal adverbs to be the $\phi$-domain, in particular around T, and applying our distinction between specifiers and adjuncts it follows that base-generation of adverbs must take place in adjoined position, not in a checking configuration. This is illustrated further in (23c), where the adjunct is merged with a T-head (where T1 and T2 could be two distinct temporal heads, à la Cinque 1999), but in specifier position, hence an obligatory checking configuration.
An appropriate topic structure like (22b) raises the element to be topicalised into a designated topic position, SpecTopP. In languages like German, the finite verbal element also raises to Top; in English, it does not. Moreover, if the topic is an argument, it must have moved through its $\phi$-position—regardless of whether it does so overtly in non-topic instances (cf. (22a)). With TopP part of the $\omega$-domain, the movement from AgrP to TopP is legitimate (as are all other movements indicated).

An illegitimate movement is illustrated in (24a): here, the topic adjoins to TP, which it cannot do for four reasons. First, an adjoined position can only ever be the result of Direct Merge, never of Move (Copy plus Merge). Second, even if it could, AgrP and TP are in the same $\Pi\Delta$, so it would violate Exclusivity. Third, AdjP is not a proper checking configuration and this movement would not be able to check a feature. Fourth, even if no feature checking were required for topicalization, and Mary could be base-generated in AdjTP, there would be two copies of Mary in the same $\Pi\Delta$. (We will see the relevance of this point in section 3.4.) (24b) and (24c) are also bad for the now obvious reasons: the outer SpecvP is not a Case position (there is no checking configuration for multi-Specs), and even if it were, the CDE could be argued to be violated in either case.

We have seen empirical evidence against multiple specifiers, and we have considered a technical implementation to rule them out on principled grounds. If not successful (that is, if they cannot be derived one way or another), I can at least make the following stipulations, which are roughly equivalent to additional assumptions standard Bare Phrase Structure must make.24

---

24 Anna Roussou (p.c.) observes that by (25b) and (25a), adjuncts could be seen as a subset of specifiers, all else being equal (such as differences in feature checking). I shall leave open for now whether this possible state of affairs is an accident of the system, or whether it expresses something deeper.
(25) 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

A more interesting question regarding Prolific Domains is their conceptual motivation and their place in the grammar, or at least their integration into the computational system $C_{\text{HL}}$, as currently perceived. This issue will be discussed in chapter 6. Chapters 4 and 5 will lay out the empirical base for Prolific Domains and the Condition on Exclusivity (along other, novel and in principle independent approaches to left dislocation constructions and Wh-question formation). The CDE, in turn, will be reconsidered in section 3.4, where I also discuss some issues pertaining to the $\theta$-domain. The overall empirical focus is the $\omega$-domain, hence the subtitle of the dissertation, with a suggestion to extend the proposal to the $\phi$-domain in a few cases, especially one considered in chapter 4.

Before going there, however, I shall examine the hypothesis that movement within a Prolific Domain cannot take place, unless it employs a repair strategy to in terms of spelling out the lower copy as a pronominal element. Our first phenomenon will be reflexivization and related issues.

3.4 Exclusivity: Derivational Reflexivization and Copy Spell Out

I propose that clause structure, beyond its intricacies of functional projections and syntactic operations in $C_{\text{HL}}$, can be divided into three Prolific Domains. One property of a Prolific Domain is the unique occurrence of an object in the phrase marker, as formulated in the definition of the Condition on Domain Exclusivity in the previous chapter. This
section deals with part (ii) of the CDE and presents some empirical evidence for *Copy Spell Out* from local reflexives in English.

### 3.4.1 Completing the CDE

In this section, I would like to suggest that in some cases the CDE can be violated—just in case the lower copy has a different PF-matrix. This part of the CDE is a repair strategy that spells out the lower copy. The empirical base presented in this section comes from movement within the θ-domain, shown initially to be ruled out. But a derivational approach to reflexivization—endorsed here, with appropriate discussion—can be pursued, if the caveat of Copy Spell Out is allowed on principled grounds.

These principled grounds will be the topic of chapter 6. More empirical evidence will be presented in chapter 4, where I subject left dislocation phenomena to an examination in derivational terms. I propose an analysis of those constructions that arguably involve movement of the left-dislocated element analogously to reflexivization: some XP moves from one position to another within the same Prolific Domain and spells out its lower copy. The relevant Prolific Domains considered there are the ω-domain, and also the φ-domain. We will then have accrued core instances of the revised CDE across all parts of the clause. Moreover, if the same instantiation of the revised CDE can be found in different constructions and languages, the caveat loses its initial ad hoc character.

To dive right in, let me repeat the CDE:

*(26)* *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.
I have suggested above how to interpret this condition (pending further discussion of AI in section 6.2). In short, the CDE says “Don’t move within a locally designated area unless it has an effect on PF.” Given (26ii), we could think of five ways to interpret this “drastic effect on the [PF] output.” These are sketched in (27):

(27) a. delete lower copy a’. [... O … Ø ☻ Ø]
b. # delete higher copy b’. [... Ø ☻ Ø … O ]
c. spell out lower copy c’. [... O … Ø ☻ X]
d. # spell out higher copy d’. [... Ø ☻ X … O ]
e. create new PF-matrix e’. [... O-X … Ø ]

Take an arbitrary object O, regardless of its phrase structure status. If it moves within a Prolific Domain, the combined operation of Copy and Merge, something must happen to either copy. Nunes (1995, 1999) argues that deletion takes place for PF-reasons: given that copies of O are non-distinct and linearization (in terms of precedence) is subject to the LCA, we cannot have essentially one element preceding and following itself. In standard applications of Move, the lower copy gets deleted (at PF). By deleting one of the two copies, they can be linearized, as we are only dealing with one copy.

This said, we still face five options. Deleting the lower copy of O is the common consequence of Move, illustrated in (27a’). Deleting the higher copy, as in (27b’), is supposedly not an option available to C_{HL}. Following Nunes, economy rules out this option:

---

25 In (27) I represent a change of PF-matrix by the arrow ‘➲’. Delete renders the PF-matrix empty (Ø). Copy Spell Out results in a different phonetic structure (X). My goal here and in subsequent chapters is to specify X, and the idea of Copy Spell Out. I thus employ ‘世代’ throughout to indicate spelling out of copies.

26 Chomsky (1995b: 227) defines distinctness in terms of different applications of Select, the step that puts a token of a lexical item into the LA. Linearization at PF requires all non-distinct copies to be unambiguously ordered in terms of precedence, conforming to the LCA (Nunes 1999: 224ff.).
the higher copy has a more complete set of checked features than the lower one (by virtue of having moved, i.e. checked a feature which is unchecked on the lower copy).

The CDE allows more alternatives: according to (26ii), all that is needed to conform to Exclusivity is a PF-effect of one of the two copies. Simply deleting the lower copy of a Domain-internal movement step is not a viable option—otherwise the illicit movements discussed in section 2.2 should be well-formed.\(^{27}\) Another effect on PF, apart from deleting, is changing the PF-matrix. This makes two non-distinct copies as distinct from each other as deleting one of the two. The cheapest way to change the PF-matrix of O, yielding a well-formed PF-expression of the languages, is replacing O with X, an item from the lexical inventory of the language that looks different, yet can act as if it were O.

This vague formulation will become clearer as we go along. Let us assume that a language can make available a (default) filler to spell out one copy with another PF-matrix. I will also address the nature of this element in the remainder of this work, after we see some cases where Copy Spell Out actually does some work.

If we can either delete or spell out the copy, we face the same two options for the latter as for the former: spelling out the lower copy, as in (27c’) or the higher one (i.e. (27d’)). The same reasoning for spelling out the lower copy applies as for deleting: the higher copy has one more checked feature than the lower copy. Lastly, a PF-effect can also be obtained by creating a different PF-matrix through Move (cf. (27e’)). As discussed above, this is what head movement does.

In short, the CDE rules out movement of a maximal phrase within a given Prolific Domain, unless it spells out its lower copy, and this is the hypothesis I will be working with, exemplifying with instances in the θ-domain (and the φ-domain in exceptional cir-

\(^{27}\) This is the definition of anti-locality. A more technical explanation will be given in section 6.2, in terms of occurrence in a contextually identical set of projections. Basically, what anti-locality does is give us a domain of evaluation in which the relevant projections can be thought of as not being hierarchically ordered. This prohibits us from forming a coherent dependency between two (or more) copies of O.
cumstances) in this section and the $\omega$- and $\phi$-domains in the next chapter. The empirical domain I consider in the remainder of this section is local reflexivization, which I propose to analyse as movement from $\theta$-to-$\theta$-position with subsequent Copy Spell Out.

### 3.4.2 A Derivational Approach to Reflexivization

Putting the Domain-driven framework and the proposal of the CDE aside for a moment, let us first consider in general terms what a derivational approach to reflexivization is, what it might look like, and why one would want to pursue this line.

#### 3.4.2.1 The rule of reflexivization

Early generative analyses aimed to relate identity of reference to, basically, non-distinctness of copies. In a pre-movement framework, there were transformational rules that applied at certain points of the “derivation” and in specific ordering relations, and replaced some lexical items with others. For example, passivization was achieved by applying a transformational rule to the underlying active structure, identical noun phrases in, say, control structures were expressed in terms of NP-deletion, *it*-replacement created certain structures assumed to be related, and so on (see, among many others, Chomsky 1965, 1968, Jacobs & Rosenbaum 1970, Rosenbaum 1967, Ross 1967, or the introduction by Jacobs & Rosenbaum 1968 and their annotated bibliography).

Likewise, there were transformational rules governing the behaviour of pronominal elements, what was later to become Binding Theory. The rule of pronominalization replaced some occurrence of identical noun phrase with a pronoun (as in (28a)), but not others (cf. (28b)); basically, this has developed into Conditions B and C.
There was also, however, a rule of reflexivization, first introduced by Lees & Klima (1963), which replaced identical noun phrases with a reflexive in some cases, namely when two identical noun phrases were in a “local” relationship:

(29)  
\begin{align*}
\text{a. } & \text{underlying representation: the boy hurt the boy} \\
\text{b. } & \text{rule of reflexivization: } \rightarrow \text{The boy hurt himself.}
\end{align*}

Given that the notions of Select and LA were introduced somewhat later than the 1960s, it is hard to relate the “derivational” histories of (28) or (29) with the modern-day equivalents. Presumably, both occurrences of the boy in (29) would be treated as distinct in a minimalist framework. However, had there been a transformational conception of movement in terms of Copy Theory, cases could be created where the two identical elements are actually non-distinct, an approach pursued subsequently. In the original transformational approach, the relevant notion was “identity” (but see fn. 29 below)

This shall suffice for the historic context. It is obvious that many of the rules of transformation did what is now done by movement—expressing traces or copies—or in terms of construal, for dependencies that apparently cannot be analyzed as the result of movement. Hornstein (2000) addresses the latter, conventional wisdom and builds on recent advances in analytical tools available in minimalism: by adopting a derivational modification of θ-roles qua features (see references in fn. 34 of chapter 2 and discussion above) and employing sidewards movement in controlled environments (Nunes 1995), he proposes a movement approach to many dependencies based on construal operations.
3.4.2.2 A derivational approach to local reflexives

While the more general extension of deriving previously construed dependencies looks promising, I cannot address the details of such analyses here (which cover relativization, control, and others; see also Boeckx 2000b, Pires, to appear). Hornstein’s (2000) work suggests, however, the possible elimination of Binding Theory, let the facts fall out from other, independently needed conditions and operations. In particular, a derivational approach to construed relations should have as few additional conditions built in as possible. Regarding binding relations, something has to be done from a minimalist perspective regardless, as the formerly crucial notion of “government” has been dispensed with (on grounds of virtual conceptual necessity; cf. Chomsky 1993, 1995a). Likewise, dispensing with the notion(s) of “binding domain” is desired, given that it involves a further step of defining domains which play a very little role outside binding.

Hornstein (2000) sets out to revive the original proposal by Lees & Klima (1963), suggesting a derivational analysis of reflexivization in minimalist terms. The idea behind it is this: given that antecedent and reflexive denote the same entity, and that there seem to be strict locality conditions (found especially in instances of A-movement), we might be tempted to derive the identity between the two elements syntactically, not solely semantically. As we have seen in section 2.2, the derivational analysis cannot work on standard assumptions about movement, which deletes lower copies—otherwise, John

---

28 See also Kayne (2000) for a similar endeavour, even if under a different set of assumptions.

29 A move questioned by Chomsky (1965: 145f.) who argues against the Less & Klima approach on the grounds of (the) recoverability (condition). Deleting a noun phrase (or other element) under identity, is one thing (viz. modern Trace Theory or Copy Theory); to replace it and introduce an additional element, be it a pronoun and/or the reflexive-marker –self, is quite another. I will address the issue of recoverability later. Let us stick for now to the guiding rationale of substantive economy, namely that if a dependency could be analysed on purely syntactic grounds, it should be superior to a purely “semantic” analysis which has to evoke additional assumptions regarding the construal of relation (reference, identity etc.).
*likes* should be a well-formed structure allowing for the interpretation *John likes himself*, which could be syntactically saturated (by copying the relevant features of *John* in theme position and re-merging them into the agent position, as discussed above ad nauseam).

Rather, Hornstein’s derivational analysis of local reflexivization can be sketched as follows (after Hornstein 2000: 159), where the pronominal element is an “elsewhere” formative and the reflexive, in particular, the residue of local A-movement:30

(30) a. \( \text{LA} = \{\text{John}, \text{like}, \text{self}, \text{T}, \text{Agr}, \nu\} \)

b. \( [\nu \text{ likes } [[\text{John}]-\text{self}]] \)

c. \( [\lambda P \text{ John likes} \nu [\lambda P \text{ likes } [[\text{John}]-\text{self}]]] \)

d. \( [\lambda P \text{ John likes} \nu [\lambda P \text{ likes } [[\text{him}]-\text{self}]]] \)

e. \( [TP \text{ John T } [\lambda \text{AgrP} (\text{himself}) \text{ Agr } [\lambda P \text{ John likes} \nu [\lambda P \text{ likes } \text{himself}]]]]] \)

The LA contains whatever it contains, say, the items listed in (30a), crucially not, though, the reflexive *himself*. The full nominal DP *John* is merged into theme position, but it comes with the reflexive-marker *self* attached to it—in adjoined position, as Hornstein suggests (i.e. (30b)). This element is a bound morpheme, and as such needs phonological material to attach to.31 *John* then raises to agent position (cf. (30c)) and spells out its trace (in (30d)), giving *self* a PF-matrix to be satisfied. And the derivation

---

30 The elsewhere approach to pronouns has been independently developed in a number of recent works. See Aoun & Benmamoun (1998) on clitic left dislocation, Aoun & Choueiri (2000) on epithet constructions, Aoun, Choueiri & Hornstein (in press) on resumption strategies, and others. The intuition behind such a view of pronouns relates to the Avoid Pronoun Principle of earlier approaches (Chomsky 1981, Aoun 1985). Note also that Hornstein (2000: 154), having proposed to dispense with Conditions B and C in earlier chapters, observes that the remaining binding condition “states an anti-locality requirement on nominal expressions” (his italics). This observation ties in neatly with the overall approach explored here, and I take it even further than Hornstein does, in terms of a more general condition on anti-locality (a notion developed here independently from Hornstein’s use).

31 The morpheme-analysis of *self* should be self-explanatory; that it should be a bound morpheme, rather than a free one, is obvious: otherwise we would expect *John likes self* to be grammatical (Hornstein 2000: 161).
continues in the usual manner, with himself raising to SpecAgrP (AgrOP) to check \(\phi\)
features and be Case-marked—either overtly or covertly, as indicated by parentheses.

Syntactically, this part of the derivation is well-formed. John checks its first \(\theta\)-
feature [TH], then its second \(\theta\)-feature [AG]. Hornstein builds on his and others’ earlier
arguments that the Theta Criterion should not exist and that theta properties constitute
roughly the same morpho-syntactic reflex as any other formal features.\(^{32}\) Spelling out the
copy is not only necessary for the reflexive morpheme to combine with, it is also driven
by Case needs. Crucially, self bears a Case feature, and this needs to be checked. By
copying the lower instance of John, all other features are transmitted to the higher copy.
The result of spelling out is that himself can then undergo Case-movement.

3.4.2.3 Reciprocals

Hornstein (2000) extends this analysis to local reciprocals, another type of ana-
phor that falls under Condition A.\(^{33}\) These can be captured roughly on a par with (30):

\[
(31) \quad \begin{align*}
\text{a. } & [v \text{ like } [[\text{the kids}]-\text{each other}]] \\
\text{b. } & [v^p \text{ the kids} \text{ like-v } [v_p \text{ like } [[\text{the kids}]-\text{each other}]]] \\
\text{c. } & [v^p \text{ the kids} \text{ like-v } [v_p \text{ like } [[\emptyset] \text{ each other}]]] \\
\text{d. } & [TP \text{ the kids } T [AgrP \text{ each other Agr } [v^p \text{ the kids} \text{ like-v } [v_p \text{ like } \text{ each other}]]]]
\end{align*}
\]

The derivation of (31), adapted from Hornstein (2000: 187) differs from the re-
flexive analysis in two main respects. First, each other is not treated as a morpheme, but

\(^{32}\) Among many others, relevant literature includes Bošković (1994), Lasnik (1995b,
1999), Bošković & Takahashi (1998), Hornstein (1998, 1999), Manzini & Roussou
(1999, 2000); see also Hornstein (2000: chapter 5) for a more accurate discussion of the
details of his approach to reflexivization.

\(^{33}\) In the bigger picture that Hornstein paints, some properties of local reciprocals (and
local reflexives) are parallel to the properties observed in obligatory control structures,
such as the requirement of local antecedents, or the restriction to sloppy readings under
rather as an adverb (and as such, part of the array, rather than introduced derivationally), specifying that the values of likers and likees must be different; this still allows for an adjunction-approach (where the reciprocal adjoins to the DP which then undergoes θ-movement), assuming that reciprocals, like reflexives, can check Case (cf. Hornstein 2000: 187; also fn. 34). Second, the θ-to-θ movement of the kids does not result in spelling out the lower copy, but in conventional deletion—due to the fact that the reciprocal is not treated as a bound morpheme, hence Case does not force spelling out.

In the next sub-section, I will give the movement analysis to local antecedent-anaphor relations a slightly different analysis, incorporating much of the above, yet in adapted form, fitting a stricter formal view of the CDE, rather than evoking Case or even lexical insertion of reciprocals.

First, though, I would like to discuss another recent approach to a derivational view of Condition A configurations. The central proposal of Lidz & Idsardi (1998) is to identify the copy left behind by A-movement, PRO as the controlled subject, and local anaphors as allomorphs whose surface forms are determined by specific conditions. Their motivation to collapse these phenomena is the same as Hornstein’s: the antecedent-anaphor relation in each case can be created from subject to object position, but not vice versa.

Following Dalrymple et al. (1998), Hornstein (2000: 187) proposes a propositional structure for the kids like each other as shown in (i), where ‘X’ is a variable over plural expressions:

(i) <the kids, λX(X like X), each other>

Hornstein suggests that the movement analysis of reflexives and reciprocals creates complex monadic predicates. Different θ-roles on the moving element and its copy ensure that identity in reference for reflexives; the reading of reciprocals, however, is not one of identity, but rather one where likers and likees are different, say from the set of kids one kid likes another, and the two kids are different.

Reflexives have a propositional structure roughly like (ii) for John likes himself. Treating reciprocals as an adverb of quantification is further compatible with Dalrymple et al.’s proposal of each other as a polyadic quantifier. See Hornstein (2000: 206, notes 105-108 for more).

(ii) <John, λx(x likes x)>
versa, into the subject of infinitival, but not finite clauses, obligatorily involves the closest possible antecedent etc. (cf. Williams 1980, Bouchard 1984, Lebeaux 1984-85).

The spirit of the proposal is the same: A-dependencies are the result of movement, where movement into θ-positions is allowed (as discussed above, adopting the movement approach to control of Hornstein 1998, 1999), deletion takes place for linearization purposes (Nunes 1995), and syntactic dependencies are created for syntactic reasons only.\(^{35}\)

The details are different, however, and incompatible with either Hornstein’s or my theory: choice of morphological form (or phonological instantiation) of the copy in the three different cases follows from a particular θ/Case ontology. This, in turn, requires (accusative) Case-assignment to take place in the base position. The object does not move to a higher position to be Case-marked, counter Checking Theory.

### 3.4.3 Exclusivity and Anaphors

My proposal to deal with reflexivization in a derivational manner should by now be obvious, given the revised CDE introduced earlier: local reflexives are not part of the LA but are formatives, inserted in the course of the derivation. As such, they are treated as the Copy Spell Out of the moving element—not for Case reasons, however, but to repair an otherwise illicit movement: to satisfy the CDE. Movement within a Prolific Domain is illicit, unless the lower copy is spelled out, and in comes the reflexive.

---

\(^{35}\) The latter follows from the Unique Lexical Access Principle which says that “each semantic entity corresponds to exactly one lexical item in the [LA]” (Lidz & Idsardi 1998: 119), placing a semantic restriction on the lexical array. Hornstein, and I, do not need to assume such a principle, following economy considerations: “keep the LA as bare as possible.” Semantic identity in the case of reflexivization is achieved through standard assumptions on movement—one lexical element, multiple copies.
3.4.3.1 Derivational reflexivization and Copy Spell Out in the θ-domain

The movement from TH to AG as sketched above is Domain-internal, it targets a position in the θ-domain from a position within the same θ-domain (cf. (30)). Such movement should be ruled out by Exclusivity—at least by one part of the CDE—, the very observation that led us to consider a formal tripartition in terms of Prolific Domains. The easy way out would be to altogether reject a derivational analysis of reflexives. However, as I will discuss in the next chapter, something very similar seems to be going on in certain cases of left dislocation. I will thus inspect the second part of the CDE and sharpen the notion of spelling out a copy along the lines sketched in section 3.4.1.

The intuition behind a derivational approach to local reflexives, of both Hornstein (2000) and Lidz & Idsardi (1998)—inasmuch as it connects to issues of substantive economy—is shared here, but I will approach the desired result from a slightly different angle. Assume that movement from θ- to φ-positions is triggered by agreement, rather than Case. This immediately puts a Case-driven approach to derivational reflexivization in doubt. If, however, the CDE can be motivated by substantial economy, we have an obvious candidate to exploit for this purpose.

This said, we can by and large adopt the derivation proposed by Hornstein from (30), recreated here with some relevant modifications I will address presently:

(32) a. LA = {John, like, self, T, Agr, v)  
    b. [v’ likes John]  
    c. [VP John likes-v [VP likes John]]  
    d. [VP John likes-v [VP likes himself]]  
    e. [TP John T [AgrP (himself) Agr [vP John likes-v [VP likes himself]]]]

The main difference is that Copy Spell Out is driven by the CDE, not by Case: if the lower copy of John were not spelled out, we would have a Domain-internal movement and the operation would violate the CDE—regardless of whether we delete the
lower copy (for linearization) or keep it (making the copies not linearizable on top). One further difference is already indicated in (32): while Hornstein takes John to be only spelled out as him, with self adjoined to the DP, I assume that self is not inserted from the lexicon, but the Copy Spell Out of John is the entire himself.

First, let us see whether we can apply this strategy to all the cases brought up in section 3.4.1 to demonstrate anti-locality in the $\theta$-domain. Recall from (10) in section 2.2, repeated here simplified as (33), which movement steps are predicted to be ruled out or, rather, be acceptable if and only if followed by Copy Spell Out (according to the CDE):\(^{36}\)

\[(33)\]
\[\begin{align*}
&a. \ [_{vP} \ AG \ V-v \ [_{vP} \ TH \ \nexists \ GO]] \\
&b. \ [_{vP} \ AG \ V-v \ [_{vP} \ TH \ \nexists \ GO]] \\
&c. \ [_{vP} \ AG \ V-v \ [_{vP} \ TH \ \nexists \ GO]]
\end{align*}\]

(33a) was demonstrated for simple transitives above (cf. John likes, i.e. (2a) of section 2.2.1). We can come up with a similar case for ditransitive constructions, where I assume, as already implied by (32), that the reflexive in its entirety is the spelled out copy of John and self is not inserted as an adjunct in the initial Merge. Exemplifying (33a) as a ditransitive, our notation allows for one compact representation:

\[(34) \ \ [_{vP} \ \text{John} \ \text{introduced}-v \ [_{vP} \ \text{John} \ \nexists \ \text{himself} \ \text{introduced} \ \text{to Mary}]]\]

Let us consider the other cases from section 2.2.1 next. On a par with (33a) qua (32) or (34), (33b) and (33c) can be derived under the derivational Copy Spell Out theory to local reflexivization I propose (cf. (4a) and (6a) from section 2.2.1, respectively):

\[^{36}\text{As a convention, the relevant copies are marked in boldface and all unpronounced ones in strikethrough.}\]
So far, this approach seems to cover the relevant data without resorting to Case-checking as the trigger for Copy Spell Out. We will see an empirical argument in favour of this over a Case-driven approach below; for now, a conceptual argument shall suffice: if the lower copy is spelled out in its θ-position because a bound morpheme adjoined to it needs to check Case at a later stage, in a different position, we would have to deal with an instance of globality, yet another unwanted case of ‘look ahead’. I take this to be less appealing than a Copy Spell Out triggered by an architectural condition. The alternative, which inserts only the relevant element and spells it out as the reflexive pronoun, will be addressed shortly, too.

As can be verified trivially, this analysis also accounts for the ungrammatical status of the following constructions:

     b. * Himself introduced John to Mary.
     c. * Himself introduced John Mary to.
     d. * Herself introduced John Mary to.

If it is the lower copy that must be spelled out (see the discussion around (27) above), these structures are not predicted, and this prediction is indeed borne out. This captures Condition A quite precisely.

---

37 Note that even though I present some examples as full sentences, the relevant derivational steps discussed here all concern the θ-domain. As such, himself in (36) must have been created in its θ-position.
This leads us to another case to consider. While the discussion above suggests strongly that lower copies get spelled out, we might have to say something about the following case under a movement approach to reflexivization:\textsuperscript{38}

(37) Pictures of himself worry John.

In these cases, we can follow Belletti & Rizzi (1988) and assume that the grammatical subject of psych-predicates is actually the internal argument. As such, the relevant steps of the derivation look as follows, where the picture-NP then moves to SpecTP:

(38) $[\text{vp John worry} - v [\text{vp worry [pictures of John} \Rightarrow \text{himself]]}]$

Thus, the reflexive in high position is derived by standard assumptions and we can maintain that the lower copy gets spelled out, as desired.

In this context, we can further assume, with Reinhart & Reuland (1993), that the reflexive in (39a) is actually a logophor, supported by the fact that it shares its distribution in these cases with a regular pronoun, as (39b) shows.\textsuperscript{39}

(39) a. John went to Mary\textsubscript{i}’s house. He saw pictures of herself\textsubscript{i} on the wall.
    b. John went to Mary\textsubscript{i}’s house. He saw pictures of her\textsubscript{i} on the wall.

\textsuperscript{38} Thanks to Anna Roussou for pointing out the potential relevance of (37). It is not clear, however, that this case really is relevant: after all, the reflexive can easily be replaced by a personal pronoun, which might tell us that we are not dealing with a local reflexive. Given that this analysis address local reflexivization only, other possibilities arise.

\textsuperscript{39} Castillo, Drury & Grohmann (1999: 95) give an additional argument for the logophoric nature of the reflexive in such cases in the discussion of eliminating intermediate SpecTP, based on the absence of complementary distribution of reflexive and pronoun in contexts such as (i):

(i) It seems to John\textsubscript{i} to appear to him\textsubscript{i}/himself\textsubscript{i} that the earth is flat.
Following up on the discussion around (36) above, two earlier assumptions are relatively important and potentially confirmed: TH is merged higher than GO, and GO is more complex. Consider:

(40)  
   a.  * John introduced to himself Mary.
   b.  * John introduced to Mary herself.

(41)  
   a.  * John introduced Mary himself.
   b.  * John introduced Mary herself.

If GO were higher than TH, the two sentences in (40) would be grammatical.\(^{40}\) Note again, that this concerns the base-generation of arguments: so far the proposal is that local reflexives are the Copy Spell Out of a \(\theta\)-domain-internal operation. At least in simple sentences such as the above, this seems to be true. Regarding (41), the violator seems to be the “preposition” (whether an underlying preposition or something else; see Baker 1997 for pointers). It can thus not be that “GO” moves, simply spelling out the \(\theta\)-marked goal argument. This supports the earlier hypothesis that non-distinctness is involved, which we take trivially to require categorial identity of the relevant copies.

In other words, (40) and (41) cannot be assigned the following derivations:

(42)  
   a.  # \([vP \text{ John introduced-}\nu [vP \text{ to } \text{ John } \check{\circ} \text{ himself introduced Mary}]]\]
   b.  # \([vP \text{ John introduced-}\nu [vP \text{ to } \text{ Mary } \check{\circ} \text{ herself introduced to Mary}]]\]

(43)  
   a.  # \([vP \text{ John introduced-}\nu [vP \text{ Mary introduced to John } \check{\circ} \text{ himself}]]\]
   b.  # \([vP \text{ John introduced-}\nu [vP \text{ Mary introduced to Mary } \check{\circ} \text{ herself}]]\]

\(^{40}\) Needless to say, the sentences should be judged with neutral intonation, that is stress on a potentially extraposed argument does not matter here. Again, the relevant movement step occurs within the \(\theta\)-domain. I address and dismiss the slight complication potentially posed by dative alternations in section 3.4.3.5.
In sum, by extending the original anti-locality hypothesis, the framework laid out so far can provide a derivational analysis of local reflexivization. Domain-internal movement can be legitimized through a repair strategy—namely exactly in case it results in spelling out its copy.\textsuperscript{41} Furthermore, the spelled out copy bears a different PF-matrix, that is to say, it looks different from the original; hence, we do not say *John likes John*.

It is clear that it is the PF-matrix that is at stake, not the LF-features, given that the two copies are otherwise identical ("non-distinct"). This might help us to reformulate the Case Filter (Chomsky & Lasnik 1977, Vergnaud 1977), namely that only overtly realized material can be Case-marked; I will return to this. We can now thus assume the CDE to really look as in (26) above, repeated here for convenience:

\begin{equation}
\text{(44) Condition on Domain Exclusivity (CDE)}
\end{equation}

An object O in a phrase marker must have an exclusive Address Identification AI per Prolific Domain $\Pi \Delta$, unless duplicity yields a drastic effect on the output.

i. A drastic effect on the output is a different realization of O at PF.

ii. An AI of O in a given $\Pi \Delta$ is an occurrence of O in that $\Pi \Delta$ at LF.

\textit{3.4.3.2 A brief comparison}

The PF-relevance of spelling out copies also follows from Hornstein’s account. In fact, it seems very hard to tease apart the two approaches on grounds other than theory-internal. Thus, the present analysis of local reflexivization conforms to two assumptions I made earlier: movement out of the $\theta$-domain is driven by $\phi$-features and movement within a Prolific Domain is ruled out, unless the lower copy gets spelled out. But these assumptions motivated a derivational look at reflexives as presented here in the first

\textsuperscript{41} Note that “repair strategies” are nothing particularly novel in the grammar. Arnold (1995) argues for do-support as such, on the grounds of Last Resort. What is new is the implementation, one not based on Last Resort but on architectural properties of phrase structure, something I come back to in chapter 6 also.
place—the argument would thus be circular. Not that there would be anything wrong with a theory-internal argument. If we could find a potential empirical shortcoming of one approach that the other can deal with would be something else entirely, however. Moreover, if the case in question would lend further support to choosing one particular assumption over another, the winner gets it all.

One phenomenon comes to mind that might just be such a case and pose a challenge to Hornstein’s analysis of reflexivization. Given that spelling out is forced purely by Case considerations, we would expect that if the relevant position would not necessitate subsequent Case checking, that element should be able to freely move to the higher position; all it would do is check an additional θ-feature. A relevant example could be (45).

(45) The fish ate.

The verb *eat* is a transitive predicate (cf. (46a)). The internal argument can, however, be left out, as in (45), which would be understood maximally as (46b), with the restriction that the eaten material be something that is generally considered to be edible and hence a possible candidate for food by the eater. Given that some fish can eat other fish, and some fish even eat other fish of the same type, (45) should be possible to be interpreted as either (46c) or (46d). The former interpretation is possible, but the latter is not: *the fish ate* cannot mean that it ate itself (or they ate themselves).

---

42 As with many points in this dissertation, Juan Uriagereka pointed this potential argument out to me. I am also grateful to Norbert Hornstein for discussing this issue, as well as everything else, with me in detail. Finally, Anna Roussou chipped in, leaving me to pick up the pieces.
(46)  a. The fish ate algae.
b. The fish ate something or other.
c. The fish ate the fish.
d. # The fish ate itself/themselves.

As Hornstein’s approach solely considers properties of formal feature checking (θ- and Case features in this case), where Copy Spell Out is driven purely by the need to create overt material that can later on check Case, this state of affairs is not expected. One way out of this dilemma—suggested also by Norbert Hornstein (p.c.)—is to assume that the two instances of *eat* in (45) vs. (46) are different. In one case, it is a regular transitive verb, and in the other it is a detransitivized verb, that is a verb which (optionally) can drop an argument. It is not clear that we want to go that route—if we do, fine. But if we want to maintain that there is only one lexical entry *eat*, we would have to find a different solution. I suggested in Grohmann (2000a: 64) an analysis along the lines sketched here, in which a highly restricted type of internal argument could optionally be deleted.

The advantage of detransitivizing such predicates is obvious: the structure conforms to Burzio’s Generalization. If a predicate could make available an internal argument and in a limited set of cases delete it before it affects the syntax, one might wonder what happens to the accusative case property of the verb, and why it could not discharge it (as it should under the generalization).

This is a good counter-argument, but we do not have to say that the unrealized argument in (45) is dropped or deleted before it affects the syntax.

Brody & Manzini (1988) identify the empty argument position as an implicit argument. Chomsky (1986b: 167) introduced this notion to account for legitimate cases of coindexation between a DP and a pronoun in, what appears to be, the latter’s Governing

---

43 Hornstein (2000: 157ff.) analyses inherent reflexives of the *dress/shave*-type in this manner also. I return to such cases shortly.
Category. Assuming an implicit (subject) argument inside the DP, with the properties of PRO, however, accounts for the following contrast immediately:

(47)  a.  * They\textsubscript{i} told [PRO\textsubscript{j} stories about them\textsubscript{i}].
     b.  They\textsubscript{i} heard [PRO\textsubscript{j} stories about them\textsubscript{i}].

The relevant interpretations are that in the ungrammatical case, the story-tellers are the topic of the stories, while in the good case, the tellers (PRO) are different from the hearers who, in turn, can be the topic of the stories. In (47b), PRO receives an arbitrary interpretation and hence is not the result of A-movement, not even under a movement approach to control endorsed here (cf. Hornstein 1999).

If we can have arbitrary subjects, we might also have arbitrary objects. If the implicit argument in (47b) is an arbitrary subject, the implicit argument in (45) should then be an arbitrary object. Dispensing with PRO altogether, Hornstein (1998, 1999, 2000) identifies arbitrary PRO as pro, a null form of a pronoun, forced by Last Resort (the “elsewhere” case). The grammatical formative pro can be inserted into the derivation whenever movement is not possible. We could then analyse (45) to have the following structure in the relevant part, the θ-domain (see Rizzi 1986 on object pro):

(48)  [\textsc{vp} the fish ate-v [\textsc{vp} ate pro]]

Detransitivization is not necessary, and neither is deletion of a specified class of arguments. Burzio’s Generalization does not pose a problem anymore either: pro is free to move to check φ-features with Agr and be Case-marked, if necessary.

To rescue the Hornstein analysis, we could dismiss (45) as a relevant case altogether. After all, one could argue, eating oneself does not constitute an appropriate edible object, not even for fish. In this case, we illustrate the point with kill. Arguably, fish are
capable of killing themselves, whether purposely or not does not play a role. *John killed himself (taking drugs)* can be the result of a planned overdose (suicide) or a careless mis-dosage (accident). But *kill* can also be used with an implicit argument, ranging over possible objects in the world that can be killed. Consider an alternative to (45):

(49)  The fish killed.

The data from above can be created analogously, receiving the same readings, that is (50c) and (50d) would still be expected to be available readings under a pure Case-driven movement analysis, but only (50c) is:

(50)  a. The fish killed the diver.
b. The fish killed something or other.
c. The fish killed the fish.
d. # The fish killed itself/themselves.

If this argument goes through, we can adopt the spirit of Hornstein’s approach, with the additional pay-off that this puzzle is taken care of. Thus, we still have a clean record for the proposal that spelling out of copies is not driven purely by formal features, to be checked at a later point in the derivation, but rather by a repair strategy to rule in an otherwise illegitimate move. I thus adopt the most important ingredient of Hornstein’s approach: the insight to derive certain construed relations. Coupled with the discussion so far, and the upcoming analysis of left dislocation structures, this serves as empirical support for the CDE as stated in (26) and a unified account for spelling out copies.

3.4.3.3 Reflexive ECM-subjects

So far, all instances of local reflexivization have been illustrated as the result of an anti-local movement step, followed by spelling out of the copy, in the same θ-domain.
This raises the question where and how a reflexive ECM-subject would be introduced, as we can barely motivate movement of the “antecedent” within the embedded θ-domain, the original merging position for the matrix subject.44

(51) a. John considers himself an idiot.
   b. Mary expects herself to win the race.

There are two ways of achieving the same result. One solution offers itself following the relevant discussion of ECM in Dutch (cf. section 3.2.3.1). If Zwart’s generalization also holds for English,45 it seems plausible that the reflexive moves to the matrix accusative position to be Case-marked. Alternatively, or in addition, there is good empirical evidence that even in English objects may move overtly, at least in some cases.46

I take the complete derivations of (51) to look as follows:

(52) a. [TP John T [? considers-ν-Agr-? [Agr John ⚫ himself considers-ν [νP John considers-ν [νP considers [TP T [SC John an idiot]]]]]]]
   b. [TP Mary T [? expects [Agr Mary ⚫ herself t_i [νP Mary t_i [νP t_i [TP to-T [νP Mary win]-ν [νP t_j the race]]]]]]]

44 Thanks to Richard Larson for making me aware of these cases, and Mark Baltin, Barbara Citko, and Marcel den Dikken for discussing my solution.

45 As noted on p. 93 above, Zwart’s generalization concerns restructuring verbs and accusative case-marking in the matrix clauses. The relevant part concerns raising predicates, possibly restructuring verbs in English, too. A restructuring verb can license Case to embedded subjects and objects in its functional domain if the verb in its complement domain has an external argument.

In other words, reflexive ECM-subjects are the result of Copy Spell Out in the matrix $\phi$-domain. As one would expect of a local reflexive by now, it starts out as its own antecedent, shown in (52a) for the subject of a small clause (SC) and in (52b) for the subject of an infinitival transitive predicate. Both have to raise into the matrix SpecAgrP overtly to check $\phi$-features and receive accusative case. However, they do so by way of the external argument position of the matrix clause, SpecvP, to satisfy the specific $\theta$-feature. As such, they should also move to matrix SpecTP to check the relevant $\phi$-features associated with that argument role, and be nominative case-marked. The move from AgrP to TP is Domain-internal and as such illicit—unless the lower copy gets spelled out: reflexive ECM-subjects thus spell out in the $\phi$-domain.

3.4.3.4 Inherent reflexives

Another case to consider here are the well-known reflexive constructions in which the anaphor occurs only optionally, exemplified by the following:

\[(53)\]
\[\begin{align*}
\text{a.} & \quad \text{John shaved (himself).} \\
\text{b.} & \quad \text{Mary dressed (herself).}
\end{align*}\]

At the face of it, one might be tempted to give the same explanation as for the eat/kill cases above, that shave/dress bear an implicit argument role which may or may not be expressed by a reflexive (which is Hornstein’s take). However, while (53) must

---

47 In (52a) I indicate all relevant head movement in detail, in (52b) I employ traces. One further assumption concerns overt verb raising to a position indicated by ‘?’ , but this assumption goes along with the hypothesis that in some cases, at least in ECM, the “object” moves overtly. I leave SpecTP of the embedded infinitival clause empty (see Castillo, Drury & Grohmann 1999 and other relevant references).

48 To my knowledge, no cases can be construed that would decide one way or another on the origin of the reflexive (i.e. matrix vs. embedded clause). Ceteris paribus, I take my suggestion to be legitimate.
have a reflexive reading in either case, (45) or (50) above cannot be construed with such an interpretation. This suggests that another analysis must be found.

Lasnik (1992) refers to these cases as inherent reflexives and analyses the optional null object position as PRO, which prompts Hornstein (2000: 167) to give it the same movement treatment as he gives to other control structures, as (54a). In case the reflexive is overtly expressed, he can apply the Case-driven spell out variant seen above, i.e. (54b):

\[
\begin{align*}
(54) & \\
n & \text{a. } [vP John \text{shaves-}v [vP \text{shaves John}]] \\
n & \text{b. } [vP John \text{shaves-}v [vP \text{shaves [John Ī him]self}]]
\end{align*}
\]

We cannot entertain this approach. If the empty position were reduced to PRO, there would be movement within a domain without Copy Spell Out and hence violate the CDE, assuming the general movement approach to control to be correct; (54a) would thus not be an option for a Domain-driven framework.

(54b) looks more promising, though. Capitalizing on Lasnik’s insight, I would like to suggest a variant of Hornstein’s analysis. If the verb in question takes an overt reflexive, nothing special needs to be said, we can simply apply the Copy Spell Out analysis, i.e. (54b) below. However, as an inherent reflexive, the verb has the option of not merging the antecedent in theme position. As movement cannot take place, barred by the CDE, our version of (54a) would look like (55).

\[
\begin{align*}
(55) & \\
n & \text{a. } [vP John \text{shaves-}v [vP \text{shaves pro}]] \\
n & \text{b. } [vP John \text{shaves-}v [vP \text{shaves John Ī himself}]]
\end{align*}
\]

Recall that if movement is barred, an elsewhere pronominal element is inserted to rescue the structure. In this case, the formative is the null form pro. The optionality of inherent reflexives can thus be recreated, following from other, motivated assumptions.
3.4.3.5 Double object constructions and other puzzles

Recall that (41b) above, at least with flat intonation, is pretty bad. However, as Barss & Lasnik (1986) observe, there is a version of it which sounds remarkably good:

(56)  a. * John introduced Mary herself. [= (41b)]
   b. John showed Mary herself (in the mirror).

A structure like (56b) forms a counterpart to the regular dative construction, known as the double object construction. Barss & Lasnik notice a number of asymmetries between the two, with respect to anaphor binding, quantifier binding, weak crossover, superiority and others. For obvious reasons, I concentrate on the first property.

The following paradigm characterizes one aspect of the dative alternation: some ditransitive verbs allow the dative construction only (with an indirect object PP), others allow the double object construction as well (indirect object PP preceding direct object):

(57)  a. John introduced the band to the audience.
   b. * John introduced the audience the band.

(58)  a. John showed the band to the audience.
   b. John showed the audience the band.

Let me briefly point out a natural incorporation into the current framework of the phenomenon which, at first glance, seems to pose some trouble for Exclusivity as told so far: if TH is merged higher than GO, how can TH ever show up as the reflexive whose “antecedent” is GO, under the assumption that Copy Spell Out targets the lower copy?

49 See also Larson (1988, 1990), Jackendoff (1990), den Dikken (1995a), Pesetsky (1995), Arad (1998) for some recent discussion. It seems to me, though, that the following discussion is an “easy way out.” I will hint at an alternative later.
We could follow Barss & Lasnik, or Arad, and base-generate TH (direct object DP) and GO (indirect object DP) in inverse positions in the two constructions. Thus, in the dative construction, the TH position could be SpecVP and GO CompVP (as assumed here throughout), but in the double object construction, GO sits in SpecVP and TH in CompVP. Arad (1998: 89f.) proposes these two structures, and supports them with data that suggest that the readings available in dative constructions are only a subset of those found in double object constructions (see especially Green 1974, Oehrle 1976).

The two pairs above would then receive the following derivations (at the relevant level, i.e. within the θ-domain):

(59)  a. \((vP \text{ John introduced-v } [\text{ VP the band introduced to the audience}])\)
    b. \# \((vP \text{ John introduced-v } [\text{ VP the audience introduced the band}])\)

(60)  a. \((vP \text{ John showed-v } [\text{ VP the band showed to the audience}])\)
    b. \((vP \text{ John showed-v } [\text{ VP the audience showed the band}])\)

The difference between the two is whether a verb does or does not allow for two different base-generated orders: (59) only allows the dative construction, while (60) makes both available.

This gives us a straightforward account for (56b) above, adjunct aside (while (56a) is still ruled out, as introduce only makes TH > GO available):

(61)  \((vP \text{ John showed-v } [\text{ VP Maria showed Maria to herself}])\)

In the appendix I suggest an alternative account, if only tentative, which treats double object constructions as “special,” in a way on a par with ECM-constructions.

Relevant for our discussion of ditransitive constructions is the following note. Consider another illicit example from above and a grammatical version of it:
The ungrammatical sentence becomes good if the direct object is focused (indicated by capitals). This still allows for a derivational analysis of the reflexives, with the ordering TH > GO, as either one, or possibly both, of the internal arguments have moved from their θ-positions. As a focused constituent, Mary might receive stress in its θ-position, forcing lower material to move over it (in accordance with the Nuclear Stress Rule of Cinque 1993; (63b) or some version of it), or the focused direct object has moved or extraposed (something like (63b’), possibly, but not necessarily, involving rightward movement). In either case, we can assume the representation given for the underlying structures above to be the source of reflexivization here as well: we can thus employ (35b) for (62b), as shown in (63a)—while (62a) cannot be derived, a desired result.50

In sum, we could account for the variation found in ditransitive constructions if we assume two different underlying structures for dative constructions and double object constructions. Following a prevalent way in the literature (but not the only one), I suggested here to treat double object constructions differently in allowing for GO > TH as the underlying order, on top of TH > GO only for the dative construction. This allows us to treat reflexives as the result of Copy Spell Out inside the θ-domain throughout.

50 Note that native speakers judge the alternative above, (40b), bad no matter what. This might not be the worst state of affairs. If some version of it turns out to be acceptable, the reflexive could be considered a logophor, which is not unreasonable at all, given what we have said so far.

(i) *John introduced to Mary herself.  (cf. John introduced Mary to herself)
3.4.3.6 Reflexive or reciprocal: A preliminary note on “default fillers”

One natural question at this point could be: how does C_HL know which item to replace a CDE-violating copy with? In other words, given that Copy Spell Out is determined straight away by the computational system (very locally), rather than by some future checking requirement, what makes one copy to be spelled out as a reflexive, and another as a reciprocal?\(^5\) Concentrating on the \(\theta\)-domain, these are the (apparently, only) options. Naturally, there is a semantic distinction: reflexives replace a lower copy whose referent is identical with the referent of the higher copy, while reciprocals pick out a particular set of the referents (see section 3.4.2.3 and in particular fn. 34 above).

For illustration, take two relatively straightforward constructions:

\[(64) \quad \begin{align*}
\text{a.} & \quad \text{John likes himself.} \\
\text{b.} & \quad \text{John and Bill like each other.} \\
\text{c.} & \quad \text{John and Bill like themselves.}
\end{align*}\]

I want to suggest that it is not unreasonable to suppose that this “semantic” distinction is encoded on the originally merged LI. In order for John to be merged into TH and subsequently move into AG position (followed by Copy Spell Out in TH), it needs two sets of \(\theta\)- and \(\phi\)-features. If this is all it has, the Copy Spell Out will be one expressing full identity. Noteworthy, though, is the fact that singular referents cannot receive a reciprocal meaning—that presupposes a plural referent sets. Following Schein’s (1993) proposal that a plural noun phrase basically expresses the coordination of all possible...

\(^5\) Another relevant question concerning Copy Spell Out as the strategy to introduce local reflexives comes from languages that do not employ reflexive pronouns, but where the verb comes with a reflexive marker (such as Finnish), or those that have both (such as Kannada). It is feasible to think of the reflexive marker as the spelled out element, possibly along the lines of the analysis presented for clitic left dislocation in section 4.4. I leave the issue open, but gratefully acknowledge discussions on this and related issues with Jonathan Bobaljik, Susi Wurmbrand, Malte Zimmermann and especially Jeff Lidz.
events involving the relevant argument structure the rough LF of (64b) look as (65a),
while that of (64c) would be something like (65b):

(65) a. $\exists e_1[\text{Likes}(e_1, \text{John}_i, \text{Bill}_j)] \land \exists e_2[\text{Likes}(e_2, \text{Bill}_i, \text{John}_i)]$
b. $\exists e_1[\text{Likes}(e_1, \text{John}_i, \text{John}_i)] \land \exists e_2[\text{Likes}(e_2, \text{Bill}_i, \text{Bill}_i)]$

Thus, merging a noun phrase denoting a multiple member set, the internal struc-
ture to $[\text{DP John and Bill}]$ presumably has these relations encoded.\footnote{These are just crude representations; the indices are just used as a convenient notation.} In that case, if the relevant information is one of a conjunction of self-liking events, the filler is a reflexive, and if it is a conjunction of transitives liking events, it is a reciprocal. There are no alter-
natives; a semi-reciprocal expression EACH OTHER that picks out only a subset of the possible liking relations, as (66) illustrates, does not exist.\footnote{This does not mean that Peter, Paul and Mary like each other could not be interpreted (exhaustively) as in (66b). Langendoen (1978) calls this the “weak” reading of reciprocals. See also Lasnik & Fiengo (1973) on each other vs. each-the-other (or our EACH OTHER). In a recent paper, Quintana (1998) presents a syntactic approach to the two readings and discusses the syntax-semantics properties of “weak” reciprocals.}

(66) a. * Peter Paul and Mary like EACH OTHER.
b. # Peter likes Paul, Paul likes Mary, and Mary likes Peter.

While a reading like (66b) may be achieved in some circumstances (see fn. 53), it is not expressed with a specific lexical item, further supporting the rather restrictive emp-
loyment of possible filler items for Copy Spell Out, namely purely grammatical forma-
tives which are subject to very little variation. We will see more instances of Copy Spell Outs in the next chapter, allowing the introduction of more such filler items, depending on the copies or involved—or rather, the Prolific Domain they occur in.
Moreover, the facts suggest that the choice of the filler, restricted as it is, depends on information internal to the noun phrases. If we tied that information to \( \phi \)-features, we would yield a further possible argument in favour of \( \phi \)- rather than Case-driven movement. Insertion of a formative in the relevant circumstance (saving a CDE violation) must be licensed by \( C_{HL} \) and a \( \phi \)-projection (Agr) seems a reasonable place to do so.\(^5^4\)

This analysis immediately accounts for the following facts. The interpretation of (67a) and (67b) is taken care of after the preceding discussion, even under a derivational approach. Curiously, though, (67c) can only have the reading of (67b), not (67a)—neither under the “weak” nor the “strong” reciprocal reading. Moreover, the pronoun in (67d) cannot refer to any of the referents given in the “antecedent.”

(67)  
\begin{align*}
\text{a.} & \quad \text{Peter, Paul and Mary shaved each other.} \\
\text{b.} & \quad \text{Peter, Paul and Mary shaved themselves.} \\
\text{c.} & \quad \text{Peter, Paul and Mary shaved.} \\
\text{d.} & \quad \text{Peter, Paul and Mary shaved them.}
\end{align*}

If pro is an “elsewhere” formative, inserted when no movement can take place, it better be unambiguous (here, reflexive) or have open reference (i.e. an arbitrary reading). As there is no null alternative, it would be impossible to construe an interpretation otherwise. There is an overt alternative, in fact, there are two, and they appear in roughly the same context: movement within a Prolific Domain, followed by Copy Spell Out, and this alternative can differentiate between a reflexive and reciprocal reading.

---

\(^5^4\)This view might be taken to imply that only nominal elements can spell out as anaphors. As the next chapter will show, there are movement-derived instances of left dislocation: some languages employ a clitic, others a demonstrative pronoun to resume the left-dislocated element. Interestingly, and a puzzle to a unified approach for the longest time, the clitic variant can only resume a noun phrase, whereas the demonstrative can serve as a legitimate Copy Spell Out for other material of different categorial status (PP, VP etc.). As it turns out, the Copy Spell Out qua clitic occurs in the \( \phi \)-position (AgrP), while the demonstrative is spelled out in an \( \omega \)-position (TopP).
Likewise, if pronouns are also elsewhere formatives (“inserted when no movement can take place”), but overt, they should not allow for the same interpretation that an alternative form could do. If we take insertion of a pronoun to be equally costly as spelling out a copy, we would otherwise face an optionality, not desired under minimalist assumptions. Applying Copy Spell Out is driven by interface conditions, otherwise the CDE would be violated (movement). Picking a particular filler element (i.e. reflexive or reciprocal) is subject to the internal information of the moving noun phrase. A pronoun can only be inserted when movement cannot take place, but in these instances, it could take place—if followed by spelling out the CDE-violating copy. Hence, a pronoun inserted into a Prolific Domain where movement could have, but has not, taken place requires independent reference, such as deictic or otherwise discourse-bound.

3.5 Conclusion

Prolific Domains are subject to the Condition on Exclusivity, barring movement from one position to another within a single domain. This alone presupposes that multiple specifiers do not exist, at least not in the cases they have been applied to: accusative cannot be checked in an outer specifier of vP because it would violate the CDE (θ-domain), topics cannot move and adjoin to TP for the same reason (φ-domain), and so on. Moreover, a one-feature-per-projection checking restriction, not unreasonable on independent grounds, cannot carry two specifiers, given that specifiers by definition enter into a proper checking relation with their head.

Note that pro is a lexical option available to some verbs, as discussed above, and as such does not enter considerations of economy or optionality.
I then considered empirical arguments against multiple specifiers, focusing on the paradigmatic case: multi-Spec\(v\)P. Reviewing data from Dutch, we found little substance in the original argument against Agr. This argument was that Agr does not carry interpretable features and should thus be deleted at LF. In contrast, \(v\) has interpretive content and is thus a more viable candidate to check accusative Case (and \(\phi\)-features, by extension). Regardless of the arguments in favour of Agr mentioned in the previous chapter, a closer look at accusative-marking showed that there are instances in which the Case-marking head does not have any other function than doing just that. As such, \(v\) would all of a sudden be on a conceptual par with Agr. We conclude that all other things being equal, Agr is a viable candidate to mark Case, check \(\phi\)-features and be a part of \(C_{HL}\).

A closer inspection of the arguments against \(X'\)-invisibility and structural relations in the phrase marker allowed us to exclude multiple specifiers on more principled grounds. These grounds were phrased in terms of natural Relations, the product of the two primitive relations, Sister and Immediately Contain, that come for “free” with the operation \(\text{Merge}\), and the first-order composition of these, yielding in addition Contain, Identity and Extended Sister. I laid out a path that integrates these relations into defining structural relations in the phrase marker, relevant for feature checking and related issues. This discussion strongly suggested that elements adjoined to XP cannot enter into a checking relation with the head, thus barring movement of adjuncts to another adjoined position, and it ruled out the possibility of more than one specifier for the same reason, holding fast to the notion that movement must be licensed (in terms of “checking”).

We also extended the discussion of anti-locality. By introducing a repair strategy to the computational system \(C_{HL}\), which spells out the lower copy on principled grounds, an otherwise illicit movement can be legitimized (CDE). By looking more closely at anti-locality within the \(\theta\)-domain we found that a derivational approach to local anaphors is possible. The empirical motivation was found to be different, however, and the discussion above can be taken as a first confirmation of the CDE.
The discussion also ties in with assumptions made earlier, such as \( \phi \)-feature checking as the trigger for argument movement or ECM-constructions supporting an Agr-approach. I extended the discussion of local reflexives to reciprocals, anaphoricity in ECM-constructions, inherent reflexives, and predicates with implicit arguments. I also discussed a possible integration of double object constructions, which turned out to be less insightful than one might hope. My hunch is that more is at stake and rather than base-generating two different structures, these constructions could involve a similarly “exceptional” Copy Spell Out as ECM-constructions, within the \( \phi \)-domain. I leave this alternative open for now.

We can now assume that the CDE does indeed quite a bit of work, especially compared with alternative approaches that have recently been proposed. I thus take the direction embarked on here a valid path to consider more seriously. Our desiderata for relations within X'-structure have also received a sound foundation, and we can work with the following assumptions (from (25)):

\[
\begin{align*}
\text{(68) } & \text{Specifiers and adjuncts are formally different objects in the phrase marker} \\
& \text{i. adjunction to XP must be the result of base-generation (Direct Merge)} \\
& \text{ii. specifiers are base-generated or result from Move (Copy plus Merge)} \\
& \text{iii. specifiers enter a checking relation with a head and must be unique} \\
& \text{iv. adjunction cannot check features with a head and need not be unique}
\end{align*}
\]

I ended with speculations on the nature of a “default filler,” the PF-matrix that replaces the lower copy of an anti-local dependency. The next chapter deals with Copy Spell Out in other Prolific Domains, where I propose an analysis for certain types of left dislocation pretty much along the lines suggested here for local reflexivization. This discussion will produce an extended set of “default fillers,” not a complete typology, but a collection coherent with the overall framework I propose.
CHAPTER 4

LEFT DISLOCATION CONSTRUCTIONS:
MOVEMENT VS. CONSTRUAL AND COPY SPELL OUT

Left dislocation constructions are interesting, even from a (purely) syntactic point of view, as they involve two phonetically distinct elements for the same referent. In the standard case, they involve a left-peripheral constituent and a pronominal element of sorts lower down in the structure, picking up its reference (and much that comes with it). Intuitively, we might want to tie this “identity” to something deeper than dealing with two lexical items that can be linked one way or another. Exclusivity seems to be a property of the grammar that might be able to cash out this intuition. I am going to play with it in this chapter, building on and extending much of what I presented in the past two chapters, thus further sharpening the concepts Prolific Domains and Copy Spell Out; also, I finally justify the long song and dance about X’-structure, specifiers and adjuncts.

4.1 Introduction

The core proposal in this chapter is that two types of left dislocation must be understood in terms of movement of the left-dislocated element, and as the relevant movement takes place within the same Prolific Domain, I propose the repair strategy Copy
Spell Out to analyse the relation between the moved element and the resumptive pronoun.

In section 4.2 I introduce tree types of left dislocation: hanging topic left dislocation (HTLD), contrastive left dislocation (CLD) and clitic left dislocation (CLLD). I first present a collection of properties, then turn to English and contrast left dislocation with an apparently very similar construction, topicalization. I then address differences and similarities among all types of left dislocation and topicalization, and discuss a battery of arguments that suggest that the topicalized, CLDed and CLLDed XP are all derived by movement. In section 4.3 I offer an analysis of CLD and HTLD in German, understanding the left-dislocated constituent to undergo all relevant movement in the former, but not the latter, followed by Copy Spell Out. I then present reasons which suggest that the two are more different than we thought, in quite interesting ways, supporting the (particulars of the) distinction between specifiers and adjuncts I proposed earlier. In section 4.4 I turn to CLLD, mainly with data from Modern Greek (henceforth, Greek). We can apply an analysis of Copy Spell Out as well, but this time it targets the $\phi$-domain. In section 4.5 we discuss the two derivational operations, their differences and similarities for CLD and CLLD. Section 4.6 concludes this chapter and prepares us for a further inspection of some left-peripheral phenomena.

4.2 Some Properties of Left Dislocation Constructions

Left dislocation constructions come in three types, each one with specific syntactic, semantic, pragmatic and even phonological properties, often different from the other. For the most part I will be concerned with the former. But the three types share at least one descriptive property: they all involve a left-dislocated phrase and a pronominal element resuming its reference somewhere lower in the structure.
4.2.1 Three Types of Left Dislocation

I first introduce the three types of left dislocation: hanging topic left dislocation (HTLD), contrastive left dislocation (CLD) and clitic left dislocation (CLLD). These are illustrated in (1) with the languages on whose syntax I shall focus on in this chapter:

(1)  a. *This man, I don’t know* him. \(\text{\it (English HTLD)}\)

\(\text{\it b. Diesen Mann, den \text{\it kenne ich nicht.}}\)
\(\text{\it \text{this.ACC} \text{\it man that-one.ACC know I not}}\)
\(\text{\it ‘This man, I don’t know [him].’} \(\text{\it \text{\it \\it (German CLD)}}\)

\(\text{\it c. Afton ton andra, dhen ton ksero.}}\)
\(\text{\it \text{\it this.ACC} \text{\it the.ACC man.ACC not} \text{\it \text{\it m.ACC know.1SG}}\}
\(\text{\it ‘This man, I don’t know [‘em].’} \(\text{\it \text{\it \text{\it (Greek CLLD)}}\)

The only LD type found in English is shown in (1a), HTLD. It involves an LDed XP which is coreferent with a regular pronominal element serving as a resumptive pronoun (RP). The LDed XP fills a topic-like position, which we will specify in due time; the RP sits in the same position where any corresponding argument could sit, be it the thematic base position or an Agr-position (depending on whether one assumes general-

---

1 A terminological note. As I will be using the terms extensively in this chapter, I will shorten as follows: left dislocation (and all its derivatives) corresponds to LD (LDed, LDing etc.). I will refer to the LDed constituent as LDed XP or just XP where appropriate, and the resumptive pronoun as RP, where it need not be more finely differentiated. Where it does, I introduce the relevant abbreviations, namely CLitic, d(emonstrative)-pronoun, and p(ersonal)-pronoun.

And a historical note. The term “left dislocation” was introduced in Ross (1967), attributed to Maurice Gross; “hanging topic (left dislocation)” is apparently due to Alexander Grosu (see Cinque 1977), “contrastive (left) dislocation” was first used in Thráinsson (1979), and “clitic left dislocation” was coined by Vat (1981). I believe. See van Riemsdijk (1997) for a brief overview of the history of LD constructions and some related issues, and the collection of papers in Anagnostopoulou, van Riemsdijk & Zwarts (1997) for a historico-contemporary perspective.

ized overt object movement for English). HTLD can also be found in German and Greek, but these two languages make one additional construction available. German has a variant which employs the corresponding demonstrative form as the RP. Leaving further details aside, this RP sits in topic position, immediately followed by the verb, and the LDed constituent is somewhere higher; moreover, the two obligatorily match in Case. The Greek variant uses a clitic as RP (hence CLLD). Like German, the LDed XP and the RP match in Case. Unlike German, the RP sits in a much lower position; again, pending further discussion, it seems to occupy an agreement-related slot, an assumption which is supported by the fact that it follows negation.³

It has often been noted that the LDed XP has an “extra-sentential” character (cf. Ross 1967, Emonds 1970 and others). For one, given appropriate identification of the pronoun in the discourse, it can be left out without changing the status of the sentence:

\[(2)\]
\[\begin{align*}
\text{a.} & \quad \text{I don’t know } \text{him.} \\
\text{b.} & \quad \text{Den} \quad \text{kenne ich nicht.} \\
& \quad \text{RP.} \text{ACC know I not} \\
& \quad \text{‘That one, I don’t know.’} \\
\text{c.} & \quad \text{Dhen} \quad \text{ton ksero.} \\
& \quad \text{not CL.} \text{ACC know.1SG} \\
& \quad \text{‘I don’t know’im.’}
\end{align*}\]

(German)

(Greek)

The word order of the LD-less remainder stays exactly the same. The Greek and English are roughly equivalent, here picking out a particular individual from the discourse, only differing in the pronominal which is a clitic CL in Greek (and can be rendered as a phonological clitic in English, as the translation indicates). The same goes for

³To bias the following discussion, I stick to the pattern in (1) and translate CLD and CLLD systematically as topicalized structures, for reasons that will become obvious soon. I also italicize the LDed XP and its RP (boldfaced, in addition) throughout. Unlike (1), I shall not gloss the RPs with their approximate English counterpart in the following, but simply indicate d-pronouns as RP.CASE and clitics as CL.CASE.
German, but as the translation indicates, the demonstrative pronoun \( d \)-pronoun is topicalized. One difference between the CL and the \( d \)-pronoun is that only the latter can be stressed, as would be expected from a full tonic pronoun.

A second indication that the LDed XP—regardless of whether HTLDed, CLDed or CLLDed—occupies a kind of irregular position with respect to the rest of the sentence comes from interaction with other phenomena, such as topicalization or Wh-questions. I will treat this property in the next sub-section, as it shall play an important role for our analysis, then turn to obvious differences among the constructions under discussion.

4.2.2 **Topicalization vs. (Hanging Topic) Left Dislocation: Some Basic Properties**

Chomsky (1977) observes similarities in the syntax of Wh-movement and topicalization. In this framework, both are derived by what we would now call A’-movement. In this context, (5) ties in with the above noted intuitive relation between topicalization and LD—which should now be differentiated, so I refer to English LD as HTLD.

(3) a. Which book should we give to John?  
   b. Who did Mary see?

(4) a. This book, we should give to John.  
   b. John, Mary saw.

(5) a. *This book*, we should give *it* to John.  
   b. *John*, Mary saw *him*.

The representation capturing (3) and (4), translated into current terminology, could in principle all look as shown in abstraction in (6), where XP refers to either the Wh- or the topicalized phrase, disregarding intermediate Case-marking for now:
The pattern of (6) applied to Wh-movement and topicalization, where XP would be moved from its base-generated position (possibly from inside VP) to the sentence-initial position SpecCP (plausibly via an agreement-related position in between, as indicated by the parentheses) is, all things being equal, not unreasonable in current frameworks (at least since Chomsky 1986a). As Chomsky argues, it is less plausible to derive HTLD in the same manner. He notes the relation between topicalization and HTLD, which is relevant here. Chomsky’s (1977: 94) illustrates with the following paradigm:

(7) a. * This book, to whom should we give it?
b. * John, who do you think saw it?

(8) a. * This book, to whom should we give it?
b. (As for) John, who do you think saw him?

According to Chomsky, one is derived by movement, the other is not. The examples in (7) are instances of topicalization. Here, movement of the topicalized element would result in a doubly filled COMP and is hence ruled out, or in a violation of the Wh-island constraint, which is equally ungrammatical. The analogous LD cases in (8) with an HT are well-formed.

---

4 I indicate the original position with a trace, for expository convenience, and disregard intermediate steps. More importantly, I refrain from a more articulated punctuation system and simply separate “topics” of all sorts by a comma in English, and all LDed XPs from the rest of the sentence in German.

5 To be precise, both constructions contain a base-generated topic. But within the framework of the time, topicalization, not LD though, also involves a Wh-operator which moves to COMP where it is later deleted. For reasons of simplicity, I call this the “movement approach” to set it apart from HTLD.

6 The as for is added for convenience only; see Villalba (2000) for arguments against a HTLD analysis (and Cinque 1977 on similarities between the two constructions).
One result of Chomsky’s study is that HTLD cannot involve movement as a number of principles assumed to be diagnostics for (non-)movement can be violated (Subjacency, the Complex Noun Phrase Constraint, Specified Subject Condition, and other islands; see Ross 1967 on the relevance of islands in syntactic computation).

4.2.3 More on Topicalization: Derivational History

Now that we have mentioned topicalization, let me lay out my assumptions regarding its structure and derivational history. In the languages to be discussed here, argument topics appear very high in the clausal structure, somewhere in the left periphery (the ω-domain). They are Case-marked and thematically specified; moreover, no other argument of the same type can appear elsewhere in the sentence. In other words, there is initial reason to assume that an argument topic enters the derivation in the θ-position, moves on to the relevant φ-position, and finally raises into the ω-domain.

In some languages, the landing site attracts the verb, in others it does not. I assume topicalization to be a movement operation like any other. That means it needs to be licensed, currently expressed in terms of feature checking. For lack of more interesting terminology, let us take [Top] to be the relevant feature. Given standard assumptions on phrase structure—and certainly under the framework I presented in section 3.2—, the landing site of a topic must be a specifier. I take the relevant position to be TopP, one projection in a finer grained COMP-complex (cf. (16) from section 2.3.3 and the accompanying discussion).

---

7 I am not concerned with the semantics or pragmatics of topics, and a finer distinction will be mostly irrelevant for the following. By ‘topic’ I simply refer to a syntactically displaced element at the beginning of a sentence with particular properties (e.g. it represents old or given information, it cannot be used as an answer to an information question, it tends to come with “comma intonation” etc.).
This analysis of topicalization is by no means the only one available. Among
other possibilities, topics have been analysed as occupying SpecCP (Chomsky 1977) or
adjoining to TP (Baltin 1982), and others (e.g. Authier 1992, Iatridou & Kroch 1992)—
all of which either base-generated or moved. In case of base-generation, the favourite
mode of θ- and φ-licensing was in terms of (empty) operator movement, as mentioned in
fn. 5. Given our desire to eliminate grammatical postulates that are not virtually concep-
tually necessary, and given that in many cases null operators can be dispensed with in
2000b, 2000c), we should be suspicious again of such an approach. Moreover, the
framework outlined in the previous chapters allows for a movement analysis of topics in
a straightforward manner, even forces it, so I assume that this is the way to go.

(9) shows the derivational histories I assume for topicalization in English, German
and Greek (where pronounced material is boldfaced for readability):8

(9) a. \([\text{Topp this man Top [TP I do-n’t-Neg-T [NegP a’-Neg [AgrP this man
like-V-Agr [vP I like-V [vP like-V this man]]]]]]\)
b. \([\text{Topp Diesen Mann mag-V-v-Neg-Agr-T-Top [TP ich mag-V-v-Neg-Agr-T
[Agr] diesen Mann mag-V-v-Neg-Agr [NegP nicht mag-V-v-Neg
[vP ich mag-V [vP mag-V diesen Mann]]]]]]\)
c. \([\text{Topp afton ton andra Top [NegP dhen Neg [TP ksero-V-v-Agr-T [AgrP afton
ton andra ksero-V-v-Agr [vP pro ksero-V [vP ksero-V afton ton andra]]]]]]\)

---

8 The structures refer to (1) without resumption. German and Greek have overt Case-
marking, thus supporting the intermediate touch down (see also chapter 6). English does
not raise the verb into the Top-head, while German does, at least in matrix Verb Second
(V2) clauses. For Greek, I assume the topic position to precede the negation projection of
which dhen ‘not’ is a specifier, both in the C-layer (see Roussou, in press and references).
Negation is arguably relatively low in German (Abraham 1995), as opposed to English
and, particularly, Greek; Abraham also argues that it is a maximal projection, sitting in
SpecNegP. All other things being equal, let us take (9) to be roughly correct, if complex.
Moreover, if movement really is the way to go for topicalization, we might wonder whether it can also be employed to understand other, apparently similar phenomena for which operators have been proposed. Naturally, I have in mind CLD and CLLD, and I will suggest exactly such an analysis, as opposed to HTLD. Next we will see how HTLD differs from topicalization in involving a base-generated XP. Then we will subject CLD and CLLD to similar comparison with respect to each other, HTLD and topicalization. I will then propose a first stab at a movement analysis for CLD. After considering more data, I will revise this analysis slightly and then apply it to CLLD.

4.2.4 Topicalization vs. HTLD, Again: Movement vs. Construal

Returning to the three types of LD from (1), a question arises at this point, one which I will explore in considerable detail: given that topicalization is derived by movement and given that LD seems to be “topic-like” (in a sense yet to be specified), is LD derived by movement also? The long and short answer to this question we will arrive at is that some types of LD are derived by movement and others are not. In the course of the discussion we will see which ones (CLD and CLLD), which element is the mover (XP), how the identity relation between XP and RP is established (Copy Spell Out), and why LD constructions are different from topicalization, yet somehow related (Exclusivity).

Before we can address these questions, we need to do some legwork, though. Once that is done, we can turn to an adequate analysis of the phenomena. While I will arrive at basically the same analysis as proposed for reflexivization, there is currently no treatment of satisfactory LD constructions on the market to draw from. The question of movement vs. base-generation (hence, construal), however, has been pertinent throughout the history of generative discussions of LD as well as throughout the changes of direction and technology in frameworks.
Lasnik & Saito (1992) approach Chomsky’s problem mentioned above from a different angle. Recall that Chomsky (1977) gave basically a derivational treatment (“movement analysis”) to Wh-movement and topicalization, but one in terms of construal for HTLD. They point out correctly that the framework following Chomsky & Lasnik (1977) and Chomsky (1981), in which the that-trace effect is accounted for by a constraint on traces, rules out the examples shown in (10a,b), but not (10c), by the ECP:

(10)  

a.  * John, I think that ti won the race.  
b.  * Who do you think that ti won the race.  
c.  John, I think that he won the race.  

(Lasnik & Saito 1992: 76)

But then they reconsider Chomsky’s (1977) analysis on the basis of a prediction which is not completely borne out empirically, namely that “under certain circumstances where [HT]LD is available, topicalization should be unavailable, for example, where Subjacency or the ECP would be violated. On the other hand, wherever topicalization is possible, [HT]LD should always be possible” (p. 76). Especially the latter prediction does not seem to be accurate, as (11) and (12) show (from Lasnik & Saito 1992: 76f.), though the ungrammaticality of (12a) may not be as severe as indicated by the star (p. 193, fn. 7):

(11)  

a.  I believe that this book, you should read ti.  
b.  … that this solution, I proposed ti last year is widely known.  
c.  The man to whom liberty, we could never grant ti…

(12)  

a.  * I believe that this book, you should read it.  
b.  * … that this solution, I proposed it last year is widely known.  
c.  * The man to whom liberty, we could never grant it…
On the basis of such data, Lasnik & Saito propose that while in general, topicalization is adjunction to IP (our TP),\(^9\) it may optionally involve movement to SpecCP in matrix contexts. By assuming that the position for topics in English is restricted to one base-generated projection per sentence (here, TopP), they can account for the contrast of multiple topicalization versus multiple HTLD (pp. 78f.):

\((13)\)  
  a. *John, Mary, he likes t.
  b. *John, Mary, he likes her.
  c. *Mary, John, he likes t.

And the same holds for fronting two objects, rather than a subject (where angled brackets indicate pronouncing either one):

\((14)\)  
  a. This book, <to> John, we should give it <to> t.
  b. *We, this book, t, should give it to John.
  c. *<To> John, this book, we should give it <to> t.

The upshot of Lasnik & Saito’s discussion with respect to the present issue is that LD in English may involve only an HT, as there is only one TopP available to (base-generated) topics, and (HT)LD undoubtedly involves base-generation. This crucially relies on assigning LDed elements a topic-like status, placed in TopP, albeit a position reserved for direct insertion only. We will come back to finer grained differences soon.

If HTLD involves base-generation of the initial element in a left-peripheral position which necessarily precedes the placement of moved Wh-elements as well as topics,

---

\(^9\) As similarly proposed by Baltin (1982), who also provided the c-examples. Recall from fn. 4 from chapter 3 that for Lasnik & Saito, adjunction of an element creates an additional maximal projection, while for Baltin it does not. This difference in detail shall not distract us, though, from what matters: an adjoined element is base-generated in that position, while a moved element targets a separate specifier (here, adjunction to TP for base-generated topics and HTs, vs. SpecCP for moved topics and also Wh-phrases).
and if there is only one base-generated topic (or LD) position per clause (in English), we would expect that multiple embedding is just as impossible. Apart from example (13b) above, repeated here as (15a), Culicover (1996) notes (15b):

(15) a. * John, Mary, he likes her.
   b. * I suggest (that) on your vacation, the beers that you drink, (that) you should keep a record of them there.

As (15) shows, multiple embedding of an HT is indeed ungrammatical. Culicover argues on the basis of (16b), however, that multiple embedded topicalization, although somewhat marginal, is possible even in English:

(16) a. * John, Mary, tij likes tji.
   b. (?) I suggest ?(that) on your vacation, the beers that you drinkj, (*that) you should keep a record of tjt i.

While there are good reasons to rule out (16a) with either topic as the subject or object, (16b) suggests that nothing in principle rules out (multiple) embedded topics. Other languages allow embedded topics more freely, as we will see.

There is thus a certain “root” character particular to HTLD, as already noted by Emonds (1970). As we will see, this is a property that not only distinguishes it from topicalization, but also from CLD (if only marginally) and CLLD (very clearly). Moreover, if the “root” is somehow connected to a specifier position—not unreasonable, as I will show—we might get some mileage out of the intricate proposal from the previous chapter that aimed to distinguish specifiers and adjuncts formally. The structures I will eventually assign to HTLD vs. CLD and CLLD bring out this difference clearly in that one targets SpecCP, the other AdjCP. But first things first.
4.2.5 How Contrastive and Clitic Left Dislocation Fit in

Now that we have good reasons to assume that topicalization involves movement of the topic, while HTLD involves base-generation, let us consider the two other types of LD. In this sub-section we will see in some detail that CLD and CLLD pattern with topicalization in that the fronted XP shows properties of movement, unlike the HT, as seen above. To keep this section—and by extension, the entire chapter—reasonably clear, I selectively discuss relevant data from the rich literature in abridged form (see Grohmann 1997a, in press for more, and references cited).\(^\text{10}\)

\(^{10}\) The literature on all three types of LD is by now very rich, often, however, inconsistent—not at least due to mixing up terminology, properties, frameworks, and what one has to do with the other. Fortunately, the data on HTLD across languages are pretty robust and apart from occasional confusion regarding CLD and CLLD with respect to HTLD, we can draw from many sources (Ross 1967, Rodman 1974 for English, van Riemsdijk & Zwarts 1974 for Dutch, Cinque 1977, 1983, 1990 for Italian, Altmann 1981, Scherpenisse 1983 for German, and many more).

The data for CLD (and various differences to HTLD and comparisons with CLLD) I present are mostly well-known—especially from Dutch. Dutch and German do, however, differ in one critical aspect: only the latter shows explicit Case-marking and -matching. Nevertheless, the properties ascribed to Dutch CLD (Vat 1981, van Haaften, Smits and Vat 1983) carry over to German and allow for finer diagnostics for movement. Relevant literature on LD constructions in German, Dutch and Greek which I draw from include—apart from the above and all relevant references mentioned in these works—Demirdache (1991, 1997), Anagnostopoulou (1994, 1997), Wiltschko (1995a, 1995b, 1997), Grohmann (1997, 2000b, 2000c, in press). Neither this list nor the data covered in this chapter claim exhaustiveness.

Also, Vat (1981), van Haaften, Smits & Vat (1983) are responsible for discovering the majority of differences between HTLD and CLD in Dutch (and German) and also made the explicit connection between CLLD, then understood as movement-derived as proposed by Cinque (1977), and CLD (see also Demirdache 1991, Anagnostopoulou 1997). A comprehensive (if incomplete) comparison as presented here cannot hurt, though, as the technical explanations and implementations have changed over the years. Moreover, my final analysis is substantially different from previous approaches which in and of itself warrants a full discussion of the phenomenon.
4.2.5.1 Contrastive Left Dislocation vs. Topicalization

Some properties of topicalization, which support an analysis that moves the topic from its base-generated position (leaving aside possible operator movement for the above reasons), are the following:

(17) Reasons for topicalization qua movement
   i. \( \theta \)-selected
   ii. Case-marking
   iii. unboundedness
   iv. island-sensitivity
   v. reconstruction
   vi. any XP

Ad (17i), (argument) topics, which I shall concentrate on for obvious reasons, are thematically selected by the verb of the clause they are interpreted in, regardless of their surface position, and, as (17ii) expresses, they also receive Case from that clause. The dependency between a topic’s base position and its surface position can span across clauses, in principle unbounded (i.e. (17iii))—but it is sensitive to intervening islands (at least some; cf. (17iv)). Most importantly, topics allow for reconstruction (= (17v)), that is they can be interpreted in a lower position which has consequences for binding relations, quantifier interaction and idiomaticity, among others.

As we will see here and in the following, CLD shares all these characteristics. The question we face is which of the two elements involved in CLD are responsible for these characteristics, given that a CLDed XP and its RP always occur right next to each other. Here, the property (17vi) comes in: any XP may be topicalized, and also CLDed.

Let us start with (17vi), then, so that we can unambiguously refer to the CLDed XP as the relevant element to check for movement properties. In (1b) we CLDed an argument DP, but argument PPs can be CLDed as well, and so can APs, and even VPs/TPs,
or whatever their categorial status (cf. Haider 1990, Hoekstra & Zwart 1994, Müller 1998).\textsuperscript{11} Compare topicalization and CLD.

(18) a. \textit{An seinen Freund} hat Martin den ganzen Tag gedacht.

\hspace{1cm} b. \textit{An seinen Freund, an den} hat Martin den ganzen Tag gedacht.

\hspace{1cm} \text{at his friend at RP has Martin the whole day thought}

\hspace{1cm} ‘Of his friend, Martin thought all day.’ \hspace{1cm} \textit{(German: PP)}

(19) a. \textit{Glücklich} war der Martin schon lange nicht mehr.

\hspace{1cm} b. \textit{Glücklich, das} war der Martin schon lange nicht mehr.

\hspace{1cm} \text{happy at RP was the Martin already long not more}

\hspace{1cm} ‘Happy, Martin hasn’t been in a long time.’ \hspace{1cm} \textit{(German: AP)}

(20) a. \textit{Billiard spielen} kann der Martin ziemlich gut.

\hspace{1cm} b. \textit{Billiard spielen, das} kann der Martin ziemlich gut.

\hspace{1cm} \text{pool play at RP can the Martin pretty well}

\hspace{1cm} ‘Play pool, Martin can do pretty well.’ \hspace{1cm} \textit{(German: VP/TP)}

Given the different categories that can front in either construction, it is very unlikely that XP is actually base-generated in its surface position: first, it is not an option for HTLD and second, there would be optional selection of a (\textit{d}-pronominal) DP.

In (21) we can see that the dependency can be long distance, even unbounded.\textsuperscript{12,13}

\textsuperscript{11} I refrain from further discussion of (20) under a Zwartian approach, as that would take us too far afield. See section 4.3.4 for a little bit more.

\textsuperscript{12} I constrain myself to singular masculine direct objects as much as possible, because unlike the homomorphemic feminine and neuter forms, we can see the difference between nominative and accusative.

\textsuperscript{13} German CLD is restricted to V2 environments and as such can only be embedded under “bridge verbs” that allow such embedding (see Müller & Sternefeld 1993). I will disregard here whether these are really instances of extraction or could fall under parenthetical structures (Reis 1996).
If an island is in the way of the dependency between surface and base position, the structure becomes ungrammatical, as (22) shows with a CNPC violation (where the island boundary is marked in boldface).

(22) a. *Den Martin hat Maria die Tatsache geglaubt mag jeder.
b. *Den Martin, den hat Maria die Tatsache geglaubt mögen alle.
   the.ACC Martin RP.ACC has Maria the fact believed like all
   *‘Martin, Maria believed the fact (that) everyone likes.’
   (German: island-sensitive)

(23) indicates that both topicalization and CLD allow reconstruction: a pronominal within the fronted XP can be bound by a lower quantifier (where intended binding/coreference relations are indicated by underlining here and in the following).

(23) a. Seinen besten Freund sollte jeder gut behandeln.
b. Seinen besten Freund, den sollte jeder gut behandeln.
   his.ACC best friend RP.ACC should everyone well treat
   ‘His best friend, everyone should treat well.’ (German: reconstruction)

I will return to these properties in more detail below. At first glance, though, it looks as if topicalization and CLD in German go hand in hand—and, by extension, in English also, as the translations show. As a working hypothesis we can thus assume that the CLDed XP is derived by movement—which throws up the questions of how the RP enters the construction, and which position the XP moves to. If the RP sits in topic posi-
tion, it is unlikely that the XP occupies a regular topic position also, otherwise we would expect the finite verb to show up immediately after the XP; after all, matrix topicalization in German is a V2 structure. Moreover, if XP has moved, it can only have moved into a specifier position and not adjoined to, say, TopP (see section 3.2.4).

4.2.5.2 Contrastive vs. Hanging Topic Left Dislocation

Let me now introduce the syntactic phenomenon of German CLD and HTLD in its full range. This sub-section introduces the patterns they come in, the main differences, and clear evidence in favour of movement of the CLDed XP.

4.2.5.2.1 Intonational break

Let me illustrate German CLD and two different types of HTLD (which, for lack of better terminology, I refer to as HTLD I and II, respectively) by way of one noteworthy difference: their intonational pattern. In both types of HTLD, XP and RP are separated by a clear pause (marked ‘#’ in (24)); this is not the case in CLD (or CLLD in Greek or, for the most part, topicalization in any language):

   this.ACC man RP.ACC have I yet never seen
   ‘This man, I’ve never seen before.’ (German CLD)

b. *Diese-r/-n Mann, # den/ihn habe ich noch nie gesehen.*
   this.NOM/ACC man RP.ACC/him have I yet never seen
   ‘This man, I’ve never seen him before.’ (German HTLD I)

c. *Diese-r/-n Mann, # ich habe den/ihn noch nie gesehen.*
   this.NOM/ACC man I have RP.ACC/him yet never seen
   ‘This man, I’ve never seen him before.’ (German HTLD II)
Instances of HTLD,\(^\text{14}\) regardless of high or low resumptive, regardless of \(d\)- or \(p\)-pronoun, regardless of Case-marking contain a clearly audible, intonational break between the HT and the subsequent part of the sentence. CLD does not have such a break.

4.2.5.2.2 The “contrastive” in CLD

One property ascribed to CLD, setting it apart from HTLD, is its “contrastive” usage. While an HT can be employed to add new information to the current discourse, a CLDed XP must be discourse-old, while a HT can also be discourse-old, it cannot be used contrastively, though. Given an information question (asking for new information), of the following six permutations, only three are felicitous (infelicity marked “#”):\(^\text{15}\)

(25) Wen hast du gestern getroffen?
‘Who did you meet yesterday?’

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Ich habe gestern \textit{den Martin} getroffen.</td>
</tr>
<tr>
<td>b.</td>
<td># \textit{Den Martin} habe ich gestern getroffen.</td>
</tr>
<tr>
<td>c.</td>
<td>\textit{DEN MARTIN} habe ich gestern getroffen.</td>
</tr>
<tr>
<td>d.</td>
<td>\textit{Der Martin, den} habe ich gestern getroffen.</td>
</tr>
<tr>
<td>e.</td>
<td># \textit{Der Martin, ich habe ihn} gestern getroffen.</td>
</tr>
<tr>
<td>f.</td>
<td># \textit{Den Martin, den} habe ich gestern getroffen.</td>
</tr>
</tbody>
</table>
‘I met Martin yesterday.’

\(^\text{14}\) In (24b,c) I indicate the optional Case-agreement in HTLD, an illustration which I drop in subsequent examples for clarity. The diagnostics we will see in the following can be used to tease HTLD and CLD further apart. The \(d\)- and \(p\)-pronouns are not interchangeable in all instances of HTLD, as illustrated here (see e.g. van Riemsdijk & Zwarts 1974, Altmann 1981, van Haafken, Smits & Vat 1983). Apparently, the preferred occurrences are high \(d\)- (HTLD I) and low \(p\)-pronoun (HTLD II); furthermore, the \(p\)-pronoun disproves of Case-matching most. I will illustrate all instances of HTLD with the HT proper (\textit{nominativus pendens}), so this should not concern us much. I will not address finer differences further, but employ for simplicity high \(d\)- and low \(p\)-pronouns throughout, pointing out relevant differences.

\(^\text{15}\) Leaving aside more intricate prosodic properties and stress patterns, the “unmarked” order in (25a) shows default accent on \textit{den Martin} ‘the Martin’. I address these different patterns in the next chapter also, clarifying why I refer to (25e) as “focused” rather than “topicalized”. For the purpose of exposition, the contrastive accents on the LDed constituents and resumptives are not marked either.
In the unmarked order (25a), *den Martin* ‘the Martin’ receives (default) focus stress, signaling new information against the broadest range of textual presuppositions. Fronting an object to the first position of a V2 matrix clause is usually taken to be an instance of topicalization, and as such an infelicitous answer to the question; cf. (25b). If it is focus-stressed, however, the reply is felicitous, pretty much on a par with the unmarked order, as in (25c). Regarding the three HTLD possibilities in German—labeled here HTLD I (with a high *d*-pronoun), HTLD II (with a low *p*-pronoun) and CLD (with high *d*-pronoun matching in Case), which I will come back to momentarily—, only HTLD I in (25d) is acceptable. Compared with HTLD II in (25e), we can already observe a difference, at least in usage, between the two types; we will see some syntactic differences in the following. Compared with CLD in (25f), we can observe the same, but as we will see, the syntactic differences between CLD and either type of HTLD outrank the HTLD-internal differences by far.

To control for felicity and contrastiveness, consider the following exchanges:

(26) Hast du gestern die Maria getroffen?
    ‘Did did you meet Maria yesterday?’

a. Nein. Ich habe gestern *den Martin* getroffen.                  (unmarked)
b. # Nein. *Den Martin* habe ich gestern getroffen.                (topicalized)
d. # Nein. *Der Martin, den* habe ich gestern getroffen.           (HTLD I)
e. # Nein. *Der Martin, ich habe ihn* gestern getroffen.           (HTLD II)
    ‘No. I met Martin yesterday.’

In this context, which provokes contrast, CLD is felicitous, while HTLD I is not. This might not be the whole story (see Altmann 1981, van Riemsdijk 1997; the obligatory “contrastiveness” in CLD has also been questioned by Frey & Grabski 1999), but I leave the discussion about the “C” in CLD at that.
4.2.5.2.3 Towards an account: The role of Prolific Domains

Putting the discussion into a somewhat larger perspective, consider the role of a clausal tripartition. Applying Prolific Domains as rough indicators, the phenomena we are dealing with can be characterized as follows (see also the explanatory note in fn. 14):

(27)  a. \[\ldots HT d\text{-}pr\ldots V \ldots d\text{-}pr\ldots \ldots [\ldots d\text{-}pr\ldots ]\ldots)] (HTLD I)

a’. \[\ldots HT \ldots p\text{-}pr\ldots \ldots [\ldots p\text{-}pr\ldots ]\ldots)] (HTLD II)

HT in NOM, appropriately Case-marked RP: high d- or low p-pronoun

b. \[\ldots XP d\text{-}pr\ldots V \ldots ??? \ldots \ldots [\ldots ??? \ldots ]\ldots)] (CLD)

XP appropriately Case-marked, matching RP: high d-pronoun

One of the goals of the following discussion is to narrow down the positions occupied by HT, XP, d- and p-pronoun. As indicated here, the first three are most likely part of the \(\omega\)-domain, while the \(p\)-pronoun is arguably in its \(\phi\)-position (where the position of the finite verb will not concern us here). We will justify these assumptions. I also indicate that in case of movement, the \(d\)-pronoun is the element that moves; that follows from the preliminary characterization of HTLD as a base-generation phenomenon. On the other hand, it is not quite clear at this point which element moves in CLD (hence ‘???’).

Appropriate Case-marking leaves the Case value open: usually, direct objects are marked accusative, indirect objects dative or genitive etc. The following examples illustrate:

(28)  a. Der nächste Präsident, \(\text{der}\) ist ja wiedergeboren.

the.NOM next president RP.NOM is PRT born-again

‘The next president, [he] really is born-again.’

b. Der Industrie, \(\text{der}\) versprechen die Kandidaten viel.

the.DAT big-business RP.DAT promise the candidates much

‘Big business, the candidates promise a lot.’
c. *Der Armut, der gedenken die Kandidaten wenig.*

The GEN poverty RP.GEN recall the candidates little

‘Poverty, the candidates recall very little.’ (German)

Thus, subjects can be CLDed (which, to translate as a topic structure, needs additional insertion of a pronoun) and both elements show up in nominative. Likewise, indirect objects show up in dative and in CLD whose verbal predicate is an irregular verb that marks genitive on its object results in genitive-marking on LDPed XP and RP. Needless to say, the corresponding HTLD structures of (28) would all involve a nominative-marked HT and an RP with Case-marking according to the predicate.

4.2.5.2.4 Topicalization, Wh-questions and CLD vs. HTLD

But the real difference between HTLD and CLD, and as such demonstrating that all instances above do indeed have the status that I assigned, is one of derivational history. The arguments in favour of base-generating the HT in English from section 4.2.4 carry over to German. Moreover, given that we have seen in the previous sub-section that CLDed XPs are likely to be derived by movement, we have a nice way of teasing the two apart.

To demonstrate the first claim, consider the following. Matrix topicalizations and Wh-questions have in common that they involve the fronted XP in first, the finite verb in second position; let us call this COMP position CP for the time being.

(29) a. *Welchen Gast hat der Oberkellner gegrüßt?*

Which guest has the maître d’ greeted

‘Which guest did the maître d’ greet?’

b. *Diesen Gast hat der Oberkellner gegrüßt.*

This guest has the maître d’ greeted

‘This guest, the maître d’ greeted.’ (German)
As in English (cf. (7)), topic and fronted Wh-phrase cannot co-occur as such:

(30) a. * Diesen Gast, wann hat der Oberkellner gegrüßt?
   this. NOM guest when has the maître d’ greeted
   *‘This guest, when did the maître d’ greet?’
   b. * Wann, diesen Gast hat der Oberkellner gegrüßt?
   when this guest has the maître d’ greeted
   *‘When, this guest the maître d’ greeted?’
   (German)

Also as in English (cf. (8)), Wh- and HTLD constructions are well-formed:

(31) a.  Dieser Gast, wann hat den der Oberkellner gegrüßt?
      this. NOM guest when has RP.ACC the maître d’ greeted
   ‘This guest, when did the maître d’ greet?’
   (German)

I illustrate this with a low RP for a good reason (HTLD II with both types of RP):
given that the high RP is in topic position and that root topicalization and Wh-question formation do not easily go hand in hand, we would have to test with some other structure.
The high RP is, as expected, not good, neither in HTLD I (32a) nor in CLD (32b); while perfectly well-formed questions, these are not information questions but rhetorical ones, hence marked with a hash mark.

---

16 At least not in pure matrix contexts under the assumption that only the first XP of a V2 structure can be a moved topic. I will address this issue in more detail in the next chapter.
Like English, the German topic and LDed XP can go together, provided that the topic follows the LDed XP (cf. (14)):

      this.NOM candidate a kick-in-the-ass should one RP.ACC/him give  
      ‘This candidate, a kick in the ass, one should give him.’

      a kick-in-the-ass this.NOM candidate should one RP.ACC/him give  
      ‘A kick in the ass, this candidate, one should give him.’  
      (German)

So far we can note that German and English HTLD (II) seem to pattern together, both with respect to topicalization and independent of it. German CLD and HTLD I do not seem to go with topicalization or Wh-movement. The violating element seems to be the high d-pronoun. Let us assume for the time being then that HTLD II must involve a base-generated HT (like English) and the reason that HTLD I and CLD are bad in the circumstances mentioned is that they involve a topicalized RP (unlike English). This does not tell us much about the nature of the LDed XP in HTLD I and CLD, however. We are going to investigate their possible differences next.

Recall the reconstruction property (17v) and the initial demonstration (23). We will now see that CLD systematically allows for reconstruction of the LDed XP, while HTLD does not. The following data also show unambiguously that it must be the XP and
not the RP that reconstructs. Moreover, within Copy Theory we take reconstruction to be a sound indication of movement (Chomsky 1995b, Hornstein 1995, Nunes 1995, Fox 1999). As such, the following data provide a strong argument in favour of XP- rather than RP-movement in CLD, and base-generation in HTLD constructions.

4.2.5.2.5 Weak crossover

CLD allows bound variable readings of pronouns, in particular in a potential weak crossover (WCO) configuration arising from a quantificational element in the matrix clause and a pronominal element inside the LDed XP, illustrated here with a strong quantifier in subject position. In other words, CLD does not give rise to a WCO effect, strongly suggesting that the LDed XP may reconstruct at LF to a position from where it is c-commanded by the quantifier. As (34) shows, this behaviour mirrors topicalization, well-known to obviate WCO effects in German.


18 As above, I mark the RP in bold for clarity and underline intended binding relations throughout. Every illustration comes in two pairs: the first pair contains topicalization and CLD, while the second has HTLD I and II. The translation of CLD qua topicalization and HTLD qua LD are further suggestive. Given the explicit presentation and notation so far, I consider it safe to dispense with explicit glosses for all of the following sentences. I will thus only gloss CLD (see also fn. 14, all else should be clear).
(34) a. Seinen Rasen mäht jeder Herforder Bürger samstags.
    b. *Seinen Rasen, den mäht jeder Herforder Bürger samstags.

    His lawn mows every Herfordian dweller Saturdays.

    ‘His lawn, every Herfordian mows on Saturdays.’

(German)

HTLD, regardless of the nature and position of the RP, does not allow the bound variable reading; as this is the only reading of interest, we can ignore the fact that the structures in (35) are well-formed, just in case the pronoun inside the LDe XP refers to a specific person (indicated by the hash mark).

(35) a. # Sein Rasen, den mäht jeder Herforder Bürger samstags.
    b. # Sein Rasen, jeder Herforder Bürger mäht ihn samstags.

    #‘His lawn, every Herfordian mows it on Saturdays.’

(German)

This observation also holds across clauses, where the relevant element is extracted and must be interpreted inside the embedded clause:

(36) a. Seinen Rasen glaubt jeder Herforder, kann er schön halten.
    b. *Seinen Rasen, den glaubt jeder Herforder, kann er schön halten.

    ‘His lawn, every Herfordian believes he can keep pretty.’

(German)

Moreover, it does not matter whether the RP in HTLD is also extracted (37a) or occurs in the embedded topic position (37a’):

(37) a. * Sein Vorgarten, den glaubt jeder Herforder, kann er schön halten.
     a’. * Sein Vorgarten, jeder Herforder glaubt, den kann er schön halten.
     b. * Sein Vorgarten, jeder Herforder glaubt, er kann ihn schön halten.

     *‘His lawn, every Herfordian believes he can keep it pretty.’

(German)
The absence of WCO effects in these contexts can be captured if it is the LDed XP itself that undergoes movement from lower down in the structure in CLD, but not in HTLD, pointing to a derivational difference between CLDed XPs and hanging topics as well as the relevant RPs.

4.2.5.2.6 Condition A

Similarly, only an anaphor inside a CLDed XP may be coreferent with a lower pronoun and/or an R-expression. Thus, the absence of Condition A effects further points to movement of the CLDed XP, as opposed to the hanging topic, on the same grounds (i.e. reconstruction). This can best be shown with the reciprocal einander ‘each other’;

(38) illustrates for CLD and topicalization, and the minimal pair in (39) for HTLD:

(38)  a.  Freunden von einander, erzählen Herforder selten Lügen.
     b.  Freunden von einander, denen erzählen Herforder selten Lügen.

‘Friends of each other, Herfordians rarely tell lies.’


*‘Friends of each other, Herfordians rarely tell them lies.’

We can find the contrast with a reflexive inside the LDed element as well (see also Vat 1981):

(40)  a.  Einen Grill bei sich im Garten hat der Alex wohl.
     b.  Einen Grill bei sich im Garten den hat der Alex wohl.

‘A grill in his own garden, Alex surely has.’
4.2.5.2.7 Condition C

If WCO and Condition A effects can be obviated by movement of the CLDed XP, we would now expect that an R-expression inside the LDed XP coreferent with a lower pronoun leads to ungrammaticality in CLD, but not in HTLD. Indeed, we can observe a Condition C effect only in CLD:

(41)  a. * Ein Grill bei sich im Garten, den hat der Alex wohl.
   b. * Ein Grill bei sich im Garten, der Alex hat ihn wohl.
   *‘A grill in his own garden, Alex surely has it.’  (German)

(42)  a. * Der Tatsache, daß Alex arm ist mißt er keine Bedeutung bei.
   b. * Der Tatsache, daß Alex arm ist, der mißt er keine Bedeutung bei.
     *‘The fact that Alex is poor, he doesn’t attach importance to.’  (German)

In HTLD, on the other hand, we find well-formedness on all levels; the absence of Condition C effects suggests base-generation of the HTLDed XP in its surface position and any relevant movement of the RP only.

(43)  a. Die Tatsache, daß Alex arm ist, der mißt er keine Bedeutung bei.
   b. Die Tatsache, daß Alex arm ist, er mißt ihr keine Bedeutung bei.
   ‘The fact that Alex is poor, he doesn’t attach importance to it.’  (German)

The data above suggest that our intuition bears some empirical merit. Given that (LF-)reconstruction can only target a position where a copy has been left behind by previous movement, we have a strong piece of evidence that connectivity in LD constructions depends on the status of the LDed XP itself, whether it has moved to or is base-generated in its left peripheral surface position.
4.2.5.2.8 Idioms

Another classic argument strongly suggesting that it really is the LDed XP that moves in CLD comes from displacing idiomatic chunks. As shown by Marantz (1984) and many others, these must be the result of movement, where the idiomatic interpretation is yielded by a strictly local (base-generated) configuration.\(^{19}\)

\[(44)\] Der Martin hat der Birgit (einst) den Kopf verdreht.
the Martin has the Birgit (once) the head twisted

‘Martin turned Birgit’s head (once).’ \((\text{German})\)

The idiom in (44) may be manipulated structurally without losing its idiomatic reading, including the relevant chunk *den Kopf* ‘the head’. (This may not hold for the English equivalent; see Schenk 1995, for instance, for discussion of syntactic properties of English idioms, and corresponding well-formed reordering.) In (45), various parts of the clause are topicalized, in (46) the idiomatic chunk is CLDed and topicalized, respectively; (47) illustrates attempted HTLDing of the relevant chunk.

\(^{19}\) As van Riemsdijk & Zwarts (1974) observe, LD with idioms leads (usually) to ungrammaticality. However, van Haaften, Smits & Vat (1983: 137f.) note that “[i]diom chunks in HTLD, where no connectedness is expected, are invariably out […] Also in the case of inalienable possession idioms, which can be read literally as well, those involving HTLD […] are either ungrammatical under any reading or can only be read literally.” And the idioms from Dutch they present show that CLD retains the idiomatic reading. Examples from the literature include the following:

(i) a. *Mijn been, dat* heb ik gebroken.
my leg RP have I broken
b. *Mijn been, ik heb* het gebroken.
‘My leg, I broke it.’ \((\text{Dutch, Vat 1981})\)

my ACC song RP ACC have I at her never wail can
‘My troubles, I’ve never been able to pour (*them) out to her.’ \((\text{German, Vat 1981})\)
(45)  a. Der Birgit hat der Martin einst den Kopf verdreht.
     b. Einst hat der Martin der Birgit den Kopf verdreht.
     c. Einst hat der Birgit der Martin den Kopf verdreht.
     ‘Martin once turned Birgit’s head.’  (German)

     ‘Birgit’s head, Martin once turned.’  (German)

     b. * Der Kopf, der Martin hat ihn einst der Birgit verdreht.
     *‘Birgit’s head, Martin turned it yesterday.’  (German)

As it turns out, those speakers that judge (45) grammatical, especially (45b,c),
also accept (46a), CLDing the idiomatic chunk, but not (47). Again, under an analysis
where the CLDed XP is derived by movement, as suggested above, the result is nothing
but expected. Let us next turn to finer differences in derivation and structure. We could
now modify (27) from above as follows, specifying the previous ‘???’:

(48)  a. [\_. HT \_d-pronoun V [\_ \_d-pronoun \_ [... d-pronoun \_ ...]]]  (HTLD I)
     a’. [\_. HT [\_ \_p-pronoun \_ [... p-pronoun \_ ...]]]  (HTLD II)
     b. [\_. XP d-pronoun V [\_ \_XP \_ [... XP \_ ...]]]  (CLD)

4.2.5.3 Contrastive vs. Clitic Left Dislocation

With this picture in mind, let us turn to CLLD, which also has been argued to in-
volve movement of the LDed XP. There are a number of differences, but also similarities,
between CLD and CLLD. The most obvious difference concerns the resumption: in one
case we are dealing with a \_d-pronoun, in the other with a bona fide clitic. The latter,
moreover, is arguably not in topic position. It appears wherever the clitic element would
appear for that particular construction in the particular language. Given that CLLD occurs in languages that have special clitics (in Zwicky’s 1977 sense), and which are verb-related, the clitic basically surfaces where the verb does.\(^{20}\)

Consider thus the following CLLD data from a number of languages:

(49) a. *Tintixi tu kathetfoxostin ekane pigenontas stin Ameriki the luck.ACC his,GEN every poor CL.ACC made going to-the America ‘The poor made their luck/fortune by going to America.’

(Greek; Anagnostopoulou 1997: 155)

b. *A se stessa, Maria non ci pensa.
of herself Maria not CL think.3s

‘*Her/Herself, Maria doesn’t think of.’

(Italian; Cinque 1990: 59)

c. *A su hijo ninguna madre desea que se lo regañe.

ACC her son no mother desires that SE CL.ACC scold ‘Her son, no mother desires that someone scold.’

(Spanish; Zubizarreta 1998: 114)

d. *Nadia u aal t-la l-m allme?

Nadia what said.3sf.-CL.dat the-teacher ‘Nadia, what did the teacher say to [her]?’

(Lebanese Arabic; Aoun & Benmamoun 1998: 570)

As indicated by these examples, CLLD shows the same movement diagnostics that we have seen above. The CL matches in Case with the XP (49a), idiomatic chunks may be CLLDed (49a), connectivity holds between an item inside the XP and the CL (49b), a bound pronoun reading is available between the XP and a lower quantificational subject (49c), and non-DPs may be CLLDed (49b,c).

\(^{20}\) I will return to the role of clitics, and why German CLD employs d-pronouns, in more detail in section 4.5.
But there are also some differences. Embedding (49c) and violation of a (weak) island (49d) are just two. Others include the obvious: CL rather than d-pronoun and low, clause-internal position rather than the high, topic position. Following Anagnostopoulou (1997: 159f.), I take the last difference to be that CLLDed XPs, unlike CLDed ones, may be stacked, i.e. they are not restricted to a unique occurrence per clause.

I will return to these differences in section 4.4 when I consider an application of my analysis of CLD to CLLD. For now, we can take (49) to indicate that there is an interesting correlation between the syntactic behaviour of CLD and CLLD—both different from HTLD and more like topicalization: the LDed XP is derived by movement.

4.2.6 Summary

This is where we stand right now. LD comes in three types: HTLD, CLD and CLLD. Concentrating on the first, and comparing it with topicalization, we found that the two share some properties, but differ in others. Topics, but not HTs, can occur in embedded contexts, and the two can co-occur, given that the HT precedes the topic. Moreover, topics cannot co-occur with fronted Wh-phrases, while HTs can (if they precede them).

A natural explanation comes from the different derivational history of the two: topicalization is a movement process that targets SpecTopP, while HTs seem to be base-generated in their surface position. We took this line of reasoning and investigated how, if at all, CLD and HTLD in German differ. We found that CLD shows indication that the LDed XP has moved to its surface position (like English topicalization), while the HT seems to be base-generated (like English HTLD). The main evidence comes from connectivity effects, instances where the XP seems to be able to reconstruct and by doing so license otherwise illicit configurations (WCO, Condition A) or destroys an otherwise available reading (Condition C), all in stark contrast to both types of HTLD, the one with a low and the one with a high RP. Moreover, idiomatic chunks may be CLDed, but not
HTLDed.\textsuperscript{21} We also saw that HTs are followed by an intonational break, while a CLDed XP and its RP can be pronounced in fast succession.

We then compared CLD very briefly with CLLD, for which it has also be claimed that it involves movement— for pretty much the same reasons. We stumbled over some differences between the two which we will come back to once we have an appropriate analysis for CLD. This we will approach in the next section, starting with a hypothetical configuration as follows (from (48) above):

(50) a. \([\ldots \text{HT} \, \text{d-pronoun} \, \text{V} \, \ldots \text{d-pronoun} \, \ldots \text{]}] \) \textit{(HTLD I)}

a’. \([\ldots \text{HT} \, \ldots \text{p-pronoun} \, \ldots \text{]}] \) \textit{(HTLD II)}

b. \([\ldots \text{XP} \, \text{d-pronoun} \, \text{V} \, \ldots \text{XP} \, \ldots \text{]}] \) \textit{(CLD)}

At first glance, the structure for CLD, (50b), is reminiscent of the structure we have seen in the previous chapter, where a full XP and a coreferent pronoun (in “identity” relation) occupy two positions in the same Prolific Domain. I capitalize on this similarity and explore an application of Copy Spell Out to account for the properties of CLD.

4.3 Copy Spell Out in the \(\omega\)-Domain

The extensive discussion so far reveals that CLD is unambiguously the result of movement of the LDed XP, while HTLD must involve a base-generated XP. In line with the framework presented in this work so far, I will propose an analysis that not only captures movement of the XP, but also resumption through a rather specific pronominal ele-

\textsuperscript{21} These findings confirm the earliest approaches to LD (such as Ross 1967, and then van Riemsdijk & Zwarts 1974, Cinque 1977, 1983), which have subsequently been used as the standard of comparison for all discussion on LD; see also references in fn. 17.
ment, much in the spirit of reflexivization. After the initial presentation of the analysis I present additional facts that ask for a slight modification in understanding the HT as adjoined to the position whose specifier may be targeted by the CLDed XP.

4.3.1 **Exclusivity: Towards a Derivational Analysis**

Now that we have the tools, the analysis should be obvious. In CLD we are dealing with an LDed XP in a left-peripheral position and a resumptive element in topic position—both arguably part of an articulated COMP, or in our current terminology: both elements sit in a position within the \(\omega\)-domain.

Let us take the surface positions of CLDed XP and RP to be as follows:

\[(51) \quad [C_P \ X_P \ C \ [_{\text{TopP}} \ R_P \ V-\text{Top} \ [_{\text{TP}} \ \ldots \ ]]] \quad (C_L D)\]

As mentioned, the RP sits in topic position, an assumption supported by the facts that (i) it is immediately followed by the inflected verb, that (ii) the construction is a well-formed sentence even if the XP is dropped, and that (iii) it behaves like a topic when compared to Wh-questions. The XP, on the other hand, has a “quasi-sentence-external” character: (i) it does not have to be present (where absence yields a well-formed topic structure), (ii) it tends to be root-related, and (iii) it cannot be preceded by anything else.

In our adaptation of Rizzi’s (1997) split CP, the highest projection is CP (his Force). This position does not seem like an implausible candidate to host the LDed XP: as we will see in the next chapter, CP usually hosts clause-typing elements (particles or operators), and CLD does not go well with clause types other than the, possibly canonical, declarative.
CLD cannot co-occur with Wh-phrases to express an interrogative, not even to yield a rhetorical question, as (52a) shows. Here the only interpretation is of a specific garden (hence ‘#’), more on par with HTLD which lacks the bound pronominal reading. Neither can the CLDed XP be used in an overt imperative, viz. (52b), nor express an ex-claimative, attempted in (52c) with an idiom chunk (regardless of the relative positions of d-pronoun and exclamative marker wie ‘how’). On the other hand, HTLD is acceptable in (rhetorical) questions, shown in (53a). A HT can also be employed in an imperative, shown in (53b), or an exclamative, cf. (53c).\footnote{Note that the structures in (53) can really (only) be instances of HTLD. While Case-matching is possible, it can only be optional. Thus, (53b), for example, can involve XP and high d-pronoun both in accusative, but as connectivity is absent in these constructions, we can conclude that we are dealing with HTLD. In (i), the bound variable reading is not possible for seinen Wagen ‘his car’:

(i) #Seinen Wagen, den wasch jetzt jeder!
    his.NOM(ACC) car RP.ACC wash.IMP now everyone
    #‘His car, everyone wash now!’ (German)
If CP hosts operators expressing these clause types, filling the position with a CLDed XP arguably predicts the complementary distribution.

Recall from the previous chapter that we analysed the identity between anaphors and their antecedents in terms of Copy Spell Out: two non-distinct copies of one element occur within the same Prolific Domain, and in order to rescue Exclusivity, the lower copy had to be spelled out. If we assume that the RP in CLD is not a lexical item, part of the initial numeration (or the lexical array LA), we might play the same game here.

Let us turn to evaluating how big an assumption this is in a moment and consider the analysis I propose for German CLD:

\[
\begin{array}{c}
\text{(54) } [\text{CP XP } \text{TopP } \text{XP } \mathbin{\circ} \text{RP V-Top } [\text{TP } \ldots \text{XP } \ldots [\text{vP } \ldots \text{XP } \ldots] \ldots]]
\end{array}
\]

In a regular CLD structure, as we have seen throughout, the originally merged argument moves from its θ- to its φ-position for the usual reasons. It then topicalizes. The derivation could stop there, and we yield a topic construction. Alternatively, the element could be equipped with an additional feature motivating further movement. For now, the nature of this feature plays no major role. Call it [F]. [F] triggers subsequent movement to CP, and we face two non-identical copies of one element—one in SpecTopP and one in SpecCP. Copy Spell Out applies and the lower copy gets realized as the d-pronoun.
This analysis, if correct, captures all the properties of CLD we have seen so far. CLD, as opposed to HTLD, requires a high RP—if Copy Spell Out applies in the highest clausal Prolific Domain, the ω-domain, this restriction can be cashed out. Moreover, there is obligatory Case-matching between the CLDed XP and the d-pronoun—if the d-pronoun is the spelled out copy which has already checked its φ-features and thus been Case-marked, we would not expect anything else. The same goes for both RPs apparently satisfying the θ-selectional requirements; we will get back to that.

More interestingly, if the CLDed XP has moved from all the way down, the relevant θ-position, we also expect to find reconstruction, as we do in any other A’-configuration (such as topicalization). In other words, with respect to WCO, Condition A, Condition C etc., the CLDed XP should behave just like a topic—and it does. The same goes for idiom chunks. If displaced idiom chunks must at one point be in the configuration that licenses the idiomatic reading, this displacement can only be understood in terms of movement—and the analysis sketched in (54) can deal with this requirement also.

In addition, all the distributional properties (concerning topics, stacking, Wh-interaction etc.) also fall out—given that the final landing site is a (unique) specifier of CP, higher than Wh-elements or other topics, it precludes any other element from preceding the CLDed XP. We will see an interesting twist in the next sub-section.

The present analysis also captures the fact that any XP can be CLDed—as long as the XP in question can be motivated to topicalize, we can posit an additional [F] for further movement, necessitating Copy Spell Out. I address this and other issues in section 4.3.3.

Naturally, such an analysis does not come without its price. One question that arises concerns the nature of the resumptive. We will come back to that later. Let us first see how this analysis fares with respect to other current analyses of CLD.
4.3.2 A Comparative Analysis

As noted already, the movement character of CLD is nothing new. It has been known for a long time (cf. Cinque 1977)—but, so far, has resisted a straightforward account for several reasons. One major technical innovation of Government and Binding Theory was the introduction of (empty) operators into the syntactic component. Whether to account for scope-marking or clause-typing (of Wh-elements, quantifiers and others), or the licensing of an element by a binding DP in COMP, operators were used all over the place. Parasitic gaps, easy-to-please constructions, relativization—these constructions had in common that they involved an invisible formative which did nothing but license otherwise illicit movement steps or selection relations. An operator and the element that would be in its position, if it had been able to move, were linked by predication, a rule of construal (Williams 1980). As a consequence, as many phenomena as possible have been analysed in terms of an empty operator, so as to justify this grammar-internal formative.

Empty operators belong to the classes of items we should be highly suspicious of in a framework that aims to dispense with anything that falls short of virtual conceptual necessity. If it turns out that empty operators are virtually conceptually necessary, because their function could not be accounted for with other, existing mechanisms, we might be justified in postulating them. Recent research suggests, however, that those constructions for which empty operators have been proposed can be analysed without resorting to these devices, by employing the tools available. This said, it comes as no surprise that operators and a rule of predication have also been the favourite analysis of those LD constructions that reflect movement character. In GB tradition, recent proposals by Demirdache (1991), Anagnostopoulou (1997) or Wiltschko (1997) come to mind.

---

Under such a predication approach, the RP could serve as the relevant operator, entering a predication relation with the CLDed or CLLDed XP. Anagnostopoulou takes the $d$-pronoun in CLD to be a $d$-operator and the clitic in CLLD as its weaker counterpart, a “syntactic operator” and as such a possible A’-binder (cf. Aoun 1981, Sportiche 1983). This, for her, derives three of the five differences between CLD and CLLD she observes (see also section 4.4). The fact that the clitic cannot be stressed, unlike the $d$-pronoun, follows if it is a purely “syntactic operator” (of which null or empty operators are a variant, which obviously cannot be stressed). The fact that clitics occur lower than $d$-pronouns follows if “syntactic operators” do not have a special position in the clause, whereas the $d$-pronoun is in the “typical [o]perator position for XPs” (p. 161). And finally, the fact that CLLDed elements can be stacked, unlike their CLDed counterparts, follows from the assumption that each LDed XP must be licensed by a resumptive operator, and these are (on reasonable grounds) uniquely tied to one SpecCP position.

As we will see below, these differences fall out of the present analysis without any mention of operator-variable chains whatsoever. The difference in position follows from introducing the resumptive element in different Prolific Domains ($\omega$- vs. $\phi$-domain), and from this it also follows that one element cannot be stressed (as spelling out in form of a clitic is forced), and stackability is a by-product of different positions of the LDed element in question. I will lay out the details and empirical support below.

Other approaches in terms of operators treat the RP as an “in situ operator” (Demirdache 1991) or as a topic operator, entering a operator-variable chain with its trace and being linked to the CLDed XP by virtue of the latter’s function as a “parasitic operator” (Wiltschko 1997). On the other hand, Zwart (1993) analyses topicalization as being CLD in disguise—with an empty operator (whose overt counterpart is the $d$-pronoun). Without further ado, I disregard all such approaches as untenable on grounds of virtual conceptual necessity—given that an alternative can be motivated to account for the facts equally well with less additional assumptions.
One alternative to the operator-approach, in fact the earliest to deal with the movement phenomenon (Vat 1981), is a version of a promotion analysis for CLD, which has been proposed for relative clause-formation (“Vergnaud-raising;” cf. Vergnaud 1974, Kayne 1994, Hornstein 2000 and others). Under this approach, both the XP and the RP are base-generated in sisterhood relation, say as the argument to be CLDed. The two move then to SpecCP (or TopP), from where the XP moves on—either to AdjCP, or a higher specifier (under a proliferated COMP-structure). The drawback of this approach is that it involves base-generation of the d-pronoun, a thorn in our side if we take “Avoid Pronoun” seriously and treat pronominal elements as Last Resort formatives. The advantage of the promotion analysis is its potential similarity to CLLD, if we wanted to pursue an analysis of CLLD as an extension of “clitic doubling” (Jaeggli 1981). Under this approach, clitic and XP are also “together” at some point (either base-generated or derived), and XP moves out at some point, just as in CLD. I will discuss this alternative in a little bit more detail in section 4.4, where I present an extension of the present analysis to CLLD, and contrast it with a doubling approach.

To anticipate, my implementation does indeed have the character of clitic doubling, though without its unwanted consequences (see Cinque 1990). As such, I believe that the present analysis takes only the best from a promotion/clitic doubling approach and places it into the current framework in a natural way.

4.3.3 CLD vs. HTLD Revisited

Before discussing CLLD, let us return to our current analysis and see how it fits in with the differences we observed between CLD and HTLD. In this section I will slightly modify the analysis, or better: I will extend it to capture an interesting property of HTLD and CLD, namely that the two can co-occur, as long as there is only one CLDed XP and any HT precedes it.
4.3.3.1 Structural implications for HTLD

Let us then revisit the distinction between CLD and HTLD. If CLD is derived by movement of the LDed element and HTLD is not, and given that the analysis of CLD as presented above has some merit and approximates (55a), (55b,c) might be expected to be the equivalents of HTLD I and II.

(55)  

a.  

\[
\begin{array}{c}
CP\;XP\;[\text{TopP}\;X\;d-RP\;V\;\ldots\;[TP\;\ldots\;X\;\ldots\;[VP\;\ldots\;XY\ldots]]]
\end{array}
\]

b.  

\[
\begin{array}{c}
CP\;XP\;[\text{TopP}\;d-/p-RP\;V\;\ldots\;[TP\;\ldots\;R\;\ldots\;[VP\;\ldots\;RP\ldots]]]
\end{array}
\]

c.  

\[
\begin{array}{c}
CP\;XP\;[\text{TopP}\;ZP\;V\;\ldots\;[TP\;\ldots\;d-/p-RP\;\ldots\;[VP\;\ldots\;RP\ldots]]]
\end{array}
\]

If movement of the LDed element takes place in CLD only, but if all other constellations are equivalent, we could expect the HTLDed element to be base-generated in the same position that the CLDed element moves to, namely SpecCP. The only (relevant) element that moves in HTLD is the RP which may either be a \(d\)- or a \(p\)-pronoun and as (55b,c) show it may be high or low. On a par with CLD, the high RP in HTLD I topicalizes, while the low RP in HTLD II moves to the regular argument position in the \(\phi\)-domain, and another XP occupies the “first” position of the V2 structure of the clause.

4.3.3.2 CLD and HTLD in interaction: Specifiers and adjuncts, once again!

We have seen above that LD and topicalization may interact in both English and German. (33) above was an example of HTLD and topicalization, repeated here as (56):

(56)  

a.  

Dieser Kandidat, einen Arschtritt sollte man dem/ihm geben.
\(\text{this.NOM candidate a kick-in-the-ass should one RP.ACC/him give}
\)

‘This candidate, a kick in the ass, one should give him.’

b.  

* Einen Arschtritt dieser Kandidat, sollte man dem/ihm geben.
\(\text{a kick-in-the-ass this.NOM candidate should one RP.ACC/himgive}
\)

‘*A kick in the ass, this candidate, one should give him.’  
\((\text{German})\)
Our account so far can deal with this example: given that the HT sits in CP, nothing bars a topic from moving to TopP and the two can co-occur—as long as the HT precedes the HT. As mentioned already, the only way to test co-occurrence with topicalization is with HTLD II, as a high RP would fight over TopP with a potential topic.  

What we could test, though, is co-occurrence of HTLD and CLD. If HTLDed and CLDed XP occupy the same CP position—argued to be unique above already—, this would be predicted to be quite ungrammatical. Interestingly, this prediction is not borne out. Consider the following data:

(57) Der Alex, den Wagen, den hat seine Mutter ihm gestern geschenkt.

‘Alex, the car, his mother gave to him yesterday (as a present).’  (German)

Unless we take a more liberal stance on what makes a topic. In previous work I have analysed material moved over the subject as targeting (recursive) TopP (see the next chapter for more). If this is so, there is no reason why CLD or HTLD I would not interact with “topicalization” (or an element that targets TopP), and indeed, under this approach the following would be predicted to be grammatical, which they are:

(i) a. Diesen Roman, den hat der Maria der Peter gegeben.
   ‘This novel, Maria, Peter gave it to.’

   b. Seinem Vater, dem will seine neue Freundin jeder vorstellen.
      ‘His father, his new girlfriend, everyone wants to introduce to.’  (German)

Interestingly, turning (ib) into HTLD yields the following binding relation:

(ii) Sein Vater, seine neue Freundin will jede Junge dem/ihm vorstellen.
     ‘His father, his new girlfriend, every boy wants to introduce to.’  (German)

In (ii), seine ‘his’ can be bound by jede Junge ‘every boy’, while the pronoun inside the HT can only refer to an aforementioned, different father. If topics can reconstruct and the relevant element in (i-ii) really is a topic, these binding (im)possibilities are predicted.
We can thus construe cases which contain more than one LDed element. Observe, though, that in these instances the two LDed elements become acceptable only if the lower one is immediately followed by its resumption:

c. * Der Alex, den Wagen, ihm hat den seine Mutter gestern geschenkt.
d. * Den Wagen, der Alex, ihm hat den seine Mutter gestern geschenkt.

‘Alex, the car, his mother gave to him yesterday (as a present).’ (German)

(58) shows various permutations of the grammatical version. The Case-marked XP must follow the nominativus pendens, and it must in turn be followed by its corresponding RP, the d-pronoun. We could present more permutations, but the gist is clear: HTLD and CLD may co-occur, as long as the HT precedes the CLDed XP.

If this is indeed the case, we have an argument that the HT sits in a different position from the CLDed XP. Holding fast to the assumption that SpecCP is unique, the position of HT could be only one of two: a specifier of an even higher projection or adjoined position to CP. Given that CP is supposedly the highest projection of the clause, the first option is unlikely. I will assume the second option and show what mileage we might get out of it. If the modification of X’-relations and checking configurations presented in section 3.2.4 is on the right track, analysing the HT as sitting in AdjCP and the CLDed XP in SpecCP immediately makes a number of predictions. In particular, (i) as a specifier, the CLDed XP must be unique, (ii) as an adjunct, the HT need not be unique, (iii) the HT can only be base-generated, and (iv) the HT must precede the CLDed XP.

The above data support prediction (iv), and the entire discussion so far has aimed to show that (i) and (iii) do indeed hold of LD constructions. This leaves us with predic-
tion (ii). While stylistically rather marked and certainly subject to the right context,\(^{25}\) we can in fact stack HTs, with or without a CLDed XP—as long as that XP follows the HTs.

\[(59)\]
\begin{align*}
\text{a.} & \quad \text{Der Alex, der Wagen, seine Mutter, gestern hat } \text{sie ihm den, geschenkt.} \\
\text{b.} & \quad \text{Der Alex, seine Mutter, den Wagen, den hat } \text{sie ihm gestern geschenkt.}
\end{align*}

‘Alex, his mother, the car, yesterday she gave (it) to him.’ \((\text{German})\)

\((57a)\) contains three HTs, \((57b)\) two plus one CLDed XP. The indices indicate the resumption relations.\(^{26}\) Naturally, any other XP may be CLDed, such as Alex:\(^{27}\)

\[(60)\]
\begin{align*}
\text{a.} & \quad \text{Seine Mutter, den Wagen, dem Alex, dem hat sie den gestern geschenkt.} \\
\text{b.} & \quad \text{Den Wagen, seine Mutter, dem Alex, dem hat sie ihn gestern geschenkt.}
\end{align*}

‘His mother, the car, Alex, she gave it to yesterday.’ \((\text{German})\)

‘The car, his mother, Alex, she gave it to yesterday.’ \((\text{German})\)

To control for prediction (iii), and the others, properly, compare the following sets of data, where all intended CLDed XPs are marked in boldface:

\[(61)\]
\begin{align*}
\text{a.} & \quad * \text{ Der Alex, den Wagen, seine Mutter, den hat sie ihm gestern geschenkt.} \\
\text{b.} & \quad * \text{ Dem Alex, der Wagen, seine Mutter, dem hat sie ihn gestern geschenkt.}
\end{align*}

*‘Alex, the car, his mother, she gave to him yesterday.’ \((\text{German})\)

\(^{25}\) Cinque (1990) cites examples from Italian which contain a number of CLLDed XPs, and so does Rizzi (1997). The general property that CLLDed XPs may be stacked is well known and attested, but I am not sure how common these constructions are in actual use. If it turns out that they, too, require special context, nothing might seem too bad about the examples from German (which, to my knowledge, have not been cited in the literature).

\(^{26}\) I will not discuss the role of word order or choice of RP in these cases, both of which follow from other aspects of German grammar (animacy, topicality, scrambling etc.).

\(^{27}\) As one might suspect, in these cases the mother cannot be Alex’s mother (see section 4.2.5.2.7). What is interesting, though, is that the parallels in English are not too bad either—given that we translate CLD as topicalization. Thus, English can have multiple HTs, followed by topicalization, but only in this order.
(62)  a. * _Dem Alex, den Wagen, seine Mutter, dem, hat den, sie, gestern geschenkt._
    b. * _Seine Mutter, _dem Alex, _den Wagen, dem, hat den, sie, gestern geschenkt._

    ‘Alex, the car, his mother, she gave yesterday.’  (German)

We can thus modify (55) roughly as in (63):

(63)  a.  \[ [\text{CP} \text{XP} C [\text{TopP} \text{XP} \supset d-\text{RP} V \ldots [\text{TP} \ldots \text{XP} \ldots [\text{vP} \ldots (\text{XP}) \ldots]]]] \]
    b.  \[ [\text{CP} \text{YP} [\text{CP} \text{XP} [\text{CP} C [\text{TopP} d-/p-\text{RP} V \ldots [\text{TP} \ldots \text{RP} \ldots [\text{vP} \ldots (\text{RP}) \ldots]]]]]] \]
    c.  \[ [\text{CP} \text{YP} [\text{CP} \text{XP} [\text{CP} C [\text{TopP} ZP V \ldots [\text{TP} \ldots d-/p-\text{RP} \ldots [\text{vP} \ldots (\text{RP}) \ldots]]]]]] \]

We can understand (63) as follows. In CLD, (63a), the XP originates below the \( \omega \)-domain (either in the \( \theta \)-position, if an argument, or somewhere within the \( \phi \)-domain) and moves to TopP, then on to CP. The latter movement would violate Exclusivity and as such can only be licit if the lower copy spells out. In these cases, the Copy Spell Out filler is a \( d \)-pronoun, matching in Case and \( \phi \)-features (if any, otherwise by default; see below). In HTLD I, shown in (63b), the element that undergoes movement is the RP itself—whether in form of a demonstrative or a personal pronoun will not be our concern here. It also comes from its \( \theta \)-position, if it is an argument, otherwise from somewhere within the \( \phi \)-domain, and it topicalizes, resulting in HTLD with high RP. The HT is base-generated in CP-adjoined position, and if there is more than one, they all are (indicated by XP, YP). HTLD II differs from HTLD I only in the last movement: the low RP does not undergo topicalization.

(63) captures all the properties of CLD and HTLD we have seen so far, in a rather straightforward manner. The structure proposed makes a number of predictions, some of which we have seen already. We will return to more discussion after CLLD has been analysed in a similar way. First, let us consider some pertinent issues which we have not yet really looked at in any depth.
4.3.4 LD and Selection

I would now like to consider Hoekstra’s (1998) attack on the movement-qua-Copy-Spell-Out approach to CLD, which is based on apparently relevant constructions from Dutch. Recall that Case-matching plays a big role in deciding whether CLD or HTLD takes place, and Dutch does not mark Case morphologically; a second indication for CLD is that the RP, exclusively of the d-variety, occurs in high, topic position.

Firstly, apparently, (64) holds for Dutch with the judgements indicated:28

(64) a. Boeken lezen, dat doe ik niet.
books read RP do I not
‘Read books, I don’t.’
b. * Ik doe niet boeken lezen.
I do not books read
‘I don’t read books.’
c. Ik doe dat niet.
I do RP not
‘I don’t do that.’

(Dutch, Hoekstra & Zwart 1998)

What this paradigm is supposed to show is that while the RP may be directly selected (as in (64c)), the potential LDed XP may not (cf. the ungrammatical (64b)), even though the corresponding LD structure is fine (i.e. (64a)). There are two immediate replies from the perspective of German for which, after all, the current proposal is supposed to apply. For one, the equivalent of (64b) is grammatical in German. While possibly not acceptable in “standard” German, it is a construction that is often used colloquially.

28 I copied Hoekstra’s example (his (15)) verbatim except that I glossed dat as RP, not as T(topic)P(ronoun), as he does; I also italicize LDed element and RP, as I have done throughout.
(65)  a. Ich tu nicht (gerne) Bücher lesen.
    I do not (with-pleasure) books read
    ‘I don’t read books (with pleasure).’

    b. Mein Freunde tun nicht viel arbeiten, die gehen lieber feiern.
    my friends do not much work they go preferably party
    ‘My friends don’t like to work very much, they rather go partying.’

(66)  a. Bücher lesen tu ich nicht (gerne).
    books read do I not (with-pleasure)
    ‘Read books, I do not (with pleasure).’

    b. Viel arbeiten tun meine Freunde ja nicht gerade.
    much work do my friends PRT not exactly
    ‘Work much, my friends don’t exactly.’

Moreover, the constituent in question may also be topicalized in German:

If, all other things being equal, topicalization is derived by movement, we would
have to find an alternative account for (66) under the assumption that the infinitival com-
plement could not be “selected” by the verb *tun/doen* ‘do’. Moreover, as Jan-Wouter
Zwart (p.c.) informs me, (66) is as well-formed as its Dutch counterpart. Also, he reports,
the “base” structure (65) is frequently used by Dutch speakers as well, especially in the
Limburg area and in other Southern dialects.

We have seen other examples where a bigger constituent than DP or PP is LDed
(cf. (20) above). Haider (1990) refers to it as “VP-left dislocation,” on analogy with so-
called “VP topicalization” in German, which, in turn, has recently been investigated in a
bigger picture as “remnant movement” by Müller (1998). It is not so clear that the fronted
constituent really is VP, or possibly something larger. For example, Hoekstra (1998), fol-
lowing Hoekstra & Zwart (1994), refers to it as “infinitival IP-constituent” for an obvious
reason: if German is indeed head-initial throughout and all arguments obligatorily move
out of their base positions, the constituent [*Bücher lesen*] ‘read books’ could not be a VP:
not only is the object Case-marked, but it would appear on the wrong side! This is even clearer with ditransitive constructions, where the only relevant constituent that could be fronted would have to be \[\textit{der Maria das Buch geben}\] ‘give Maria the book’. Naturally, one solution is to give in and assume the German (and Dutch) VP to be head-final. But as this is not a major topic in this work, I dispense with such considerations and follow the Zwartian line further (cf. Zwart 1993, 1994, 1997a) assuming that it is not unreasonable to take the constituent in these cases to be something bigger, such as a TP.

Hoekstra’s second argument against a movement analysis also involves “selection,” but this time it concerns selection of PP over DP in Frisian (his example (16)):

(67) a. \textit{Reduzem dër woe ik wol wenje.}  
\textit{Reduzem RP would I sure live}  
‘Reduzem, there I sure would like to live.’

b. * \textit{Ik woe wol Reduzem wenje.}  
\textit{I would sure Reduzem live}  
‘I sure would like to live *(in) Reduzem.’

c. \textit{Ik woe wol dër wenje.}  
\textit{I would sure RP live}  
‘I sure would like to live there.’     \textit{(Frisian; from J. Hoesktra 1995)}

Here I agree with Hoekstra: it is unlikely that the DP \textit{Reduzem} is LDed by movement. On the other hand, nothing rules out the LDed element in (67a) being base-generated, i.e. being a HT. Hoekstra considers this option but, on the basis of Dutch and Frisian again, rules it out immediately on the same grounds as above. Consider:
In both cases, the LDed XP does not satisfy the “selection criteria,” which I would like to dismiss as above. The two sentences in (68) are relevant, though, in that they show connectivity effects. The reciprocal within the LDed XP in (68a) is coreferent with the lower pronominal subject, and the pronoun inside the LDed in (68b) can be interpreted as a bound pronominal, referring to elkenien ‘everybody’.

Consider the German equivalents of (68):

(69) a.  *Einander helfen*, *das* tun *sie* hier nicht.
   each-other help RP do they here not
   ‘Help each other, they don’t do (that) here.’

b.  *An seinem Geburtsort da* würde *jeder* wohl leben wollen.
   at his birthplace RP would everybody sure live want
   ‘His birthplace, everybody sure would like to live.’ *(German)*

(69a) works as expected: the LDed constituent is moved, spells out its copy as the RP and binding of the reciprocal can take place. Moreover, if it is true of German also that only DPs can be HTs, the fronted XP must be a CLDed constituent. (69b), as the German version of (68b), is a little bit trickier. I translated it as a PP, and everything is fine, for the obvious reasons. However, were we to translate literally, we would have to employ a DP. Given that Frisian does not have overt Case-marking, we would face the problem which Case to employ: if we use nominative, the construction could be HTLD and a bound pronominal reading would not be expected; if we use another Case, we face the problem of selection. Thus, if (67a) is a HTLD structure, because of the nominative
LDed XP and as such would not violate selection, playing the same trick on (68b) would rule in a bound pronominal reading for HTs. Needless to say, the only possible structure in German would be with a nominative, but here the connectivity effect gets lost.

Note that the connectivity property of CLD would be the ideal candidate for comparing German and Dutch (and Frisian) LD constructions. Due to the absence of overt Case-marking, it would be hard to use the Case-matching criterion. Motivating the Copy Spell Out analysis on the basis of connectivity seems a more plausible way to go about it (see also Vat 1981, van Haaften, Smits and Vat 1983 for ample data). This does not help us with Hoekstra’s problem, which I take to concern only (68b).

Naturally, to say something definite, LD constructions in Frisian (and Dutch) would have to be scrutinized, similar to what we did here for German. I cannot do that here and leave the discussion where it stands, possibly unresolved, but at least for German extremely promising and clear. How far the details of the analysis of CLD that I propose extend to Dutch and Frisian remains to be seen. 29

29 One possibility might be that (68b) contains a preposition at some point, which got deleted in the course of the derivation, or was empty to begin with. I suspect that Hoekstra illustrated his point with Frisian for a good reason. Dutch might not make this option available either. If that is the case, we would really have to look some more at Frisian, something I cannot do here. Note that the three languages surely differ in many respects. One property that distinguishes Dutch from German, for example, is that prepositions can be stranded under certain circumstances, namely when the extracted element bears locative morphology, a property ascribed to the feature [+R] by van Riemsdijk (1978). Consider (i):

(i) \textit{Jan, ??(daar/\*die) heb ik een hekel an.}  
\textit{Jan RP have I a disgust for}  
‘Jan, I have disgust for.’ \textit{(Dutch, adapted from Zwart 1997c)}

The complement of the preposition \textit{Jan} cannot be extracted without resumption, and the resumptive must be the R-form, not the regular RP. I take the RP \textit{d\`er ‘there’} from above to be the Frisian R-equivalent. Continuing this pure speculation, maybe this could be a clue as to why not only the DP-complement of the preposition can be CLDed, but also the preposition deleted. But, as I said, this is only speculation at this point. (Note that if valid, this line of reasoning could also account for Hoekstra’s first case, viz. (67a) and understand \textit{Reduzem} not as a HT but as a CLDed XP, followed by P-deletion.)
4.3.5 Summary

To sum up this section, we have seen that the distinction between the two types of LD in German still holds, and should hold, characterized as we have throughout: obligatory Case-matching and high RP of the d-variety for CLD and at best optional Case-matching, high or low RP in either d- or p-form for HTLD. On empirical grounds, I propose that HTLDed elements are adjoined to CP, while CLDed elements move to SpecCP. We can take (70), adopted from (63) above, as a final structural representation for the two, giving us the only possible ordering relation HTLDed XP followed by CLDed XP:

(70)  a. \([CP \, XP \, C \, [TopP \, XP \, \supseteq d-RP \, V \, \ldots \, [TP \, \ldots \, XP \, \ldots \, [vP \, \ldots \, (XP) \, \ldots]]]]\)
     b. \([CP \, YP \, [CP \, XP \, [CP \, C \, [TopP \, d-/p-RP \, V \, \ldots \, [TP \, \ldots \, RP \, \ldots \, [vP \, \ldots \, (RP) \, \ldots]]]]]]\)
     c. \([CP \, YP \, [CP \, XP \, [CP \, C \, [TopP \, ZP \, V \, \ldots \, [TP \, \ldots \, d-/p-RP \, \ldots \, [vP \, \ldots \, (RP) \, \ldots]]]]]]\)

We have also seen some of the predictions this approach makes turn out to be borne out. Next, I apply the gist of the analysis to LD structures with resumptive clitics.

4.4 A Derivational Approach to Clitic Left Dislocation

CLLD has often been likened to CLD (or the other way around), and we have seen initial arguments above. In this section, I present a collection of data from CLLD in Greek, suggesting that the CLLDed XP is the element that undergoes movement. I then consider the syntactic environment of the clitic, suggesting that it enters the derivation in the \(\phi\)-domain. I then propose an analysis in terms of Copy Spell Out very much like the one for German CLD, but Exclusivity is threatened in another Prolific Domain, and the derivation involves spelling out a copy of the XP in the form of a clitic.
4.4.1 Movement vs. Construal in CLLD

The same issue can be raised for CLLD as for CLD. Is there any movement involved, and if so, what moves where when? As mentioned in section 4.2.5.3, the same reasons that motivated a movement analysis for CLD have been put forward to analyse CLLD in similar terms (cf. Cinque 1977). In this sub-section, I am going to present the relevant data from Greek (from Anagnostopoulou 1997), and I also show how Greek CLLD differs from German CLD, an issue further investigated in section 4.4.2.

(71) seems to support the movement assumption. The XP and the CL are both appropriately Case-marked and selected by the relevant predicate (*agapo ‘love’). (71a) shows unboundedness and reconstruction effects (bound pronominal reading, as indicated by underlining), while (71b) suggests island-sensitivity (CNPC with a relative clause). (These sentences are modeled after data from Iatridou 1990, Anagnostopoulou 1997.)

(71) a. *Tin mitera tu skeftika oti o Petros ipe oti the.ACC mother his.GEN thought.1SG that the Peter said.3SG that kathenas tin agapai. everyone CL.ACC loves

‘His mother, I thought that Peter said that everyone loves.’

b. *Tin mitera tu sinantisa tin kopela pu ipe oti the.ACC mother his.GEN met.1SG the girl who said.3SG that kathenas tin agapai. everyone CL.ACC loves

*‘His mother, I met the girl who said that everyone loves.’ (Greek)

In all these respects CLLD differs from HTLD: HTLD does not involve obligatory Case-matching in Greek, and it is not unbounded nor sensitive to islands, and certainly does not allow reconstruction (or show any other connectivity effect).
So far, CLLD seems to match CLD pretty well. There are a number of differences, though, as alluded to above. Greek CLLD may violate some islands, what Anagnostopoulou calls “selective islands” (Wh-islands, basically). Unlike CLD, CLLD may occur in any embedded context and more than one XP may be CLLDed. Contrast:

(72) a. \textit{To forema} dhen ksero \textit{pu} na \textit{to} valo. \\
the.ACC dress not know.1SG where to CL.ACC put \\
*‘The dress, I don’t know where to put.’ \\
\textit{(Greek)}

b. * Den Rock weiß ich nicht \textit{den wohin} ich legen soll. \\
the.ACC dress know I not RP.ACC where I put shall \\
*‘The dress, I don’t know where I should put.’ \\
\textit{(German)}

(73) a. … oti \textit{ton Janni} dhen \textit{ton} ksero. \\
that the.ACC John not CL.ACC know.1SG \\
‘… that John, I don’t know.’ \\
\textit{(Greek)}

b. * … daß \textit{den Martin} \textit{den} ich nicht kenne. \\
that the.ACC Martin RP.ACC I not know \\
‘… that Martin, I don’t know.’ \\
\textit{(German)}

(74) a. \textit{Tis Marias} to vivlio \textit{tis} to edosa. \\
the.DAT Maria the.ACC book CL.DAT CL.ACC gave.1SG \\
??‘Mary, the book, I gave to.’ \\
\textit{(Greek)}

b. * \textit{Der Maria, das Buch, der das} gab ich. \\
the.DAT Maria the.ACC book RP.DAT RP.ACC gave I \\
??‘Mary, the book, I gave to.’ \\
\textit{(German)}

Coupled with the other two observations mentioned above also—clitic rather than \textit{d}-pronoun and clause-internal rather than left-peripheral resumption—, we are facing five discrepancies (see also Anagnostopoulou 1997: 159f.). These are listed in (75):

---

\textsuperscript{30} One aspect in which Greek CLLD differs not only from CLD, but also from CLLD in many other languages, is that only DPs can be CLLDed (see Cinque 1990 on Italian).
Differences between CLD and CLLD

i. resumptive is a stressable $d$-pronoun in CLD vs. unstressed clitic in CLLD
ii. RP in CLD ($d$-pronoun) is in $\omega$-domain, RP in CLLD (clitic) in $\phi$-domain
iii. CLLDed XPs can be stacked, CLDed XPs cannot
iv. CLLD can be freely embedded, CLD cannot
v. CLLD can span across selective islands

4.4.2 CLLD vs. CLD: A Difference in Prolific Domains

If we want to evoke a similar analysis as for CLD, namely in terms of Copy Spell Out driven by Exclusivity, these differences will have to be accounted for. The first two differences can be accounted for relatively easily: we know that Copy Spell Out makes pronominal elements available, and that the choice might differ should not come as a surprise. I suggest for now to stipulate that CLLD involves spelling out an otherwise illicit non-distinct copy with a clitic and CLD with a $d$-pronoun. We might be able to say something deeper soon. Naturally, the RP is not even expected to be stressable if it is a clitic, unlike the $d$-pronoun which is, after all, a maximal phrase. Regarding the occurrence of the resumptive element, the obvious way to go would be considering whether the timing of Copy Spell Out could be different. I suggest exactly that: CLD involves an illicit movement within the $\omega$-domain, while CLLD involves the same one Prolific Domain lower, inside the $\phi$-domain.

Basically, this allows for an initial, rough adaptation of (54) for CLD in terms of (76) for CLLD, a hypothesis which we will consider next:

\[
\begin{align*}
(76) \quad & [_{CP}\, XP\, C\, [_{Top}\, Top\, [_{TP}\, \ldots\, XP\, \ldots\, XP\, \overset{\text{CL-V}}{\ldots}\, [_{vP}\, \ldots\, XP\, \ldots]]]]
\end{align*}
\]
This rough structure needs elaboration on (at least) two counts. First, if Copy Spell Out occurs inside the $\phi$-domain, what are the two $\phi\Delta$-internal positions (here within TP)? One is presumably close to the verb, as indicated here, for reasons we will look at soon. Second, if the final position of XP is the same as in CLD, does it pass through TopP? If so, why does the lower copy (in topic position) not spell out? But perhaps, the two positions are different. In that case we would expect the difference to be testable, or at least have some empirical basis. I suggest that it does. I will tackle the second question first and address the first question afterwards, in section 4.4.3.

Clearly, the CLLDed XP cannot move to SpecCP via SpecTopP, otherwise it should face the same fate as the German CLDed XP and be forced to spell out the lower copy—this movement would take place within the $\omega$-domain. We have noted several times already that the LDed XP in all constructions is “topic-like,” without being more specific. Rather than become more specific, I would like to capitalize on this deliberately vague formulation. Presumably, a topic becomes a topic when it moves to TopP. However, if in German CLD the $d$-pronoun is in TopP and hence the syntactic topic, what does the LDed XP do in CP? One answer could be that the two structures (topicalization vs. CLD of XP) simply differ in some respect, thus motivating two ways of expressing two slightly different state of affairs. The notion “contrastive” comes in handy, again. Recall that only CLD must be contrastive; there is no such presupposition on simple topicalization. We can thus stick to the assumption that the syntactic topic is indeed the $d$-pronoun and its referent, in slightly higher position, serves some other function, possibly facilitating the contrastive function.

This said, CLLD is not exclusively contrastive either. In fact, it presumably is not contrastive at all: one aspect of CLD is the tonic resumptive element. The $d$-pronoun, in

---

31 Call this the “syntactic topic,” driven by a formal feature; this is the only kind of topic we are interested in here. Of course, something can be a topic without moving to a specific position (see e.g. Reinhart 1981, Vallduvi 1992, Erteschik-Shir 1998).
topic position, may be stressed, supporting the contrastive nature of the entire construction. The clitic, being phonetically extremely reduced, cannot be stressed or emphasized in any way. Moreover, the clitic in CLLD does not sit in a topic position (see above, also next sub-section). This suggests that the topic-like nature in CLLD comes from the XP itself. Indeed, it has been argued that many languages that have CLLD use that strategy to express topicalization, some languages even exclusively—Greek is one of them (Tsimpli 1995).

If CLLD is the strategy to express topicalization, it is not unreasonable to assume that the CLLDed XP sits in SpecTopP—a different position from the CLDed XP. Moreover, as is well known, many languages allow for multiple topics; even English, in some circumstances (cf. 16b). Employing Rizzi’s (1997) analysis, among others, we could then account for this state of affairs by invoking an additional TopP—basically, as many topic projections as there are (syntactically displaced) topics in a sentence.

It should be clear now where this is going. If the CLLDed XP sits in SpecTopP, rather than SpecCP, and if TopP is “recursive” (unlike CP), we would actually expect that more than one CLLDed XP may be present in a clause, but not more than one CLDed XP, regardless of how many topic projections the German clause might make available. Consider the following visualization of this hypothesis:

\[(77)\]
\[\begin{align*}
(77a) & \quad [_{\text{CP}} \text{XP} [_{\text{TopP}} \text{XP} \supset \text{RP} V [_{\text{TP}} \ldots \text{XP} \ldots]]] \quad (CLD) \\
(77b) & \quad [_{\text{TopP}} \text{YP} [_{\text{TopP}} \text{XP} [_{\text{TP}} \ldots \text{YP} \ldots \text{XP} [_{\text{YP}} \supset \text{CL} \text{XP} \supset \text{CL}] V \ldots]]]] \quad (CLLD)
\end{align*}\]

(77a) is the well-known structure for German CLD, allowing for exactly what we see: one CLDed XP and one corresponding RP. (77b) is a first pass for CLLD in Greek: the CLLDed elements occupy topic positions, and all resumption takes place low. This very rough sketch will be considerably sharpened in the next sub-section, where I also address the position and introduction of the clitics more accurately than here.
4.4.3 Copy Spell Out in the $\phi$-Domain

Let us pick up the discussion from (77b). This structure indicates the occurrence of any CLLDed elements within the $\omega$-domain and of the related clitics within the $\phi$-domain. It also illustrates clustering all clitic occurrences around the verb. One property of clitics in Greek (and Italian, Spanish etc.) is their verb-relatedness. That is to say, clitics attach to the verb and form a complex head that might or might not undergo subsequent movement.

The relevance to CLLD should now be clear. What if the clitic enters the derivation (in form of a spelled out copy) in its canonical clitic position? This derivational approach to CLLD has been considered in the literature (for critical discussion, see Cinque 1990, Anagnostopoulou 1997). However, the way the issue was presented there need not be the only way. Introducing the resumptive clitic in the course of the derivation might have a strong smell of a clitic doubling analysis to it, but need not. As the criticism against a clitic doubling approach to CLLD is well taken, I will suggest an alternative movement analysis for CLLD. Let us take a look at (some of) the issues involved.

Cinque considers two movement approaches to capture CLLD, but discards both of them in favour of a base-generation analysis involving construal. One possible analysis considers CLLD essentially as a variant of clitic doubling. We can disregard a clitic doubling analysis on the same grounds as Cinque (1990: 60f.) does, though I will return to this issue once our alternative is on the table. The main objection is that while some languages, such as Greek or Spanish (cf. (78)), exhibit both CLLD and clitic doubling, Italian only has the former. That is, a clitic never doubles a full nominal phrase, regardless of whether the DP has moved (shown in (79)). Restricting the clitic doubling approach to Italian CLLD only seems also ill motivated, given that clitic doubled nominals across the
clitic doubling Romance languages are always introduced by a preposition, while CLLDed DPs are not (as in (80)).

(78) a. Ta vrika, ta klidia.
   CL found.1SG the keys
   lit. ‘I found them, the keys.’ (Greek; Villalba 2000: 135)

b. La vi muy cambiada, a Maria.
   CL saw.1SG very changed to Maria
   lit. ‘I saw her very changed, Maria.’ (Spanish; Villalba 2000: 134)

(79) a. * (A) chi lo conoscete?
   (to) who CL know.2SG
   lit. ‘Who do you know him?’

b. * Lo conosciamo (a) Gianni.
   CL know.1PL (to) Gianni
   lit. ‘We know him, Gianni.’ (Italian; Cinque 1990: 60)

(80) Non so se il vino lo volete adesso o dopo.
   not know.1SG if the wine CL want.2SG now or later
   lit. ‘I don’t know whether the wine, you want it now or later.’
   (Italian; Cinque 1990: 61)

The other movement approach to CLLD Cinque considers treats the clitic RP as a spelled out form of a trace (or copy) of the CLLDed element. Obviously, this is similar to the analysis I will provide, but it differs in one crucial respect. The only argument Cinque presents against this option is the fact that the clitic RP does not license parasitic gaps. Given the conception of spelling out the trace he has in mind, this is a powerful argument. In the present context, however, it dissolves. Cinque’s line of reasoning is that if CLLD were the result of topicalization of the CLLDed XP which spells out its copy as the clitic RP, the clitic would be an A’-bound variable and as such expected to license a parasitic gap. This does not turn out to be the case, though.

Consider (81):
If the clitic were really an A’-variable, Cinque’s argument goes through. However, as (77b) and the preliminary discussion so far imply, we do not take the clitic RP to be an A’-bound variable. That is, we do not consider the copy left behind by topicalization to be the relevant copy to spell out. In the current framework of Prolific Domains, there is no formal reason why this copy should spell out—the one argument for Copy Spell Out we have considered so far is a repair strategy to save Exclusivity: not movement across Domain boundaries, but within a Domain undergoes Spell Out.

Indeed, take Copy Spell Out of the CLLDed XP to take place entirely within the $\phi$-domain, which would make the spelled out clitic at most a A-bound element, not an A’-element. Thus, Cinque’s objection to a Copy Spell Out approach, well taken as it is, does not concern our proposal. I then take the stakes to be level: if CLLD displays movement diagnostics and if there is a movement analysis which has not been considered yet, we should explore this first before retreating to a base-generation analysis. I lay out this alternative next.

The basis of the Copy Spell Out approach to CLLD is a particular analysis for cliticization. I would like to suggest that argument clitics arise from DP-movement to the relevant $\phi$-head. I take a clitic to be a DP of sorts—basically, a “neutralized maximal projection” (Uriagereka 1995a: 113), an element that can behave like a head or a like a maximal projection (see, among others, Chomsky 1993). Consider the following:
(82)  a.  Ton ksero.
    CL.ACC  know.1SG
    ‘I know him.’  \((\text{Greek})\)
  
    b.  \(\left[\text{AgrP} \left[\text{DP ton-D\text{-}}\text{ksero\text{-}}\text{V\text{-}}\text{Agr} \left[\text{vP pro ksero-V \left[\text{VP ksero-V DP}\right]\}}\right]\right]\right] \]

    Take the clitic to be generated as a DP inside the \(\theta\)-domain. The verb successively cyclically moves to Agr, at which point the DP can move, and it does so by adjoining to Agr, forming a complex head structure (boldfaced). As this is movement of a maximal phrase, no locality restrictions (in the form of the HMC) are violated, and the D-head can manifest as the clitic. As such it forms a cluster with the verb and undergoes any subsequent movement with it (needless to say, subject to the HMC).\(^{32}\)

    In CLLD, the derivation is similar, but with a twist. Rather than a regular clitic like (83a), the initial DP looks as in (83b):

(83)  a.  \([\text{DP D'} [\text{ton-D}]]\)
  
    b.  \([\text{DP D'} [\text{ton-D [NP Janni]]}]\)

    The entire DP ton Janni ‘the John’ adjoins to the verb in the Agr-head. Given our checking relations, the D-head can check anything it needs to with Agr, such as \(\phi\)-features (by virtue of Agr Containing D\(^{0}\); cf. section 3.2.4.3). But unlike a clitic, as in (83a), this DP contains more than a simple head, namely an NP. Let us assume that the entire DP then needs to enter a checking relation, not only D\(^{0}\). In this case, DP moves from the complex Agr-head to the Agr-specifier—indeed a very local, if not a paradig-

\(^{32}\)This is a variation of the original XP-movement analysis of clitics (Kayne 1989), later modified by Sportiche (1995), among others. It differs in that I do not assume XP-movement to a specifier position where the clitic is the result of a “strong” agreement relation or a base-generated Agr-head. I will address some technical issues momentarily.
matic example of anti-local, movement. The result of this movement is the necessity to spell out the lower copy of DP.

A more accurate derivation up to this point is represented in (84); cf. (82):

\[ (84) \quad \begin{align*}
\text{a. } & [\text{Agr} [\text{DP-V-v-Agr}] \text{-Agr} [_{\text{VP}} \text{pro V-v [VP V DP]]}] \\
\text{b. } & [\text{Agrp DP} [\text{Agr} [\text{DP-V-v-Agr}] \text{-Agr vP]}}] \\
\text{c. } & [\text{Agrp DP} [\text{Agr} [\text{DP ☰ CL-V-v-Agr}] \text{-Agr vP]}}]
\end{align*} \]

Note that we are still dealing with two non-distinct copies. Arguably, DP in Agr has different properties from the original DP in CompVP in (84a).\(^{33}\) That is irrelevant, however, as the two copies are in distinct Prolific Domains. What is relevant is that DP, the interesting object in the phrase marker at the point (84b), is identical to the one in (84c), when it has moved to SpecAgrP—if nothing else, by virtue of this movement being an instance of substitution.

Consider the relevant derivational steps under the arboreal telescope:

---

\(^{33}\) Recall the discussion from section 2.3.2 on head movement and distinctness.
(85) a. 

```
(85) a. 

(85) a. 
```

b. 

```
(85) b. 
```

```
(85) b. 
```
The initial $V'$ in (85a) consists of $V$ and DP, the result of standard Merge. The $V$-head undergoes successive-cyclic head movement, yielding (85b), and at that point DP moves from CompVP to adjoin to the complex Agr-head, with $V$-$v$ in Agr. In a simple clitic case, DP would consist of $D^0$ only and $D^0$ could enter into a checking relation with $Agr^{0\text{max}}$ (via Contain, the Natural Relation for Head-Head feature checking); cf. (82). However, this DP is more complex, and while the relevant feature on D is able to get checked, the categorial DP-feature(s) are not—and we assume that full phrases, rather than heads, check $\phi$-features and get Case-marked. As a consequence, DP moves on, to SpecAgrP, sketched in (85c). This movement is a regular application of Copy and Merge: it copies the entire DP as detailed here (i.e. the circled object) and merges it with the next possible position, the specifier of that projection. As a result, two non-distinct copies of one object in the phrase marker are in an anti-local relationship, violating the CDE. The constellation can be saved by changing the PF-matrix of the lower copy, and as a result
*ton Janni* ‘the Janni’ can stay in SpecAgrP, while the copy inside the complex head gets spelled out. Let us assume, for lack of more accurate procedure, that such configurations warrant Copy Spell Out in terms of a clitic (see section 4.5 for a brief discussion).

One aspect of this analysis is that a full DP can adjoin to a head. Note that if the DP contains only the clitic, this should not pose much of a problem (cf. (82)), as it is a neutralized maximal projection, in Uriagereka’s (1995a) terms. It is a head and a maximal projection at the same time. This is arguably not the case here. Rather the DP seems to be a full DP. We can appeal to well-formedness at PF to account for these cases. The theory does not prohibit full projections to adjoin to heads on categorial grounds; after all, the categorial status of a clitic DP is still a DP, regardless of its internal structure (which is not relevant for merging two objects). What seems to be at stake is the heaviness of the adjoined material. Take PF to worry about the prosodic properties of the objects to be linearized. Appealing to a Prosodic (or Phonological) Hierarchy (Selkirk 1984, Nespor & Vogel 1986), there is at least a difference between a prosodic word and a phonological phrase. A head, regardless of how many morphemes it contains, is treated as a prosodic word, possibly a clitic group, but in any case less than maximal projections, whose prosodic structure starts at the level of the phonological phrase. A DP adjoined to a head should result in the formation of a prosodic word/clitic group, and if the DP consists of the (clitic) head only, this does not pose a problem. However, if this DP is a fully structured maximal projection, a conflict arises. Leaving further details aside, I take this to mean that the full DP must leave the head it adjoined to before the head moves on. There is only one possibility, given that adjunction to XP by movement is ruled out: the specifier of that head. Thus, while syntactically well-formed, a full DP cannot be sustained by the head it adjoins to. These possibilities are sketched here for convenience:

\[(86)\]
\[
\begin{align*}
\text{a.} & \quad [\_YP \_XP [\_X_P -Y]-Y \ldots \_XP \ldots] \\
\text{b.} & \quad # \quad [\_Z_P \_XP Z [\_YP \_Ø [\_X_P -Y]-Y \ldots \_XP \ldots]]
\end{align*}
\]
The final structure we end up with at this point is simplified in (87):

(87) \[
\begin{array}{c}
\text{AgrP} \\
\text{DP} \quad \text{Agr'} \\
\text{ton Janni} \\
\text{Agr0} \quad \ldots \\
[t\text{-ksero}-V\text{-v}\text{-Agr}] \\
\end{array}
\]

The DP is now free to undergo further movement; relevant for us is topicalization, i.e. movement to SpecTopP. As this projection is in the next higher Prolific Domain, it is well-formed. The complex head stays behind and may undergo further successive-cyclic head movement (which it presumably does at least up to T in Greek).

To recap, the relevant derivational steps are as follows:

(88) a. \[\text{[Agr} \ [\text{DP}\text{-v}\text{-Agr}]-\text{Agr vP}]\]

b. \[\text{[AgrP DP} \ [\text{Agr} \ [\text{DP}\text{-v}\text{-Agr}]-\text{Agr vP}]]\]

c. \[\text{[AgrP DP} \ [\text{Agr} \ [\text{DP} \ \text{CL}\text{-v}\text{-Agr}]-\text{Agr vP}]]\]

d. \[\text{[T' [CL}\text{-v}\text{-Agr}]-T \ldots [\text{AgrP DP} \ [\text{Agr} \ [\text{CL}\text{-v}\text{-Agr}]-\text{Agr vP}]]]\]

e. \[\text{[TopP DP Top} \ldots [\text{T' [CL}\text{-v}\text{-Agr}]-T \ldots [\text{AgrP} \ldots ]]]]\]

f. Ton Janni ton ksero.

‘Janni, I know him.’

(Greek)

---

34 While in spirit compatible with other derivational analyses of cliticization, and there have been many since Kayne (1975), this analysis differs in some respects. As I focus on CLLD structures, I cannot lay out a detailed presentation for other clitic constructions. The current approach, which takes the clitic to be derived by DP movement to a \(\phi\)-head, does not, however, imply an analysis of clitic doubling along the same lines. This is, of course, desired, given Cinque’s (1990) objections to a clitic doubling approach to CLLD. In other words, clitic doubling does not follow the derivational steps sketched here; in particular, it does not arise from DP movement to a \(\phi\)-head, then its specifier, followed by Copy Spell Out (yielding DP and doubled clitic). See Uriagereka (1995a) and references for arguments to adjoin the doubled element to the DP containing the clitic. Following the PF-requirements mentioned above, this makes the DP too heavy to adjoin to a head in the first place. Moreover, it would imply that the doubled element always surfaces in its \(\phi\)-position, which is not the case either. As it turns out, doubled material has a distinct “extraposed” character, better analysed in terms other than assumed here for CLLD.
4.5 Discussion of Left Dislocation Constructions and Copy Spell Out

To sum up, I have proposed a unified analysis of three types of LD constructions: based on the evidence that CLLD and CLD, in contrast to HTLD, exhibit connectivity between the LDed XP and its RP (Case-matching, lack of WCO and Condition A effects, sensitivity to Condition C), that the XP can be an idiom chunk only in these strategies, and that the RP in these two constructions is a specific pronominal element, I second a derivational analysis for CLD and CLLD, but a base-generation analysis of HTLD. In this respect it is also noteworthy that we take topicalization to be derived by movement in the current framework, and as such the movement variants of LD can be compared with this construction.

(89) is a rough representation for topicalization of an argument XP. What is important here is that it starts out in a thematic position (specifier or complement within v/VP, the θ-domain) and passes through an agreement position (e.g. TP or AgrP, within the φ-domain) before landing in topic position, here taken to be part of the discourse layer (ω-domain). I take the position for syntactically displaced topics to be TopP and languages differ whether the Top-head must be filled or not.35

(89) \[ \text{Topicalization} \]
\[ [\text{TopP XP Top} \ldots [\ldots \text{XP} \ldots [\ldots \text{XP} \ldots ]]] \]

This analysis—here forced by Full Interpretation qua Prolific Domains (see section 6.2 on Address identification)—allows the following derivational histories for the movement varieties of LD constructions, CLD (as in German) and CLLD (e.g. Greek):

35 I will come back to topics in chapter 5, where I also briefly discuss other discourse-driven displacement in a comparison with “scrambling.” Verb-relatedness of Top is further contingent on other properties of the language (such as V2).
(90) *Contrastive Left Dislocation*

\[ [\text{CP XP C} \ldots [\text{TopP } \text{XP} \ni \text{RP} \text{ Top} \ldots [\omega \ldots \text{XP} \ldots [\omega \ldots \text{XP} \ldots]]] \]

(91) *Clitic Left Dislocation*

\[ [\text{TopP XP Top} \ldots [\omega \ldots [\text{RP } [\omega [\text{XP } \ni \text{CL}]-\phi] \ldots [\omega \ldots \text{XP} \ldots]]] \]

CLLD and CLD differ from topicalization, and each other, in several respects. Topicalization unanimously displaces an element into a position within the next higher Prolific Domain (e.g. from \( \theta \)- to \( \phi \)- to \( \omega \)-position), and lower copies are deleted under standard assumptions (Nunes 1995).

Up to the point of “topicalization,” CLD works the same, but it then involves an additional step: the XP moves from topic position to a higher, left-peripheral position, outside the relevant domain for V2 (“quasi clause-external”). We identified this position as the specifier of C, the head that types the clause. This movement violates Exclusivity (it is Domain-internal), but can be rescued by the repair strategy Copy Spell Out. The lower, non-distinct copy thus undergoes PF-Spell Out and surfaces as a resumptive. The Copy Spell Out filler in this case is a *d*-pronoun (German, Dutch). It is a full maximal projection and hence can be stressed in its topic position. One by-product of this derivation is the contrastive character that has been observed for CLD. The topicalization-part of CLD then concerns the RP.

In CLLD, it applies to the CLLDed XP itself, which is taken to target SpecTopP and stay there. The difference between CLLD and CLD is the point of Copy Spell Out. I suggested a way to introduce the clitic RP inside the \( \phi \)-domain, clearly the clausal domain where it (first) surfaces. This derivation generates a full nominal DP in its \( \theta \)-position and moves it into the relevant \( \phi \)-head qua head-adjunction, an operation I motivated by comparison with the regular cliticization process. The D-head enters a checking relation with the \( \phi \)-head (such as T for subjects and Agr for objects), but leaves the categorial features of DP unchecked (\( \phi \)-features). This warrants movement to the specifier of that head, a
paradigmatic anti-local movement which must be followed by Copy Spell Out. The Exclusivity-violating copy is the one within the complex φ-head which is non-distinct from the copy in the specifier position and has less features checked. (Recall that it is not non-distinct from the original copy in the θ-domain, but that is irrelevant here.) As a head of sorts, the pronominal filler in this case is expected to be head-like element, hence a clitic. The XP then topicalizes by moving into SpecTopP.

These two derivational histories allow for a unified account for the similarities between CLD and CLLD (basically, in terms of reconstruction), while at the same time permitting an account for the differences observed. The fact that one RP is in topic position and the other clause-internal is not mysterious anymore, and neither is the observation that one can be stressed, while the other cannot. The difference in categorial status is also taken care of. The last two, arguably more severe differences, fall out as well. As the landing site of the CLLDed XP differs from that of the CLDed XP in being a recursive topic projection rather than a unique clause-typing position, iteration of CLLDed XPs is expected, but not of CLDed ones. Moreover, the root character of CLD, as opposed to CLLD, follows from the same fact.

HTLD does not suggest the possibility of XP-reconstruction (lack of connectivity) and should hence be analysed differently. In addition, the RP does not need to come in one specific pronominal guise\(^\text{36}\) and its position is not pre-determined either. We took these properties to indicate that the HT is base-generated in its surface position, while the RP is the originally merged argument that may or may not undergo movement (depending on the language, and it may even be topicalized). We thus yield the following possibilities for HTLD constructions, all borne out in the languages we considered:

\(^{36}\) In fact, it need not even be a pronominal, if one would want to push a similar derivation for epithet constructions. I cannot discuss the properties of these constructions, but refer the reader to Aoun & Choueiri (2000) for recent discussion and ample references.
Hanging Topic Left Dislocation

a. \[[CP \text{XP} [CP \text{C} \ldots \text{[\ldots]} \text{RP} \ldots]]]\]
b. \[[CP \text{XP} [CP \text{C} \ldots \text{RP} \ldots \text{[\ldots]} \text{[\ldots]} \text{RP} \ldots]]\]
c. \[[CP \text{XP} [CP \text{C} \ldots \text{[TopP} \text{RP Top} \ldots \text{[\ldots]} \text{[\ldots]} \text{RP} \ldots]]\]]

In contrast to topicalization, CLD and CLLD, the HT is adjoined to a maximal projection, rather than sitting in a specifier. Following our strict distinction between specifiers and adjuncts, and their properties, this framework makes a number of predictions. First, as an adjunct the HT must be base-generated and cannot result from movement, a prediction independently borne out from the data. Second, as an adjunct we expect the HHT to always precede any other LDed XP, should the two co-occur in the same position; clearly, this concerns CLD, as we take the CLDed XP to sit in SpecCP and the HT in AdjCP. Third, if the two occur in the same projection, we expect one to be unique, the other not. This prediction is also borne out. New data from German show that instances of multiple HTLD can be construed, even in combination with CLD, provided that the structural relations are as follows: there is only one CLDed XP and any HTs that occur must precede the XP.

Let us now turn briefly to some arising issues. One question that might be asked is why German does not allow CLLD. Given that German does have special clitics (in the sense of Zwicky 1977; see Zwart 1991, 1996a), at least in dialects (Abraham & Wiegel 1993, Abraham 1995, 1996a, 1996b), why could the derivation of CLD not proceed analogously to CLLD? While German does have special clitics, they differ from those found in other languages, such as the Romance languages, Arabic, Greek etc., those languages that allow CLLD. The main difference boils down verb-relatedness (Haegeman 1996): clitics in these languages are en- or pro-clitic to a verbal element, while German(ic) clitics tend to occur much higher, close to the complementizer. We can take this to indicate that “verbal” clitics are introduced in the \(\phi\)-domain, while “complementizer” clitics are licensed in the \(\omega\)-domain (see also Rivero 1997 on I- vs. C-clitics).
In Grohmann (2000a, 2000c) I sketch an analysis of German clitics along these lines: building on Uriagereka’s (1995a) proposal that certain clitics serve a discourse function (such as encoding “point of view”), I identified the ω-head “F” as the licensing head for clitics in Germanic. German and Dutch show an intricate pattern of prosodically reduced pronouns, roughly partitioned into weak and clitic pronouns (along the lines of Cardinaletti & Starke 1995, 1999). While weak pronouns surface in the φ-domain, clitics result from movement into the ω-domain, comparable to the approach to cliticization advocated here. The head-character of clitics is derived, resulting from DP movement into a head. As such, weak pronouns sit in specifier positions at Spell Out, while clitics sit on heads—and these heads are in a different domain from Romance or Greek clitic hosts.

Still, one might want to push this further and ask why the CLDed XP does not spell out its ω-copy as a clitic, if the RP in CLD is within the ω-domain, the same domain clitics are introduced. In other words, why can (93) not be a possible derivation for CLD?

(93)  

a. # [\[Φ XP \[ [XP ⊔ CL]-ω \] \[ ω \] \[ XP \]\[ ω \] \[ XP ... \]]]]  
b. # [TopP XP Top ... [\[Φ [XP ⊔ CL]-ω \] \[ ω \] \[ XP \]\[ ω \] \[ XP ... \]]]]  
c. # [TopP XP Top ... [Φ XP [\[XP ⊔ CL]-ω \] \[ ω \] \[ XP \]\[ ω \] \[ XP ... \]]]]

If clitics are introduced within the ω-domain, the corresponding XP might be expected to adjoin to that head and move on to its specifier. However, I identify the clitic-licensing ω-head as F. This head is lower than Top, which licenses syntactic topics. This rules out (93a) on the grounds that XP could not check [Top]. We could then hypothesize the derivation in (93b), where XP would move from F to SpecTopP. As our discussion above showed, this is not an option for PF-reasons. The hypothetical structure is the same as (86b) which we ruled out on sound grounds. Lastly, (93c) is ill-formed as it involves two violations of the CDE, an aspect which could suggest that a given object in the phrase marker may undergo Copy Spell Out only once.
A more serious question regards the possibility of CLLDing non-arguments. Greek is a nice language to illustrate CLLD with, as it only makes this strategy available to argument DPs.\textsuperscript{37} One difference between Greek and, say, Italian is that the former does not have corresponding clitics to resume (or double) PPs, APs etc. Given our assumption that all arguments must raise into the $\phi$-domain at some point, an extension to these cases in other languages does not seem unreasonable.\textsuperscript{38}

A more general issue concerns the form of the spelled out copy. In the LD constructions considered here, we are dealing with two types: a full pronominal resumptive element and a head-like clitic. Deciding on which form is inserted could depend on structural considerations. If the Exclusivity-violating copy sits in a head, a clitic is the Copy Spell Out filler, and if it sits in a specifier, a phrasal pronoun is inserted. This answer begs the question, of course, why in some specifier positions, the outcome is a (demonstrative) pronoun and in others a reflexive or reciprocal. One might extend the answer by defining the nature of the filler element in terms of Prolific Domains: spell out as reflexive/reciprocal in the $\theta$-domain, as a regular (or otherwise specified) pronoun in the $\omega$-domain. However, the analysis we gave to reflexive ECM-subjects (and possibly other constructions) concerns the $\phi$-domain. Moreover, generalizing on the basis of three types of Copy Spell Out might not be the wisest thing to do in the first place. I have no answers to this and related questions and leave them open for future research.

\textsuperscript{37} CPs can also be CLLDed, as in (i), but the nominal character of (argument) CPs warrants the same derivational history as argument DPs.

\begin{center}
(i) (To) oti o Jannis agapai ti Maria to ksero. \\
\textit{the} that the Jannis love.3SG the Maria CL know.1SG \\
`I know that Jannis loves Maria.' \textit{(Greek; Anna Roussou, p.c.)}
\end{center}

\textsuperscript{38} However, it poses the question of how adjuncts can ever be CLLDed followed by Copy Spell Out. If this process is the same as for arguments, something else needs to be said. But note that an analysis of CLLD with adjuncts qua Copy Spell Out in the $\phi$-domain makes one prediction: a C-related adjunct should not be able to be CLLDed. This can be tested with speech act adverbs \textit{(sincerely, frankly)}, which arguably are Direct Merged into the $\omega$-domain. I leave the details for future research.
4.6 Conclusion

In this chapter we have looked at left dislocation constructions, in particular at three different strategies: HTLD, CLD and CLLD. On grounds of connectivity between an LDed constituent and its resumptive pronoun we were able to distinguish two types of LDed elements: one derived and one base-generated. CLD and CLLD pattern alike and further comparison with topicalization strongly suggests that all three elements (the CLDed XP, the CLLDed XP and the topicalized XP) are the result of movement into the \( \omega \)-domain. HTLD, on the other hand, does not exhibit connectivity effects, which in turn suggests that the HT is merged directly into its surface position, and any movement in the clause involves the RP. Regarding the RP, we assume it is the originally merged argument in HT-constructions. For the movement variants, however, we can employ an alternative: the RP enters the derivation as the spelled out copy of an otherwise illicit copy. We saw that the differences between CLLD and CLD boils down to Copy Spell Out applying inside the \( \phi \)- or ther \( \omega \)-domain. We cashed out this difference further by accounting for other differences between the two (embedding, stacking).

The data we looked at also suggest a structural difference between CLDed XP and HT, where we identified the latter, base-generated element to sit in AdjCP and the former to move to SpecCP. This analysis not only offers empirical support for the difference between specifiers and adjuncts laid out in its details in chapter 3, but it also has empirical consequences (co-occurrence, precedence and stacking of HTs).
CHAPTER 5

(MULTIPLE) WH-QUESTION FORMATION:
MOVEMENT, INTERPRETATION AND TYPOLOGY

This chapter deals with the formation of Wh-questions from two perspectives. On the basis of cross-linguistic variation in the formation of multiple (binary) Wh-questions, I propose a tripartition into zero, singular and multiple Wh-movement languages. The primary criterion is the (non-)availability of single-pair interpretation, which can be made sense of once we dissociate Wh-movement from Q-typing. The different patterns, and the tripartition that emerges, suggest that German is a multiple Wh-movement language, contrary to popular belief. The second part of this chapter consists of empirical arguments for exactly this assumption. In sum, this chapter argues for interpretive differences as the result of Wh-displacement targeting the $\omega$-domain vs. the $\phi$-domain.

5.1 Introduction

In the following I will be concerned with a classification of languages into three types of strategies to form multiple Wh-questions. One type of language forms a multiple Wh-question by moving all Wh-phrases to a Wh-licensing position, another type moves only one Wh-phrase, and a third type does not move any Wh-phrase to that position. I
suggest dissociating Wh-displacement from clause typing, so moving a Wh-element is not necessary to license an interrogative interpretation. I further distinguish Wh-fronting from real Wh-movement. The latter targets a specific position, by default a focus position in the $\omega$-domain, while the latter does not. One instance of non-Wh-driven movement is so-called “Wh-scrambling,” as found in e.g. Japanese. In section 5.2 I present the basic facts about multiple Wh-questions, including the relevant data and a brief overview of some approaches. I propose a cleaned up tripartite typology in section 5.3, after reviewing a quasi-tripartition that has been suggested recently. One important aspect of classifying languages into multiple, singular and zero Wh-movement is the interpretation that languages make available. I concentrate on the difference between pair-list and single-pair readings, and compare the availability of especially the latter across languages. One consequence of the tripartition is that German should be viewed as a multiple Wh-movement language: German is like Bulgarian. I provide empirical support for this view in section 5.4. The upshot is that non-topicalizable material cannot occur between two Wh-phrases, supporting the analysis that all Wh-phrases move into the $\omega$-domain. I offer a twist to the proposal in section 5.5, suggesting that German is also like Italian in not forming “real” multiple questions to begin with. The conclusion in section 5.6 wraps up our discussion of the left periphery, addressing also relevant issues for Exclusivity.

5.2 Three Types of (Multiple) Wh-Question Formation

This section introduces the basic background. Data from a range of languages suggest that the strategy to express Wh-questions comes in three guises: languages either move one Wh-phrase, all Wh-phrases or none. After a presentation of the core data, I quickly review current approaches to the syntax of Wh-question formation.
5.2.1 A Brief Overview of Wh-Questions

At first glance, one might be tempted to hypothesize that in order to ask an information question—as opposed to a rhetorical question or one with echo interpretation—languages either move a Wh-phrase (henceforth, WH) or leave it in situ.\footnote{I use “in situ” in the liberal way it is standardly understood, namely roughly as “not moved to the same position as the equivalent element in other languages.” With the rise of functional projections and Checking Theory, the originally in situ interpretation of “base position” has been lost; other factors contribute also, which we will address shortly.} Before we refine this hypothesis, consider the following two sets of data. (1) represents some languages that move a WH to a sentence-initial position, and (2) others that do not.\footnote{French I and II designate two strategies available, both used as information questions (i.e. the one that leaves the WH in situ does not yield an echo interpretation, as the English equivalent would.}

(1) a. What did John buy? \hspace{1em} (English)
b. Was hat Peter gekauft?
what has Peter bought
‘What did Peter buy?’ \hspace{1em} (German)
c. Qu’est-ce qu’il a donné à Jean?
what has-he given to Jean
‘What did he give to Jean?’ \hspace{1em} (French I)

(2) a. Tanako-wa Mitsue-ni nani-o ageta no?
Tanako-TOP Mitsue-DAT what-ACC gave Q
‘What did Tanako give to Mitsue?’ \hspace{1em} (Japanese)
b. Zhangsan mai-le shenme?
Zhangsan buy-ASP what
‘What did Zhangsan buy?’ \hspace{1em} (Chinese)
c. Il a donné le livre à qui?
he has given the book to who
‘Who did he give the book to?’ \hspace{1em} (French II)
Since Huang’s (1982) seminal work on Wh-questions in Chinese, the “standard” account of Wh-questions is that all WHs are licensed in COMP (CP) at some point in the derivation—some languages do this overtly, others covertly. The Wh-element is taken to be an operator which must scope over the entire sentence at LF and the topmost projection in the clause, CP, is the position where operator-variable chains are created and wide scope can be established. Other things being equal, this view allows for a bipartition of the formation of Wh-questions into overt Wh-movement and Wh-in situ languages.

However, there arise some complications with this view. One is posed by so-called “Wh-scrambling.” Some Wh-in situ languages allow displacement of WH’s that does not seem to take place to check a Wh-feature (such as Japanese or Hindi). Another comes from languages that move all WHs. As we will see very shortly, these languages differ in the position the WH moves to (e.g. Bulgarian vs. Serbocroatian). A further complication has to do with an appropriate typology of Wh-question formation, a topic I tackle from one particular perspective, namely based on the (non-)availability of single-pair interpretation (Wachowicz 1974). In this context, the question arises whether there are (limited) parameters that can successfully group languages into clear patterns of Wh-formation or not. I suggest there are, given not unreasonable assumptions about the syntax and semantics of Wh-questions, and I will lay out all of these in the following.

### 5.2.2 Multiple Wh-Questions: Zero, Singular and Multiple Wh-Movement?

The clean bipartition that (1) and (2) suggest gets messy as soon as we add another WH, yielding a multiple Wh-question.\(^3\) Some languages move only one WH, others still do not move any, but a third class of languages moves all WHs. Throughout I will

---

\(^3\) To keep the discussion brief but clear, I restrict the discussion to Wh-questions with exactly two WHs, what Pesetsky (2000) calls “binary Wh-question.”
refer to the first type as “singular Wh-movement languages,” the second as “zero Wh-movement (or Wh-in situ) languages,” and the third type as “multiple Wh-movement languages.” Moreover, all varieties come in two flavours: those that obey Superiority and those that seem to violate it. Moreover, all varieties come in two flavours: those that obey Superiority and those that seem to violate it. 

English is a good example of a singular Wh-movement language which exhibits Superiority effects, that is, no lower WH2 may move across a higher WH1.

(3)  

a. Who bought what?  
b. * What did who buy?

German is an example of a language that seems to move also one WH only, like English, but does not show Superiority effects: either WH may front rather freely (Haider 1983, 1993, 1996, Müller 1995, Grohmann 1997b, Grewendorf 1998, Sabel 1998).

(4)  

a. Wer hat was gekauft?  
b. Was hat wer gekauft?  
   what has who bought  
   ‘Who bought what?’ (German)

Some Wh-in situ languages also lack Superiority effects, others do not (see e.g. Huang 1982, Tsai 1994, Watanabe 1993, Dayal 1996). Compare Japanese and Chinese:

4 Chomsky (1973: 246) formulates the Superiority Condition as follows:

(i) **Superiority Condition**
   i. No rule can involve X, Y in the structure ... X ... [ ... Z ...—WY] ... where the rule applies ambiguously to Z and Y, and Z is superior to Y.
   ii. The category A is ‘superior’ to the category B if every major category dominating A dominates B as well but not conversely.

5 I take Superiority purely descriptive here, referring to moving a lower WH2 over a higher WH1, regardless of the type of movement that takes place. That is, neither Japanese nor Chinese would move a WH to SpecCP in these instances, for example, but only the former allows fronting of WH2 over WH1. We will get back to this in detail.
French is an oddball: as mentioned above, it makes two strategies available, singular or zero Wh-movement. The same two strategies are also available in multiple Wh-questions, but regardless of whether one WH moves or none, Superiority must not be violated (see also below).

The third category consists of languages that move all WHs, which include, for example, the Slavic languages and Romanian (among others, Wachowicz 1974, Rudin 1988, Comorovski 1996, Bošković 1997b, Richards 1997, Citko 1998, Stepanov 1998, Stjepanović 1999). Multiple Wh-movement languages also show the split with regard to Superiority: some must obey Superiority, others can apparently violate it.
Romanian and Bulgarian are languages that are sensitive to Superiority, like English, but front all WHs obligatorily (but see Bošković 1997b, Richards 1997 on relaxation when more than two WHs are involved: apart from WH1, all WHs may be ordered in any combination; I will largely ignore Wh-questions with more than two WHs)

(9) a. Cine ce a vazut?
   b. * Ce cine a vazut?
      what who has seen
      ‘Who saw what?’ (Romanian)

(10) a. Koj kogo viđa?
    b. * Kogo koj viđa?
       whom who sees
       ‘Who sees whom?’ (Bulgarian)

But other, apparently multiple Wh-movement languages, such as Serbocroatian or Polish, may front all WHs in any order.

(11) a. Ko je keta kupio?
    b. *ta je ko kupio?
       what AUX who bought
       ‘Who bought what?’ (Serbocroatian)

(12) a. Ktgo kogo zobaczył?
    b. Kogo ktgo zobaczył?
       whom who saw
       ‘Who saw whom?’ (Polish)

This brief overview shall suffice for an illustration of the phenomenon. I now address briefly some theoretical issues concerning these strategies, consequences that arise, and an outline of the proposal that I present in the remainder of this chapter.
5.2.3 *Clause Typing and Q*

One question is whether Wh-movement is syntactically or semantically driven. Under the view that the Wh-operator (or interrogative clause-typer) sits on the WH, all WHs must move to SpecCP at some point.\(^6\) This approach goes back to Huang (1982) who proposes LF-movement of WHs in Wh-in situ languages. There is an alternative, namely that another element types the clause, possibly independent of the WH. Baker (1970) suggests a Q-morpheme, elaborating on an idea by Katz & Postal (1964), which was developed further by Cheng (1991). Under the latter analysis, all that is needed to license a question is Q in C, and languages allow either one of two strategies: (i) move a WH, which by default contains Q, to SpecCP or (ii) generate Q in C, which comes in the form of a Q-particle. This Q can be a phonologically pronounced morpheme such as Japanese *no* in (2a) or an unpronounced, empty morpheme. An implementation of this approach need not postulate LF-movement of the Wh-phrases.

We can schematize the possibilities for the generalized movement approach and the Q-typing approach as follows (where the copy pronounced at the point of Spell Out is marked in boldface and the LF-moved WH parenthesized):

\[(13) \quad LF\text{-movement approach}\]

\[
a. \quad \begin{array}{c}
CP \\
\text{WH} \\
\text{who}
\end{array} \quad \begin{array}{c}
C' \\
\text{C} \\
\text{did} \\
\text{…}
\end{array} \quad \begin{array}{c}
\text{… WH} \\
\text{… you see}
\end{array} \quad (= (1a))
\]

\(^6\) Or only one WH in every language, with lower WHs interpreted, for example, functionally (Chierchia 1991, Hornstein 1995) or by absorption (Higginbotham & May 1981).
Thus Q could sit on the WH in SpecP, as in (13a), or in C, as in (13b). The latter can be covert, as in Chinese, or overt, as in Japanese, for example.\footnote{This presentation is, of course, a gross simplification. The goal of this section is simply to motivate the deviation from “standard” approaches that I propose. Apart from the references cited here, see Watanabe (1993), Comorovski (1996), Dayal (1997), Richards (1997), Grewendorf (1998), Sabel (1998), Simpson (2000), and many more.}
Another issue concerns the landing site of moved WHs: do all displaced WH’s target the same position (such as SpecCP, for scopal reasons)? Under the LF-movement approach of Huang, all WHs target SpecCP. In multiple Wh-questions, all non-moving WHs raise at LF, possibly adjoining to one another (cf. May 1985). Rudin (1988) extends this view and proposes that in multiple Wh-movement languages, all WHs move overtly to SpecCP. In languages like Bulgarian, where Superiority is obeyed, the first element moves into SpecCP and all subsequent WH-movements right-adjoin to this WH. Those languages that lack Superiority are not taken to involve multiple WH-movement, but only singular WH-movement, with the lower WHs adjoining to TP. As this point will be relevant later, I postpone the discussion. In general, this approach is not incompatible with the alternatives outlined above, but the shift of attention is different. The Q-typing/LF-movement approach is concerned with licensing the clause type and establishing the appropriate scope. This approach focuses on the licensing condition for each separate WH, arguing that all WHs must be licensed in CP, and some languages have to do this overtly. We could thus extend the above structure with the following equally rough structure, which holds either at Spell Out or at LF:

![Diagram](image)

A modern implementation of this approach involves multiple specifiers of CP, rather than adjunction structures (Richards 1997, 1999). Under this view, the highest WH targets SpecCP first, and subsequent WH-movement “tucks in” additional WHs. Again, there are a number of variations available to deal with some of the variation we have seen already, and plenty more. As the multi-Spec approach is untenable in the current frame-
work, I will not discuss the details any further, but will return to some relevant aspects of related proposals by Stjepanović (1995, 1999), Bošković (1997a, 1997b, 1998b, 1999, 2000a, to appear), Citko (1998), and others.

I will argue in the following that Wh-movement is independent of interrogative force. Rather, the clause is typed interrogative by a question morpheme, the famous Q-particle (overt or covert). I am going to integrate Cheng’s approach into a more articulate structure of CP (à la Rizzi 1997), here understood as the ω-domain, the domain that encodes certain discourse properties. But the present approach does not require WHs to move to yield a well-formed question, not even in languages that do not make a Q-particle available (see also Hagstrom 1998).

Displacement of WHs takes place for an additional discourse effect, driven by a special feature, the feature [Wh], which might be related to “focus.” By separating [Wh] from [Q], we can license the interrogative clause across all languages without resorting to any kind of movement of WHs, and no necessity to move at LF either (see also Brody 1995, Hornstein 1995, Kayne 1998 for arguments against covert A’-movement). If WHs move, they do so for other reasons. It has been argued that languages that move WH to a C-related position (or ω-position), canonically target FocP. One argument comes from the complementary distribution of displaced WHs and displaced focus phrases (Horvath 1986, Brody 1990). I will extend the landing site possibilities by projections motivated by discourse effects. Singular or multiple Wh-movement languages then move one (or all) Wh-phrase(s) into the ω-domain. Zero Wh-movement languages may move WHs to some lower position, within the φ-domain. Japanese allows for “Wh-scrambling” and, as we will see, motivation can be found to front WHs in other Wh-in situ languages. As we will see, a distinction between “Wh-movement” (to check [Wh] in the ω-domain) and “Wh-fronting” (movement to a φ-position for some other purpose) will be necessary. In sum, WHs may move into one of two Prolific Domains: they may target the ω-domain or the φ-domain, and I lay out the syntactic and interpretive differences involved.
Following Wachowicz’s (1974) insights, recently recast with minimalist assumptions by Bošković (1998b), I will capitalize on the correlation between overt movement into the ω-domain on the one hand and the availability of single-pair vs. pair-list readings in multiple Wh-questions on the other. I follow Cheng (1991) in assuming that Q types the clause, and I extend Hagstrom’s (1998) proposal that the Q-morpheme is inserted in either one of two positions, yet to be specified, and moves to C. I will show that it is possible to tease apart the landing site of WHs. Anticipating the analysis, I will refer to displacement of a WH into the ω-domain and ω-movement and into the ϕ-domain as ϕ-movement. English, thus, ω-moves one WH, Bulgarian ω-moves all WHs, while Chinese does not ω-move any WH. We will see finer grained differences among languages which lead us to the tripartition alluded to:

(16)  

*Typo- and topological Wh-displacement*

i. multiple Wh-movement languages ω-move all WHs
ii. singular Wh-movement languages ω-move one WH and might ϕ-move others
iii. zero Wh-movement languages do not ω-move any WH, but may ϕ-move them

We will see, though, a peculiar state of affairs in German: this language tends to pattern with multiple Wh-movement rather than singular or zero Wh-movement languages. I propose that German be considered a multiple Wh-movement language and provide empirical support in the second part of this chapter.  

This said, we face a number of possibilities, all of which I will explore here, sketched for convenience in (17). I argue that αP refers to some projection within the ϕ-domain and βP within the ω-domain, and I will refine the position and relevance of Q (where “unmoved” refers to Spell Out position, as possibly in Chinese, for example).

---

8 The material in this chapter is an elaboration from early research (Grohmann 1997b, 1998), which evolved in material that I first presented as Grohmann (1999b), and then continued in collaboration (Citko & Grohmann 2000, forthcoming).
With respect to Q, we will see that while it universally moves to C—in the overt syntax or at LF—, the relevant point in the derivation for computing interpretation is the position it moves from. The next section boils down to the following: if Q moves from within the φ-domain and if only the lower WH2 is in the scope of the variable Q leaves behind, pair-list reading is available. If it moves from within the φ-domain and all WHs are in the scope of the variable, single-pair reading obtains. If it moves from within the ω-domain and all WHs are in the scope of the variable, single-pair reading is available, but this is blocked by any WH moving past it.

5.3 A Tripartite Typology

In this section I propose a typology of Wh-question formation into singular Wh-movement languages, multiple Wh-movement languages and zero Wh-movement or Wh-in situ languages. The core ingredient is the (non-)availability of single-pair readings across languages. Adapting recent proposals on the syntax and semantics of (multiple) Wh-questions by Hagstrom (1998) and subsequently Bošković (1998b), I push the emerging generalizations into a clean, tripartite typology of Wh-question formation.
5.3.1 Pair List vs. Single Pair Readings

Before we address syntactic variation in forming multiple Wh-questions across languages presented above, consider interpretive differences found both in different languages and with different strategies. Wachowicz (1974) found multiple Wh-questions to be ambiguous between a pair-list (PL) and a single-pair (SP) interpretation. An SP answer consists of a single proposition, whereas a PL answer consists of sets of propositions. A question like *Who bought what?* could thus potentially be answered either way:

(18) a. Mary bought a sweater. \((SP)\)
    b. Mary bought a sweater, Jane bought shoes, Anne bought a skirt… \((PL)\)

The availability of SP and PL readings is subject to cross-linguistic variation, an aspect that Bočković (1998b) considers in quite some detail, following Hagstrom (1998). He proposes a syntactic analysis based on Hagstrom’s semantic approach to multiple questions, which allows a tripartition across languages as envisioned here. I extend this analysis, but utilize the ingredients made available by Hagstrom and Bočković as much as possible; I will henceforth refer to this line as the “Hagstrom-Bočković approach.”

To illustrate differences in SP-availability across languages, English contrasts with Japanese, for example, in that it disallows SP readings in multiple Wh-questions. Thus, in Scenario I, which requires an SP answer like (18a), the English multiple Wh-question (19a) is infelicitous, while its Japanese counterpart (19b) is a legitimate way of requesting the desired information (leaving aside quiz question contexts for now).\(^9,10\)

---

\(^9\) I indicate infelicitous questions for the particular scenario with a hash mark and add the relevant readings in parentheses on the right-hand side, starring non-available readings.
(19) Scenario I: John is in a store and off in the distance sees somebody buying an article of clothing, but he does not see who it is, and neither does he see exactly what is being bought. He goes to the shop assistant and asks:

a. # Who bought what? (*SP)
b. Dare-ga nani-o katta no?
   who-NOM what-ACC bought Q
   ‘Who bought what?’ (Japanese)

We thus reach a first descriptive generalization which we will refine and extend as we go along:

(20) Descriptive generalization

Japanese allows, but English disallows, SP readings in multiple Wh-questions.

As mentioned above, one obvious difference between the English question and its Japanese counterpart in (19) is that the first WH has overtly moved to the clause-initial position in English, but not in Japanese, and it is safe to assume this to be an ω-position (e.g. FocP or CP). On the assumption that this movement checks the feature [Wh] in English, but not in Japanese, I will then investigate the following hypothesis, where by “Wh-movement” we understand movement to check [Wh], or—to put it in more general terms—movement targeting the ω-domain:

---

10 A conjoined question is acceptable in this situation, leaving aside D-linked questions which I return to. As a general disclaimer, the following judgements might be idealized somewhat, they are certainly subtle. The majority is taken from Bošković (1998b).

11 I will not consider the Vacuous Movement Hypothesis (VMH) of George (1980), Chomsky (1986a). As will become clear in the following, I assume that one WH moves overtly in English, whether subject, object or adjunct. See Agbayani (2000) for recent discussion and relevant references. The issue of the VMH is irrelevant here for other reasons. For example, (i) is equally infelicitous in this context:

(i) #What did you sell to whom?
(21) *Working hypothesis (to be revised)*
Overt Wh-movement forces PL interpretation.

The Hagstrom-Bošković approach argues explicitly for such a correlation between overt Wh-movement and interpretation. One prediction the hypothesis makes is that French should make both readings available, depending on the strategy employed: French I should behave like English, while French II should pattern with Japanese (cf. (1c) and (2c) above).

Scenario II shows that this correlation is on the right track.\(^\text{12}\)

(22) *Scenario II*: Mary sees John down the corridor. Not wearing her glasses, she notices that John gives something to someone but can’t figure out what the item nor who the person is. So she asks her friend to find out:

a. *Il a donné quoi à qui?*  
   *he has given what to whom*  
   ‘What did he give to whom?’  
   \((SP)\)

b. *# Qu’a-t-il donné à qui?*  
   *what has-he given to whom*  
   ‘What did he give to whom?’  
   \((*SP)\)

Turning to multiple Wh-movement languages, Bulgarian should also behave like English in that it allows only a PL answer—after all, movement of both WHs still falls under (21). This prediction, too, is confirmed:

\(^{12}\) (22) represents basically the same scenario as (19), but as French does not employ subject-verb inversion, it is hard to unambiguously test for Wh-movement of a subject-WH. The observation that overt Wh-movement in French destroys the SP reading is due to Bošković (1998a). See also Boeckx (1999b) and references cited there.
Likewise, Serbocroatian should not make an SP answer available, as both WHs front overtly. This prediction regarding Serbocroatian is not borne out, however:

\[\text{(24) Scenario II} \quad \text{(cf. (22))}\]
\[
\begin{align*}
\text{a. } & \quad \text{Ko je %\v{c}ta kupio?} \\
& \quad \text{(SP)} \\
\text{b. } & \quad \text{\v{c}ta je ko kupio?} \\
& \quad \text{what AUX who bought} \\
& \quad \text{\textquoteleft Who bought what?\textquoteright} \\
& \quad \text{(Serbocroatian)}
\end{align*}
\]

In Bulgarian—a language that fronts all WHs, yet obeys Superiority—a multiple Wh-question is infelicitous in a situation that requires an SP answer (cf. (23a)). Naturally, (23b) is purely unacceptable. Interestingly, in Serbocroatian the situation is different. Not only does it allow either order of the WHs, it also makes available the SP reading. In fact, (24b) only makes the SP reading available, while (24a) allows both SP and PL.

Bo\%kovi\c\ calls the loss of PL interpretation by moving a lower WH2 over a higher WH1 “Interpretive Superiority,” contrasting it with “Syntactic Superiority,” which does not allow non-D-linked WH2 > WH1 to begin with, (holding in English and Bulgarian, for example). To complete the terminology of Superiority types, Bo\%kovi\c calls the possibility of fronting a bare (or D-linked) WH2 over a bare (or D-linked) WH1 “Apparent Superiority,” as German (or English) make available. As I will employ this terminology, consider the breakdown as follows:
Types of Superiority

i. Syntactic Superiority: *non-D-linked WH2 > non-D-linked WH1
   (such as English, Bulgarian, Romanian, Chinese, French)

ii. Apparent Superiority: (any) WH2 > (any) WH1
   (such as Japanese, German, Serbocroatian, Polish)

iii. Interpretive Superiority: loss of PL reading when WH2 > WH1
   (such as English, Japanese, Serbocroatian)

In this respect, Serbocroatian seems to have more in common with languages that do not move any WH—either strict Wh-in situ languages like Japanese or the alternative strategy available in French (i.e. French II): both allow an SP answer.

This observation becomes even more apparent when we consider situations that clearly ask for a PL answer. Consider the following scenario for Japanese, another language that allows for Apparent Superiority:

(26) Scenario III: John sees Jim finishing off his daily sales. He sees a bunch of people that he knows walk away and Jim stacking his left-over merchandise, so he asks:
   a. Dare-ga nani-o katta no? \( (\text{PL}) \)
   b. # Nani-o dare-ga katta no? \( (*\text{PL}) \)
      what-ACC who-NOM bought Q
      ‘Who bought what?’ \( (\text{Japanese}) \)

As observed by Bošković, the only felicitous question is the one which does not exhibit Apparent Superiority, i.e. where the order of WHs is maintained. If the question involves Apparent Superiority, an Interpretive Superiority effect arises: the PL reading gets lost and the SP interpretation is the only one available.

It should come as no surprise that in this PL-provoking context, Serbocroatian shows the same pattern as above, namely like Japanese it makes the PL reading available only with the order WH1 > WH2, conforming to Apparent Superiority:
We can actually find a similar effect in English. Recall that the ordering that obeys Syntactic Superiority allows for a PL reading (in fact, it does not allow SP). While bare WHs can never violate Syntactic Superiority, a D-linked WH2 may move over a D-linked WH1 (Pesetsky 1987). In this instance of Apparent Superiority in English, however, we can detect Interpretive Superiority:

As expected, Bulgarian, a language that is sensitive to Syntactic Superiority, patterns like English with respect to PL-availability, shown in (29) with bare WHs.

---

13 Barss (1990, 1999) notes that the preferred, if not only, reading for (28c) is SP. Apparently, the PL interpretation is not out for all speakers, though. I might thus idealize these cases, for the purpose of presentation. See section 5.5.2 for discussion, however.

In either case, it is noteworthy that (28c) is still infelicitous in a context which forces an SP answer, such as Scenario I in (19). Interestingly, Anna Roussou (p.c.) observes that in “quiz questions” of the sort illustrated in (i), which clearly ask for an SP reading, the multiple Wh-question is well-formed:

(i) In the battle of Salamina, who defeated whom?

However, as she notes further, this seems to be only possible with a subject-WH:

(ii) a. * Who did the Greeks defeat in the battle of Salamina when?
    b. * When did the Greeks defeat who in the battle of Salamina?

Pending further discussion, it is not unreasonable to assume that the fronted WH in (i) does not undergo movement into the \(\omega\)-domain, but remains in the subject position, an option which presumably is also available in instances of D-linking (see section 5.5.3).
(29) **Scenario III** 

a. Koj kakvo e kupil?

b. *Kakvo koj e kupil?

‘Who bought what?’

As Citko & Grohmann (forthcoming) note, Bulgarian D-linked Wh-questions do allow movement of either WH. Furthermore, if WH2 moves over WH1 (i.e. Apparent Superiority), the preferred reading is the PL reading (Roumyana Izvorski, p.c); the same applies to Romanian (Ileana Comorovski, p.c):

(30) **Pair-list reading**

a. Koga kniga koj covek e kupil?

b. La care cînd te ai gîndit?

‘Which book did which person buy?’

‘Which one have you thought of when?’

In sum, if Japanese is a well-behaved Wh-in situ language that does not involve movement of any WH into the ω-domain to check [Wh] (standard assumption) and if Serbocroatian patterns alike with respect to interpretation, it should also be a Wh-in situ language (Bošković 1998b). On the other hand, if English moves one WH into the ω-domain and Bulgarian moves all (both standard assumptions), but if the two differ in those contexts which seem to lack ω-movement of a WH in English, but not Bulgarian the patterns we have seen so far allow for an extension of the initial hypothesis:
(31) *Working hypothesis (extension)*

i.  ω-movement of any WH forces PL interpretation, regardless of Superiority

ii. φ-movement of all WHs showing Apparent Superiority destroys PL reading

iii. φ-movement of all WHs without Apparent Superiority allows SP reading

Thus, (31i) denotes instances of high displacement of a WH, presumably triggered by checking a Wh-feature. Instances we have seen here are, for example, Bulgarian or Romanian in any multiple Wh-question (where both WHs move) and English or French I with bare WHs (where only one WH moves). If a WH moves for any other reason and targets a lower position (which we will both come back to), SP interpretation is available. Hypothesis (31ii) basically yields Interpretive Superiority effects, enforcing the SP reading when a bare WH2 moves over WH1, but extends to instances of D-linking in these languages as well. The relevant examples we have seen were from Japanese, Serbocroatian and English. (31iii) addresses multiple Wh-question which conform to any Superiority, i.e. Japanese, French II or even Chinese (not illustrated).

5.3.2 *A Typological Tripartition*

We can summarize the descriptive findings of the Hagstrom-Bošković approach as follows. In Wh-in situ languages, the SP reading is always available, but the PL reading only when (Syntactic or Apparent) Superiority is obeyed. The latter falls under Interpretive Superiority: if a lower WH moves over a higher one, only the SP interpretation is acceptable. An apparent exception to this generalization is Serbocroatian, which allows SP readings although both WHs are fronted. Moreover, in cases of Apparent Superiority (WH2 > WH1), the SP reading is actually preferred (Interpretive Superiority)—if the PL reading is at all available (Bošković notes dialectal variation)—just as in Japanese. English also exhibits Interpretive Superiority, once we look at D-linked Wh-phrases. We might want to assume that D-linked WHs do not undergo the same movement as bare
WHs (Pesetsky 1987). In this case, we could account for the latter observation: D-linked WHs do not (need to) move into the \( \omega \)-domain. Arguably, this is true of Japanese regardless of whether the WHs are D-linked: as a well-behaved Wh-in situ language, WHs do not move to check [Wh]—if they move, they do so for other reasons. This gives us one possibility to fit Serbocroatian into the generalization. If the WHs do not move to check a Wh-feature in this language, Interpretive Superiority is explained, and the reason why Serbocroatian patterns with Japanese is because both are Wh-in situ languages.

This is essentially the conclusion Bo\(\ddot{\text{e}}\)ković draws, yielding a clear diagnostic for zero Wh-movement languages. The Hagstrom-Bo\(\ddot{\text{e}}\)ković approach allows for the following classification of strategies to form multiple Wh-questions across languages:

<table>
<thead>
<tr>
<th>If a language has:</th>
<th>It allows readings:</th>
<th>But:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Zero Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL/SP</td>
<td></td>
</tr>
<tr>
<td>(Japanese, Chinese, Serbocroatian, Polish, Russian, French II…)</td>
<td>WH2 &gt; WH1: SP</td>
<td>(such as Wh-scrambling)</td>
</tr>
<tr>
<td><strong>Singular Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL</td>
<td>German:</td>
</tr>
<tr>
<td>(English, Greek, Spanish, Italian, French I, German, …)</td>
<td>WH2 &gt; WH1: SP</td>
<td>WH2&gt;WH1:PL (even D-linking) (such as D-linking)</td>
</tr>
<tr>
<td><strong>Multiple Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL</td>
<td></td>
</tr>
<tr>
<td>(Bulgarian, Romanian…)</td>
<td>WH2 &gt; WH1: PL</td>
<td>(even D-linking)</td>
</tr>
</tbody>
</table>

Table 1: Quasi-tripartition à la Bo\(\ddot{\text{e}}\)ković-Hagstrom

Apart from the third column, the exceptional one, we have seen everything we need to know to follow this classification.\(^{14}\) The tripartite split envisioned here seems to

\(^{14}\) I am perfectly aware that this typology is far from complete, and some languages I mention I do not consider further. However, my main goal in this chapter is to revisit the Hagstrom-Bo\(\ddot{\text{e}}\)ković approach from a German perspective, which I am going to do next. As such, I take the preceding discussion at face value. I will address some questions that arise at this point in the following, but by and large present the main analysis as close to the original of Hagstrom (1998), extended by Bo\(\ddot{\text{e}}\)ković (1998b), as possible.
fit in nicely with the data. The third column anticipates the following discussion, and is
the reason for one half of this chapter: to show that German fits the tripartition without an
ad hoc repair strategy. Basically, if German is, as standardly assumed, a singular Wh-
movement language just like English, we would expect it to pattern like English with re-
spect to the availability of SP/PL interpretation.

However, German, a language that seems to differ from English mainly in that it
apparently need not obey Syntactic Superiority, does not meet our expectations cf. (26)):
(32) *Scenario III*: John sees Jim finishing off his daily sales. He sees a bunch of people
that he knows walk away and Jim stacking his left-over merchandise, so he asks:
a. Wer hat was gekauft? (PL)
b. Was hat wer gekauft? (PL)
   what has who bought
   ‘Who bought what?’ (German)

In German, which also allows for Apparent Superiority, both orders yield the PL
reading. Interestingly, neither order allows SP interpretation: if we replace Scenario III
with Scenario I (from (19)), neither question is felicitous, not even with D-linked WHs:

(33) *Scenario I*: John is in a store and off in the distance sees somebody buying an ar-
ticle of clothing, but he does not see who it is, and neither does he see exactly
what is being bought. He goes to the shop assistant and asks:
a. # Wer hat was gekauft? (*SP)
b. # Was hat wer gekauft? (*SP)
c. # Welcher Kunde hat welches Produkt gekauft? (*SP)
d. # Welches Produkt hat welcher Kunde gekauft? (*SP)
   which product has which customer bought
   ‘Which customer bought which product?’ (German)

While Serbocroatian and Japanese pattern together, German does not pattern with
English, which seems to form multiple Wh-questions just like German, namely by mov-
ing one WH overtly into the ω-domain and leaving the other one(s) in situ. Instead, it appears to pattern with Bulgarian, another language that fronts all WHs, like Serbocroatian, but unlike Serbocroatian seems to really Wh-move these elements. (32) and (33) resemble all discussion of Bulgarian and Romanian so far, but not the other languages.

We seem to be forced into adopting one of two possibilities: German is very different, and the odd behaviour with respect to the non-availability of SP readings is due to some other factor, or German is a multiple Wh-movement language, just like Bulgarian. Hagstrom (1998), and by extension Bošković (1998b), adopt the former hypothesis and relate this behaviour to the fact that German is a “scrambling” language. Apart from the fact that we do not really know what “scrambling” really is, Serbocroatian is also a “scrambling” language, and so is Japanese. Unlike Japanese, however, Serbocroatian fronts all WHs, yet we seem to get some mileage out of the hypothesis that Serbocroatian is actually a Wh-in situ language.

By parity of reasoning, I opt for the second possibility: German is like Bulgarian and moves all WHs overtly into the ω-domain. I will lay out the details in the following, suggesting that this approach gets us places and manages to account for some intricate data regarding the interaction of WHs with other elements in the ω-domain.

In the remainder of this chapter I will thus address the following two questions that fall out from the data presented above and the tentative tripartition in Table 1.

(A) Why is there a correlation between syntactic movement, such as scrambling or fronting, and the availability of an SP interpretation in Wh-in situ languages?

(B) Why does German not pattern with other singular Wh-movement or at least scrambling languages regarding the availability of PL in WH2 > WH1 questions?

15 Or somewhere inside the φ-domain. As mentioned already, “in situ” refers to any position below the ω-domain, be it the thematic position or the agreement position, depending on whether the language moves all arguments overtly or not. We will address the syntactic and positional details shortly.
What is particularly puzzling is the German pattern. If, however, the point could be made that German is actually a multiple Wh-movement language, we would yield a clean tripartition, as shown in Table 2. This is the direction I embark on and argue for in quite some detail next (Grohmann 1999b, Citko & Grohmann 2000, forthcoming).

<table>
<thead>
<tr>
<th>If a language has:</th>
<th>It allows readings:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Zero Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL/SP</td>
</tr>
<tr>
<td>(Japanese, Chinese, Hindi, Serbocroatian, Polish, French II…)</td>
<td>WH2 &gt; WH1: SP</td>
</tr>
<tr>
<td><strong>Singular Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL</td>
</tr>
<tr>
<td>(English, Swedish, Spanish, Greek, Italian, French I…)</td>
<td>WH2 &gt; WH1: SP</td>
</tr>
<tr>
<td><strong>Multiple Wh-movement</strong></td>
<td>WH1 &gt; WH2: PL</td>
</tr>
<tr>
<td>(Bulgarian, Romanian, German…)</td>
<td>WH2 &gt; WH1: PL</td>
</tr>
</tbody>
</table>

Table 2 An improved tripartition

Bošković (1997b) argues that Wh-phrases in Serbocroatian do not involve Wh-movement, but focus-driven movement to some position within TP, our φ-domain. Wh-movement, i.e. movement that checks [Wh], targets an ω-projection. As Serbocroatian does not move WHs to check [Wh], but nevertheless displaces all WHs, I refer to it as φ-movement of WHs. The same goes for rearrangement of WH-ordering in other zero Wh-movement languages (like Japanese or Hindi). What is important is that multiple Wh-movement languages move all WHs into the ω-domain. Implementing ideas from the Hagstrom-Bošković approach to multiple questions, I propose an analysis that motivates all required movements and captures the different landing sites.

---

16 In Citko & Grohmann (forthcoming) we further consider whether Slavic languages like Polish or Serbocroatian always pattern with Wh-in situ languages, and compare these languages with Russian. The conclusion we reach is that all three languages can safely be classified as Wh-in situ languages, in the context expounded here.

17 See also Stjepanović (1995, 1999) for the original proposal that Serbocroatian Wh-questions involve focus movement, related work by Bošković (e.g. 1997a, 1997b, 1998a, 1999) for more arguments, Rudin (1988) for the original idea that the φ-domain is targeted in Serbocroatian, but not Bulgarian, and Bošković (1997c), Boeckx & Stjepanović (1999) for the suggestion that the focus projection is part of the φ-, not the ω-, domain.
We can now revise our working hypothesis from above in a way that seems to not only capture the discussion so far, but exhaustively extends, as we will see shortly:

(34) Working hypothesis (revised)
   i. movement of any Wh-phrase targeting the \(\omega\)-domain forces PL interpretation
   ii. movement of any Wh-phrase maximally targeting the \(\phi\)-domain allows SP
   iii. a. movement of any Wh-phrase within the \(\phi\)-domain forces SP if \(WH2 > WH1\)
       b. any Wh-phrase within the \(\phi\)-domain also allows PL if \(WH2 > WH1\)

5.3.3 A Basic Semantics for Questions and the Role of \(Q\)

The answer to question (A) from above—why there seems to be a correlation between syntactic movement and the availability of SP reading in Wh-in situ languages—follows from the semantics for Wh-questions proposed by Hagstrom (1998) and extended by Bošković (1998b). It relies on the universal existence of a Q-morpheme, which is responsible for interrogative clause-typing. In English-type languages, the Q-morpheme is a covert counterpart of an overt particle found in many Wh-in situ languages, such as Japanese (see Cheng 1991 for a detailed description). As Bošković (1998b) shows, the position of a Q-morpheme correlates with the availability of SP/PL readings with minimal assumptions, an approach I will make critical use of here.\(^\text{18}\)

The Hagstrom-Bošković approach makes basically two options available for insertion of the Q-morpheme, henceforth \([Q]\). Next I present what I take to be the most important ingredients of Hagstrom’s semantics to derive the following characterization:

(35)   a. \([Q]\) moves from some clause-internal position to \(C^0\) \(\rightarrow\) pair-list
   b. \([Q]\) is generated above both WH’s and moves to \(C^0\) \(\rightarrow\) single-pair

\(^\text{18}\) Out of context, these minimal assumptions make for a rather complex system, however. In the interest of clarity, I will review the gist of Hagstrom’s (1998) proposal along the lines presented by Bošković (1998b), including the latter’s simplified adaptations.
Following the assumption that WHs do not need to move, but that [Q] universally undergoes movement (by LF), Hagstrom provides arguments from various Wh-in situ languages that the clause-internal position corresponds to the lowest WH (chapters 2 and 3). I thus take the “low Q” to sit on WH2 in binary Wh-questions. The “high Q” position is somewhere above all WHs, for reasons I lay out momentarily. The structure of the articulated C-layer, our $\omega$-domain, which we assumed so far, looks as sketched in (36):

(36) \[ CP > \text{TopP} > \text{FocP} > \text{TopP} > \text{FP} \]

Given that we identified this Prolific Domain with discourse properties, I suggest looking here for a WH-independent position of origin for [Q]. I explore a specific insertion site for high Q, namely F, the lowest C-related head, acting as the interface between the $\omega$- and the $\phi$-domains.\footnote{As we will see shortly, I will modify the structure of the $\omega$-domain slightly. Moreover, I suggest parameterizing the function of “F” as the mediator between agreement and discourse properties. As such, nothing really hinges on the labels of the projections in this or any other domains for the purposes of discussion.} As we will see next, the SP interpretation is destroyed by any WH crossing [Q] in F. [Q] is thus not only the universal clause-typer for interrogatives but also the determining element for SP/PL-interpretation.

For more on the nature of [Q] and the detailed semantics briefly sketched here, see in particular chapters 5 and 6 of Hagstrom (1998). The gist of the analysis is the following, adapted from Bočković (1998b). A single question is a set of propositions, which can be responded either by choosing one proposition from the set or by denying the presupposition that there is an answer. Call this the rule of “Single Question Recognition.” By extension, the rule of “Multiple Question Recognition” identifies a (pair-list) multiple question as a set of questions, which can also be responded to in two ways: by selecting one proposition from the set $A$ for each member set $A$ by denying the presupposition that there is an answer in $A$.

The following are the relevant definitions Hagstrom provides.
(37) *Single Question Recognition*
If the semantic value of an utterance is of type <pt> (a set of propositions),
then the utterance is a (single) question.
To respond: (a) one proposition from the set is selected,
or (b) the presupposition (that there is an answer) is denied.

*(Hagstrom 1998: 148)*

(38) *Multiple Question Recognition*
If the semantic value of an utterance is of type <pt, t> (a set of questions),
then the utterance is a (pair-list multiple) question.
To respond: For each member set $A$,
(a) one proposition from the set $A$ is selected,
or (b) the presupposition (that there is an answer in $A$) is denied.

*(Hagstrom 1998: 148)*

Hagstrom treats [Q] as an existential quantifier over a choice function (where the
semantic value of a Wh-element is a set of individuals):

(39) *Choice Function*
A function $f$ is a *choice function* if it applies to a (non-empty) set and yields a
member of that set.

*(Hagstrom 1998: 130)*

[Q] always ends up in C by moving from one of two positions: the low position,
attached to WH2, or from the high position, here tentatively taken to be F.\(^\text{20}\) The trace left
behind by Q-to-C movement, regardless of its original position, is the choice function
variable which determines the interpretation (SP vs. PL).

\(^{20}\) This is Bošković’s (1998b) simplified adaptation of Hagstrom’s analysis which in-
volves more movement, in terms of “Q-migration.” The model presented here should be
enough for our purposes, though.
If [Q] starts out from low position, the choice function variable has only WH2 in its scope; this yields a multiple Wh-question calling for PL interpretation. The SP reading is enforced by merging [Q] high; the choice function variable it leaves behind scopes over all WHs (which are arguably generated lower than the \( \omega \)-domain).

The following is a very rough outline of the analysis Hagstrom provides, adapted for ease of exposition from Bošković’s discussion.

Merging with WH2 and moving on to C, the choice function variable left behind by [Q] takes WH2 as its argument—here understood as a set of individuals—and returns an individual, which is taken as an argument by the verb. The value of the higher WH1 is a set, as it is outside the scope of the choice function variable. Further composition of WH1 yields a set of sets of propositions, one set of propositions for each value \( x \) in the set of WH1, shown in (40) below. Hagstrom calls this composition “Flexible Function Application,” a repair strategy for semantic type mismatches. It applies in particular when a function receives a set of arguments instead of a single argument, in which case the function is applied to each argument in the set of arguments. The result of this function application yields a set.\(^{21}\)

By the Multiple Question Recognition rule, an appropriate answer to a question such as (40a)—where [Q] is merged low (attached on WH2)—provides an answer to each of the constituent questions of the set \{What did A buy?, What did B buy?, …\} or, more formally, (40b):

\(^{21}\) The formal definition is as follows:

(i) \( \text{Flexible Function Application} \)
\[
\ll f \ a \rr = \begin{cases} 
(f(a) & \text{(where } f \text{ and } a \text{ are sisters)} \\
\lambda m \exists x. [m = f(x) \land a(x)] & \\
\lambda m \exists g. [m = g(a) \land f(g)] & \\
\lambda m \exists g \exists x. [m = g(x) \land f(g) \land a(x)] 
\end{cases}
\]

\( \text{whichever is defined.} \)

\( \text{(Hagstrom 1998: 142)} \)
a. Who bought what?
b. \{\{A \text{ bought } f_1 (\text{WHAT}), A \text{ bought } f_2 (\text{WHAT}), \ldots\}, \\{B \text{ bought } f_1 (\text{WHAT}), B \text{ bought } f_2 (\text{WHAT}), \ldots\}, \ldots\}\n
Merging \[Q\] low thus yields a set of questions. Merging \[Q\] high, though, yields a set of propositions. Both WHs are contained within the argument of the choice function variable. Composition of WH2 (qua Flexible Function Application) returns a set of properties which, once applied to the set WH1, yields one proposition for each possible subject with each possible object. In other words:

\begin{align*}
\{A \text{ bought } a, A \text{ bought } b, \ldots, B \text{ bought } a, B \text{ bought } b, \ldots\},
\end{align*}

(\text{where WH1=}\{A, B, \ldots\} \text{ and WH2=}\{a, b, \ldots\})

The Single Question Recognition rule then gives an SP answer to this question.

In sum, if the choice function variable left behind by movement of \[Q\] to \(C\) scopes over WH2 only, we get PL interpretation (“low Q”); if it scopes over both WHs, we get SP interpretation of the multiple Wh-question (“high Q”), or expressed as follows, where bold strikethrough indicates the position of the choice function variable:

\begin{align*}
\text{(42) Choice Function Variable and Question Interpretation} \\
\text{a. } [Q]-C \ldots WH1 \ldots WH2-[Q]-\ldots \quad (\text{low } Q: \text{PL})
\end{align*}

\begin{align*}
\text{b. } [Q]-C \ldots [Q]-\ldots WH1 \ldots WH2 \quad (\text{high } Q: \text{SP})
\end{align*}
The interesting case for us is high [Q] and the (non-)availability of the SP reading across languages, and this is what I shall concentrate on (for low [Q], see Hagstrom 1998). I assume that all relevant movement is overt, except for [Q].

In the next sub-section I lay out the basic assumptions regarding the derivations of multiple Wh-questions across languages, including the identification of the relevant positions. I suggest on the one hand that all ω-movement targets FocP and FP in the canonical cases, and on the other that all φ-internal positions relevant here either involve the regular agreement projections (yielding WH1 > WH2) or a “syncretic” category, one which contains several relevant features (for the order WH2 > WH1).

5.3.4 The Syntax of Q, Prolific Domains and Anti-Locality

In this sub-section I tentatively apply the Hagstrom-Bošković approach to the Domain-driven framework presented here. While Bošković’s (1998b) discussion centered around a line of reasoning, rather than detail of the analysis, the details of some of the derivations presented in the following will also require additional research. However, I supply some motivation to pursue the current line of reasoning, which I take to be worthwhile for the reason that it offers a way of classifying languages into one of three types of multiple Wh-question formation. In this section I restrict the discussion to a simple conversion of the results from the Hagstrom-Bošković approach.

5.3.4.1 Zero Wh-movement languages

The two options—Wh-scrambling within the φ-domain (WH2 > WH1) or retaining base order (WH1 > WH2)—are available to Wh-in situ languages, which explains why both SP and PL readings are felicitous in either context (section 5.3.1). With high [Q], the choice function computed over the variable ranges over all Wh-phrases in situ (yielding SP), whereas with low [Q] it ranges only over WH2 (yielding PL). I suggest
(44) to be the relevant structures for a multiple Wh-question in Chinese, such as (43), with a PL reading and an SP reading, respectively (where CP $\approx \omega \Delta$, TP $\approx \phi \Delta$, vP $\approx \theta \Delta$).\(^{22}\)

(43) Shei mai-le shenme?

who buy-ASP what

‘Who bought what?’ (Chinese = (6a))

(44) a. $[[\text{CP} \ [Q]-C \ldots [\text{TP} \ WH1 \ T \ldots [\text{vP} \ WH1 \ldots WH2 \ [Q] \ldots]]] \quad (PL)$

b. $[[\text{CP} \ [Q]-C \ldots [\text{FP} \ [Q]-F \ [\text{TP} \ WH1 \ T \ldots [\text{vP} \ WH1 \ldots WH2 \ldots]]] \quad (SP)$

In (44a), [Q] is attached to WH2, which arguably does not move overtly, and ends up in C. The relevant trace it leaves behind is on WH2. Consider this trace the relevant choice function variable which takes only WH2 as its argument. The result is a PL interpretation. In (44b), on the other hand, [Q] is merged high—somewhere within FP (see fn. 22). In this case, both WHs are contained within the choice function variable’s argument and through Flexible Function Application, we yield an SP reading.

Other Wh-in situ languages have very similar structures. While French II might look exactly like Chinese in the relevant aspect,\(^{23}\) there is good evidence that at least in Serbocroatian all WHs leave their $\theta$-positions. After all, we know that all WHs obligato-

\(^{22}\) I will not get too specific about the (syntactic) nature of [Q]. It might not play a role here whether it is a head or something else, or how exactly it moves. The point I make in this section regards a potential solution which takes the liberty of leaving some questions open. I address this issue briefly in section 5.3.4.4, suggesting that [Q] starts out joined.

\(^{23}\) Cedric Boeckx (p.c.) informs me that the case can be made that even French in situ WHs undergo overt movement. This allows for an approach in terms of topicalizability, proposed first in Grohmann (1998), to certain intervention effects which have been noted for German by Beck (1996a, 1996b), Korean by Beck (1996a), Beck & Kim (1997), Japanese by Lee & Tomioka (2000), French by Chang (1997). See also section 5.4 below.
rily front in Serbocroatian—even if this fronting does not target the ω-domain. A more detailed discussion of the clause structure and relevant derivational details of Serbocroatian is beyond the scope of this chapter, and the same goes for Japanese. However, both languages are so-called “scrambling” languages, which we can pre-theoretically take to indicate that all arguments move out of their θ-positions into the φ-domain. I suggest tentatively that this kind of scrambling involves movement within the φ-domain.\footnote{As we will see below, I assume that the comparable type of “scrambling” in German targets the ω-domain. Details concerning the nature of movements and positions in scrambling structures for both Prolific Domains cannot be discussed here. I leave the present discussion deliberately vague, simply adopting the results from the Hagstrom-Bošković approach to the syntactic impact of multiple question interpretation.}

In Japanese, the WHs in multiple questions do not front the way they do in Serbocroatian. In fact, by the look of it, they do not necessarily need to move at all. For simplicity, though, and for comparative reasons, let us assume that WH1 sits in SpecTP at Spell Out, when it is the subject, and WH2 (when it is the object) in the relevant AgrP. (The following structures are further shortened for relevance.)

(45) a. Dare-ga nani-o katta no?
   who-NOM what-ACC bought Q
   ‘Who bought what?’

b. Ko je ko to kupio?
   who AUX what bought
   ‘Who bought what?’

(46) a. [CP [Q]-C [TP WH1 T [AgrP WH2 [Q]]] Agr [vP WH1 v WH2 [Q]]]] (PL)

b. [CP [Q]-C [FP [Q]-F [TP WH1 T [AgrP WH2 Agr [vP … ]]]]] (SP)
Suppose that for a Japanese question with WH1 > WH2, these structures are legitimate. Abstracting away from the fact that all WHs are fronted in Serbocroatian, and ignoring the position of the verb, for example, similar structures can be envisioned for this language (and, by extension, for Polish and Russian also; see Citko & Grohmann, forthcoming). What might be different are the labels of the projections. We are building on two premises, however, which are worth recalling. First, the hypothesis we are exploring at the moment is that these languages do not involve movement of any WH into the ω-domain. Thus, all WHs are at most in some position within the φ-domain. This might turn out to be a wrong hypothesis, of course, but once the entire picture is painted we will see some virtue coming out of this hypothesis. Moreover, this is simply reporting the result from Bošković’s (1998b) study. The second assumption concerns anti-locality. As I have argued in the previous chapters, a tripartite clause structure, following and extending an “intuitive” partition into Prolific Domains, is not unreasonable. In fact, we achieved some results by the proposal, and the Condition on Domain Exclusivity. One aspect of the theory of anti-locality is that Domain-internal movement is ruled out. This second assumption may also turn out to be wrong. But the goal of the present work is to take the framework seriously and explore its merits, predictions and consequences. We can thus assume that, all other things being equal, all WHs in Japanese and Serbocroatian are somewhere within the φ-domain at Spell Out, and the only copy each WH relates to resides within the base-generated θ-domain.25

This said, TP and AgrP are possible candidates to host non-scrambled WHs in Japanese. In this case, (46) is a plausible structure relevant for us. Merging [Q] low yields a PL interpretation and merging it high an SP reading, just as for Chinese. However, the order can be reversed. The only (rough) structures we can imagine are thus the following:

25 Again, for exposition I restrict the discussion to argument-WHs. Recall also that we explore the hypothesis that LF-movement of A’-elements can be dispensed with. As such, the position at Spell Out is also one of the (two) possible LF positions in these cases.
(47) a. Nani-o dare-ga katta no?
    what-ACC who-NOM bought Q
    ‘Who bought what?’
    (Japanese = (5b))

b. ta je ko kupio?
    what AUX who bought
    ‘Who bought what?’
    (Serbocroatian = (11b))

(48) a. \[
    [CP [Q]-C [AgrP WH2-[Q] Agr [TP WH1 T [\vP WH1 v WH2-[Q]]]]]
    (SP)
\]

b. \[
    [CP [Q]-C [FP [Q]-F [AgrP WH2 Agr [TP WH1 T [\vP ...]]]]]
    (SP)
\]

The derivations sketched in (48) proceed as follows. Following Stjepanović (1995, 1999), Bošković (1997a, 1998b, 1999, 2000a) and others, Wh-fronting in Serbocroatian is discourse-driven, possibly by some kind of a focus feature. Moreover, as Bošković (1997c), Boeckx & Stjepanović (1999) argue, this projection is within IP, what we call the \( \phi \)-domain. They identify it as AgrP, and I tentatively adopt the label. However, the choice of the label really plays no role: movement from the base position to this position must be in one step. It is possible that the “AgrP” in (48) is different from the “AgrP” in (46) above. What is not possible is that both projections are hosts of the relevant WH. I leave open details regarding a possible implementation of syncretic categories into the current framework (as proposed by Giorgi & Pianesi 1997; see also Zubizarreta 1998, Boeckx & Stjepanović 1999). Let us take (48) as it is for now—again, on the grounds that I am merely reporting other scholars’ work.

In sum, the proposal I have just sketched takes all WHs in Wh-in situ languages to be licensed maximally within the \( \phi \)-domain (following Bošković 1998b), and the assumptions are compatible with the larger framework presented here. In this sense, (48) does indeed account for the facts we have observed in section 5.3.1: merging \([Q]\) with
WH2, followed by “scrambling” over WH1, can only result in an SP reading, because both WHs are contained within the argument of the choice function variable. The same situation arises when [Q] is merged above all WHs, into FP, and no WH moves into the ω-domain.26

So far, I have suggested a way of implementing the main result of Bošković’s (1998b) extension of Hagstrom’s (1998) proposal regarding multiple questions and interpretation: Wh-in situ languages do not involve displacement of WHs into the ω-domain and as a result do not make a PL reading available if WH2 moves over WH1.

5.3.4.2 Singular Wh-movement languages

Our discussion of singular Wh-movement can be kept rather short, basically adopting the Hagstrom-Bošković analysis. If we assume that the approach sketched above is on the right track, we yield the following structures:

(49) Who bought what?

(50) a. \[ [CP [Q]-C [FocP WH1 Foc [TP WH1 T [vP WH1 v WH2-\{Q\}]\]] ] \] (PL)

b. \[ [CP [Q]-C [FocP WH1 Foc [FP \{Q\}-F [TP WH1 T [vP WH1 v WH2]]\]] ] \] (PL)

These are the only two possible derivations for the English question (49) given that a bare WH2 may never move over a bare WH1.27 Regardless of whether WH2 stays

26 Just to remind the reader, the PL reading is, of course, obtained by merging [Q] with WH2 and not move WH2 over WH1—again, conforming to the data reported above.

27 In fact, this statement can presumably be extended to all singular Wh-movement languages. The option to move a lower WH over a higher WH, if neither is D-linked, seems to be a property common to all singular Wh-movement languages. These also tend to be non-scrambling languages, that is to say, word order in these languages seems to be rather fixed, as opposed to many Wh-in situ languages (Serbocroatian, Japanese, Hindi).
really in situ at the point of Spell Out (here referring to its θ-position) or whether it moves into the φ-domain overtly, movement of low Q to C leaves behind the choice function variable which invariably only has WH2 in its scope and thus takes WH2 as its argument, returning a PL interpretation. This much can be achieved without any additional discussion.

We have seen, however, that merging [Q] high and moving WH1 into the ω-domain destroys the SP reading (Interpretive Superiority). This is sketched in (50b). Recall that an SP interpretation is the result of the choice function variable taking an argument that contains all WHs. When [Q] moves over WH1, however, the variable it leaves behind has only WH2 in its scope, thus ruling out the SP interpretation, if the boldfaced position of [Q] (the choice function variable) and the non-strikethrough copies of all WHs are the relevant elements to compute interpretation, which I assume to be the case.28

28 This is a simplified version of Bošković’s explanation, fitting not only the current framework but also making one less assumption. For Bošković, [Q] and WH1 target the same position, CP. (Note that he Bošković does not specify the position of high Q at all.) As Wh-movement is not necessary to derive interrogative interpretation, which is completely achieved by universal Q-movement, Bošković appeals to a unified LF structure (“interpretive operations at the interface should be as simple as possible,” Chomsky 1995b: 359): if a non-moved WH can be interpreted in situ, then a moved WH should also be interpreted in situ. With this assumption, the loss of SP interpretation by over Wh-movement must derive from something else. Bošković suggests that moving WH1 over [Q] results in a relativized minimality violation, further assuming that [Q], WH1 and C all bear the same feature, [Wh]. A number of questions arise with this move.

For one, it is not clear that [Q] really “interferes” with moving WH1 to C. Presumably [Q] moves to C before WH1 moves, and only the choice function variable might interfere (regardless of overt or covert Q-movement). In addition, it need not be the case that [Q] and WH1 really bear the same features. Under the approach endorsed here, they do not: [Q] types the clause in C and WH1 basically moves to FocP, where we take [Wh] to be an interrogative focus feature, a feature that asks for new information (as opposed to a declarative focus feature, one that introduces new information).

In sum, ruling out PL interpretation on the basis of a minimality violation is not feasible here, and neither is the assumption that the moved WH must be interpreted “in situ.” I see the same confusion alluded to earlier: “in situ” is employed in the literature with a certain degree of liberty. If we characterize Serbo-Croatian as a Wh-in situ language, this is certainly not the way we should understand the term, and neither does it apply to Japanese in a straightforward way, unless we analyse all scrambled word orders as base-generated. Basically, an “in situ” element should be one that is interpreted in its base, its θ-position, but clearly that makes the wrong prediction, even for “Wh-in situ” languages (which should really be referred to as zero Wh-movement languages): as in some in-

(continued...)
To sum up our current assumptions, we take \([Q]\) to move to \(C\) to type the clause, and this movement takes place either overtly or covertly across all languages. We further assume that WHs in English move to check a Wh-feature,\(^{29}\) which Japanese WHs, for example, do not bear.\(^{30}\) Following much research since Horvath (1986) and Brody (1990), we take FocP to be the canonical landing site of overtly moved WHs, which check \([Wh]\). As we will see below, we can conceptualize this assumption on the grounds that the kind of focus targeting FocP is basically the same kind of focus employed in Wh-questions: new information—in interrogatives this information is inquired, in declaratives it is asserted. Languages simply differ as to whether they must express either type of focus by displacement to a particular position or not.

stances, WH1 moves to TP and WH2 to AgrP, obligatory interpretation of WH1 “in situ” results in the configuration in which the choice function variable would always scope over all WHs—in which case the SP reading would be the only one available:

\[
(i) \quad [CP \quad [Q]-C \quad [TP \quad WH1 \quad T \quad \ldots \quad [AgrP \quad WH2\{Q\} \quad Agr \quad \ldots \quad [vP \quad WH1 \quad v \quad \ldots \quad WH2\{Q\}]]]
\]

Strict “in situ” interpretation takes the \(\theta\)-positions of both WHs. But the choice function variable is computed over the trace \([Q]\) leaves behind from moving to \(C\) from the intermediate Agr-position, \textit{not} from the original \(\theta\)-position, as indicated here (and, as far as I can tell, assumed by Bošković also).

Furthermore, this approach presupposes that all elements moved into the \(\omega\)-domain for discourse reasons (and by extension, those that target a \(\phi\)-position for the same reason, as in Serbocroatian) can only ever be interpreted in their base position. Clearly, this is not the direction we want to pursue, scope being only one reason. Admittedly, the Q-approach sketched here leaves many questions open, but if we take the Hagstrom-Bošković approach seriously and scrutinize some predictions it makes—the goal of this chapter—we can only touch on discrepancies in passing. I do exactly this and leave the remaining details to be worked out in future research.

\(^{29}\)Addressing the phenomenon of D-linking in any detail would take us too far afield. The predictions of this, and the Hagstrom-Bošković, approach concerning D-linked WHs is apparent: if, all other things being equal, an SP reading is always possible we might be tempted to not evoke the \(\omega\)-domain as the landing site of D-linked WHs. That D-linked Wh-phrases should differ in their syntax from bare ones is not unreasonable.

\(^{30}\)The reason for this might lie in the morphological make-up of Wh-words, an idea that might want to be extended to all languages. That is to say, whether a language moves one, all or no WH into the \(\omega\)-domain could be related to some property other than a stipulated correlation between form and function, but that leads us too far here.
5.3.4.3 Interlude: German

After reviewing the Hagstrom-Bošković approach to zero and singular Wh-movement languages and slightly modifying some details of the analysis fitting current assumptions, it is now time to address the German paradox. As we have seen, if loss of PL reading is indeed the effect of moving a lower WH over a higher one, the only way to account for why German WH2 > WH1 questions retain their PL reading (no Interpretive Superiority effects) is to resort to a Q-stranding mechanism which would allow WH2 to scrambling over WH1, but leave the Q-morpheme in situ. This is indeed the line of thought taken by Hagstrom (1998) and Bošković (1998b).

Under the assumption that English and German only differ in that the latter allows WH2 to move over WH1, targeting FocP, the resulting configuration would look like (51) where, unlike (48), WH2 and [Q] move separately:

\[
\text{(51) } \left[\text{CP} \left[\text{Q}\right]-\text{C} \left[\text{FocP WH2 Foc} \left[\text{TP WH1 T} \left[\text{AgrP WH2} \left[\text{Q} \text{ Agr} \left[\text{vP ...} \right] \right] \right] \right] \right] \right] \right] (PL)
\]

In this case, Hagstrom (1998) argues, Q-movement is unaffected by Wh-movement and WH1 will remain outside the scope of the choice function variable. Apart from other technical problems, my main objection to this line of explanation is that Q-stranding seems to be an ad hoc mechanism, fixing a mysterious state of affairs for a scrambling language—which differs in this respect from other scrambling languages. It is not clear, for example, why Q-stranding should be allowed only in German, but not in Japanese or Serbocroatian (as illustrated by the contrasts in (26) and (27), respectively). We know that not only do these languages show reordering effects of non-interrogative elements, but they also allow WHs to move freely over one another. If Q-stranding were an option made available by Universal Grammar, we would expect to see it in more cases than German—aside from obvious learnability issues. Moreover, the core case for which
Hagstrom proposes this option involves apparent “intervention effects” as noted by Beck (1996a, 1996b) and discussed in much work since (e.g. Grohmann 1998, Boeckx 1999b, Pesetsky 2000, Miyagawa, in press). I will suggest an alternative approach in the next section and will illustrate the relevant data there. It suffices to anticipate that a non-stranding approach of [Q] in these cases is superior to Hagstrom’s analysis primarily as it allows us to capture these “Beck effects” in a more natural way than Beck suggests.

Another problem [Q], and stranding of [Q], might pose has to do with structural or even locality complications. Hagstrom suggests that [Q] is a head, supported by the fact that its overt counterpart for the most part is a monosyllabic element (such as *no* in Japanese or *-GA* in Okinawan). However, movement to C, regardless of its launching site (from WH2 or FP) runs into trouble with locality conditions on head movement. While Hagstrom addresses this issue, there is another way of inserting [Q], particularly when it attaches to WH2. I suggest that [Q] is adjoined to the lowest Wh-phrase. As such it may still be a “head” or, basically, an $X^{0\text{max}}$ element, very much like a clitic. If [Q] starts out adjoined to WH2, it must move with WH2 and cannot be stranded on structural grounds. This might not be a necessary consequence, though. I will thus (tentatively) assume that low Q looks as follows (where X is the appropriate categorial label of [Q], possibly D):

\[
\begin{align*}
(52) & \quad \text{DP} \\
& \quad \text{DP} \quad X^{0\text{max}} \\
& \quad \text{WH2} [Q]
\end{align*}
\]

Hagstrom also discusses Q-stranding in some (head-final) Wh-in situ languages. A discussion of these data is beyond the scope of the present investigation, however, especially as many ancillary assumption hinge on the analysis, such as overt vs. covert Q-movement and a structural analysis for head-final languages.
If this is the structure, Q-stranding is not an option.\textsuperscript{32} If \([Q]\) is an \(X^{\text{omax}}\), it does not matter much where within FP high \(Q\) is inserted. Movement to the head of CP, which share the same Prolific Domain with FP, is possible in either case: Exclusivity is not jeopardized by moving it to C (see section 2.3.2). I will postpone a discussion of locality concerns for movement of low \(Q\) to C to future work, which will also address other “long local” movement operations, such as “scrambling” of the sort necessary to derive (48b), for example. See chapter 6 for initial pointers.

5.3.4.4 Multiple Wh-movement languages

We have seen good reasons to assume that Serbocroatian WHs do not undergo the same kind of movement that Bulgarian ones do, but rather pattern like Japanese “in situ” WHs. Rather than implementing Q-stranding and maintaining that German multiple Wh-questions are formed just like English ones, qua singular Wh-movement, I suggest an alternative explanation.

We have seen above that in singular, but not zero, Wh-movement languages, the SP reading of bare WHs in a multiple Wh-question is disallowed. Given the correlation between the scope of \([Q]\) and the position of Wh-phrases, it becomes clear why this should be the case. Simply, the configuration needed for an SP reading, in which the choice function variable left behind by Q-movement has scope over both Wh-phrases, never arises. Wh-movement always displaces one WH to a position out of the scope of the choice function variable, even if \([Q]\) is generated high (cf. (50b)).

Similarly, in multiple Wh-movement languages, Wh-phrases escape the scope of \([Q]\), which destroys the SP reading configuration. On a par with the structures above, we can sketch this as in (54), where \([Q]\) is merged low in (54a) and high in (54b):

\textsuperscript{32} Again, this means that Hagstrom’s analysis for Japanese, Okinawan, Sinhala and others must be revisited. It remains to be seen how compatible the present assumptions are with the original idea.
Apart from FocP as the canonical position for Wh-phrases in singular Wh-questions or for WH1 in multiple ones, I assume that WH2 targets FP, the specifier of the projection hosting high [Q]. I leave aside more specific motivation, such as an additional focus-like feature or the licensing of “point of view” (cf. Uriagereka 1995a, 1995b; see also Agr1P of Cardinaletti and Roberts 1991 or FinP of Rizzi 1997 for a low C-position). Rather, I build on the intuitive connection between F and Wh-properties in the sense that clause-typing properties must be identified in the finer grained ω-domain of the clause, and if F may host [Q], it might not be an unreasonable landing site for WHs also. We can employ (54) as the rough structure for all relevant questions in multiple Wh-movement languages: WH2 > WH1 is not attested, as the languages in question are sensitive to Syntactic Superiority, and after the discussion about the properties of German

---

Roussou (1998) investigates multiple Wh-questions in Classical Greek and comes to the conclusion that all WHs are fronted into the ω-domain here also. She further identifies two separate C-projections as potential landing sites, providing not only additional conceptual and empirical arguments against a multi-Spec analysis for multiple fronting of WHs (Koizumi 1994, Mulders 1996, Richards 1997), but also evidence in favour of separate projections in the ω-domain. Roussou does not settle on the labels of these heads, but in recent work, she lays out an articulate C-structure which is compatible with the one assumed here, all details aside (Roussou, in press).
multiple Wh-questions in the next two sections, we will have a natural implementation at
our disposal for motivating ω-movement of all (fronted) D-linked WHs as well. Thus, if
all Wh-elements in these languages target the ω-domain, availability of SP interpretation
in multiple Wh-questions is not even expected.

This discussion suffices for the more “traditional” multiple Wh-movement lan-
guages. The reason is simple: as I will argue next that German is also a multiple Wh-
movement language, albeit with slightly different properties, I will address some issues
related to the ω-domain anyway. In addition, as we will see in section 5.5, the trigger to
move all WHs in German into the ω-domain turns out to be different from Bulgarian or
Romanian, for example, and I will concentrate on these.

5.4 **German = Bulgarian: Multiple Wh-Movement**

It is now time to subject the claim that German is a multiple Wh-movement lan-
guage to some scrutiny. If German really were a multiple Wh-movement language, as the
data in (32) and (33) suggest on interpretive and, by extension, typological grounds, we
might expect certain grammatical reflexes, such as syntactic evidence for this very high
occurrence of all Wh-elements as can be found in Bulgarian, for example. I show in this
section that such evidence can indeed be found and revolves mainly around the relative
placement of adverbs and quantifiers with respect to the Wh-phrases in the clause. This
suggests that German is like Bulgarian in that both languages obligatorily move all WHs
into the ω-domain. The properties of the Wh-elements involved in multiple questions will
be scrutinized in the next section; in that section, I consider the equally novel proposal
that German is (also) like Italian in that no “real” multiple questions can be formed!

---

34 Recall that the traditional assumption regarding multiple questions in German is pretty
much on a par with English: one WH undergoes movement to FocP, any other WHs stay
“in situ” (i.e. the relevant Agr-projections, somewhere within the φ-domain in our terms).
5.4.1 German as a Multiple Wh-Fronting Language

Given what we have seen so far, under a multiple Wh-movement analysis of German we would expect all Wh-phrases to appear in some C-related position at the point of Spell Out, i.e. all Wh-phrases obligatorily move into the $\omega$-domain in the overt syntax. Following, but not necessarily relying exclusively on, the assumption that FocP is the default host for overtly moved Wh-phrases, FP is a possible host for [Q]. Some technical details and predictions will be addressed in section 5.5.3. Let us see now how we could justify both WHs moving there in the overt syntax.

Apart from the syntactic evidence introduced momentarily, there are other “grammatical reflexes” worth mentioning. One is semantic in nature and follows the discussion, namely the role of SP/PL interpretations following the Hagstrom-Bočković line. Pushing this approach further turns this point into a typological concern, as suggested throughout. Another is pragmatic in nature and was first explored by Grohmann (1998). It focuses on the observation that the discourse requirements, which allow a multiple Wh-question in German to be felicitous, are different from those applying to English or even the more closely related Dutch. This will be explored in section 5.5.

In essence, I propose a structure like (54) above for any multiple question in German, which might now be worthwhile considering in more detail.\(^{35}\)

In other words, all Wh-phrases move into the $\omega$-domain in the overt syntax and regardless of whether [Q] is merged high or low, it can never scope over all WHs—leaving only a PL interpretation. As the placeholder ‘…’ between FocP and FP indicate, there is room for material between the two WHs. Following Rizzi’s (1997) structure, this room

\(^{35}\)I ignore Apparent Superiority constellations or a treatment of D-linking for the purposes of presentation, but also in anticipation of the next section.
is reserved for topics (cf. (36)). This suggests that only topics might be placed in between two Wh-elements, and I will capitalize on this prediction in the following. (55) is the structure relevant here, where I indicate the highest possible “second position” as the target for the finite verb, which I will largely ignore in the discussion.\(^{36,37}\)

\[(55)\]

\[
\begin{array}{c}
\text{CP} \\
\downarrow \\
\text{C'} \\
\downarrow \\
\text{[Q]-C}^0 \\
\downarrow \\
\text{FocP} \\
\downarrow \\
\text{WH1} \\
\downarrow \\
\text{V-...-Foc}^0 \\
\downarrow \\
\text{TopP}^* \\
\downarrow \\
\text{Top'} \\
\downarrow \\
\text{Top}^0 \\
\downarrow \\
\text{FP} \\
\downarrow \\
\text{WH2} \\
\downarrow \\
\text{F'} \\
\downarrow \\
\text{F}^0 \\
\downarrow \\
\phi\Delta \\
\end{array}
\]

\(^{36}\) The situation in the Slavic multiple Wh-movement languages is quite different, of course. I address the difference, and a way to capture it, in section 5.5.3. To anticipate, one major difference regards topic positions: while topics may target a position in between the two WHs (a rather low \(\omega\)-position), all topics precede WH in Bulgarian.

\(^{37}\) Regarding the V2 position, I assume that whatever derives this phenomenon will ensure proper placement in these cases also. I indicate the two different launching sites for [Q] with the dashed lines: the one coming off \(F^0\) represents high Q and the one originating within the \(\phi\)-domain low Q. An ancillary assumption is that Q-movement from low, WH2-joined position must take place from within the \(\phi\)-domain. I cannot discuss the details, but here a minimality effect might be evoked, or possibly economy: given that F is a potential host for [Q], movement of the WH2-[Q] complex past F must be ruled out.
5.4.2 Quantifier Interaction

The two WHs in a German multiple question must not only be non-adjacent, but other elements may be placed in between—in some cases even quite a lot of material (a matter I return to in section 5.5.3). Beck (1996b) observes that apart from non-quantificational elements, some quantifiers may appear in between the two Wh-phrases, whereas others may not. I cannot review the details of her analysis here nor delve further into patterns involving other elements than two WHs (cf. Grohmann 1998, Hagstrom 1998, Boeckx 1999b, Pesetsky 2000, Miyagawa, in press). I concentrate on the type of intervention effect that arises from placing a particular quantifier in between two WHs. The essence of Beck’s proposal is that negative quantifiers induce a blocking barrier: the WH left in situ at Spell Out must move to CP at LF, and if blocked it cannot do so.

As (56) shows, many different elements can occur between two WHs—even more than one, as the parenthesized adverb indicates—as well as follow them. This is not so for all quantificational phrases, though:

(56) a. Wer hat Maria/seine Freundin/einen Porsche (gestern) wo geküßt?
   b. Wer hat wo Maria/seine Freundin/einen Porsch (gestern) geküßt?
   who has where Maria/his girlfriend/a Porsche (yesterday) kissed
   ‘Who kissed Maria/his girlfriend/a Porsche (yesterday) where?’ (German)

(57) a. Wer hat alle Bücher wo gekauft?
   who has all books where bought
   ‘Who bought all (the) books where?’
   b. * Wer hat kein Buch wo gekauft?
   who has no book where bought
   ‘Who bought no book where?’
   (German)
Universal quantifiers seem to be able to appear in between the two WHs, while negative ones do not. If all WHs must end up in CP at one point, one possibility is the barrier-approach. Note that nothing rules out co-occurrence of two WHs and a negative quantifier: if the quantifier follows all WHs, the question is well-formed.

(58) a. Wer hat gestern wo alle Bücher gekauft?
   who has yesterday where all books bought
   ‘Who bought all books where yesterday?’
   b. Wer hat gestern wo kein Buch gekauft?
   who has yesterday where no book bought
   ‘Who bought no book where yesterday?’

   (German)

However, not only negative quantifiers are prohibited from “intervening.” And neither is the class of possible interveners restricted to universal quantifiers.

(59) a. Wer hat viele/die meisten/mehr als drei Bücher wo gekauft?
   who has many/the most/more than three books where bought
   ‘Who bought many/the most/more than three books where?’
   b. * Wer hat wenige/höchstens drei/weniger als drei Bücher wo gekauft?
   who has few/at-most three/fewer than three books where bought
   ‘Who bought few/at most three/fewer than three books where?’

   (German)

And again, if the quantifiers follow the WHs, everything is alright:

(60) a. Wer hat wo viele/die meisten/mehr als drei Bücher gekauft?
    who has where many/the most/more than three books bought
    ‘Who bought many/the most/more than three books where?’
   b. Wer hat wo wenige/höchstens drei/weniger als drei Bücher gekauft?
    who has where few/at-most three/fewer than three books bought
    ‘Who bought few/at most three/fewer than three books where?’

   (German)
These contrasts allow for the following generalization. Monotone increasing quantifiers may appear in between two Wh-phrases, decreasing ones may not.\(^{38}\) As laid out above, we do not share Beck’s assumption on LF-movement of WHs here. Hence, any resort to an intervening or blocking barrier must be discarded immediately, as the lower WH does not raise at LF. There is an additional problem with such an approach, which basically boils down to a reintroduction of (artificial) barriers, a configuration which a minimalist approach would rather deduce than stipulate.

Under the hypothesis that all Wh-phrases overtly move into the \(\omega\)-domain, such as FocP and FP, we might be able to go another route towards such a deduction, albeit in quite different terms. We can restate the “intervention effect.” Given our architecture of this Prolific Domain, as sketched in (55), iterative topic phrases may appear in between the two WHs. It follows that any material intervening between the two WHs must thus target TopP. A natural explanation for the ungrammaticality of the strings shown here, namely \(WH > \text{decreasing quantifier} > WH\) would be that decreasing quantifiers cannot be topicalized, that is, they cannot ever target SpecTopP. This is the route I pursue here.

\subsection*{5.4.3 Topicalizability}

If TopP is the projection ranked between the two WHs, we expect only topicalizable material to occur between two Wh-phrases. As it happens, increasing quantifiers fit this condition, while decreasing ones do not. Consider the following data:

\(^{38}\) While noted in passing by Beck (1996b: 30), I extend this classification in Grohmann (1999b, to appear), building on Grohmann (1998), where all kinds of constructions were considered, including Apparent Superiority structures and the combination of arguments and adjuncts (see also Citko & Grohmann 2000, and Boeckx 1999b, in progress for quite similar effects found in French II).
(61) a. **VIELE Bücher hat Peter gestern gelesen**, (nicht ALLE).
    many books has Peter yesterday read (not all)
    ‘**MANY books Peter read yesterday**, (not ALL).’

b. **Viele Bücher hat Peter gestern gelesen.** (Er ist fleißig.)
    many books has Peter yesterday read (he is hard-working)
    ‘**Many books, Peter read yesterday.** (He is a hard worker.)’ **(German)**

(62) a. **WENIGE Bücher hat Peter gestern gelesen**, (nicht VIELE).
    few books has Peter yesterday read (not many)
    ‘**FEW books Peter read yesterday**, (not MANY).’

b. * **Wenige Bücher hat Peter gestern gelesen.** (Er ist faul.)
    few books has Peter yesterday read (he is lazy)
    *‘**Few books, Peter read yesterday.** (He is lazy.)’ **(German)**

Commonly, the first element in a non-subject-initial V2 matrix clause is referred to as the “topic” of the clause, a loose notion that has sometimes also been referred as “focus.” Thus, as we can see in (61a), this can be a contrastive position. However, decreasing quantifiers may not appear in initial position, unless they are contrastively used.

Take the following questions:

(63) a. What happened?

b. What did Peter do yesterday?

c. Did Peter read few/many books yesterday?

d. How many books did Peter read yesterday?

(63a) asks for new information in a completely out-of-the-blue context. We will get back to that momentarily. (63b) also asks for new information, but in a slightly more restrained way, the discourse already establishes the participant and a time specification related to the event inquired about. In a felicitous answer in German, this new information appears in a default focus position, where the type of focus is “information focus” (see É. Kiss 1998 for discussion and references) and the “default” position is rather low
in the structure (see Abraham 1995 and references on theme/rheme structures). In this case the value for ‘what’ is the new information. Neither construction in (61) or (62) can be used to answer this question. Disregarding pitch accents, the only word order of a felicitous answer is the one in (64):

(64)  
   a. Peter hat gestern viele Bücher gelesen.  
      Peter has yesterday many books read  
      ‘Peter read many books yesterday.’  
   b. Peter hat gestern wenige Bücher gelesen.  
      Peter has yesterday few books read  
      ‘Peter read few books yesterday.’

It is not implausible that the position of viele/wenige Bücher ‘many/few books’ is not a particular syntactic projection, such as FocP, but related to the information structure in German. I shall largely ignore this point (see Abraham 1995, Meinunger 2000 and many others).

What is of interest here is that (61) and (62) can be used as felicitous answers to questions, namely to something like (64c,d). Here, the value of the new information asked for is specified, either in form of a yes-no question or a Wh-question. Along with many other possible replies, (61a) and (62a) can be uttered in response to (64c), so as to deny the presupposition that Peter read few books (61a) or many books (62a). In this sense, the answer contains a contrastive element as part of the new information.

(61b) and (62b) are simple topic structures. There is no contrastive stress, and the fronted elements can only pick up a referent already established in the discourse. Both are infelicitous responses to any of the four questions in (64).\(^{39}\) I will use the notions topic, focus and contrastiveness in this sense.

\(^{39}\)(61b) can be used to respond to a question like Who read many books?, but the ungrammatical (62b) cannot serve as an answer to Who read few books?.
Under a single COMP structure, we do not need to distinguish the landing site of the first XP in German V2 clauses further. Whether it is a focalized or a topicalized constituent, or something else (such as a Wh-phrase) does not really matter. In a more articulated system, we might want to reserve some projections to certain elements. This is one part of the motivation to split this layer into separate projections. On the basis of Italian, Rizzi, for example, proposes a unique FocP hosting contrastively focalized phrases and recursive TopP for topics, where one distinguishing property between the two is resumption: a topic can be coreferent with a clause-internal clitic (see chapter 4), a focalized element cannot.

To come back to the issue at hand, the difference between (61) and (62) is that some elements can be fronted into the first position as “pure” topics (see fn. 39) as well as contrastively used elements, while others can only be used contrastively. Decreasing quantifiers fall into the latter category, and these are also the elements that cannot appear in between two WHs. I refer to the elements in this class as “non-topicalizable material.”

This said, it is interesting to note that once contrastively stressed, the quantifiers identified by Beck as “interveners” become suddenly good, even clearer with nur ‘only’. While I abstain from pinning down the exact position of contrastive elements (whether it is an additional topic or focus projection, for example), the absence of contrast in the following data casts serious doubt on a barrier-driven intervention effect.

(65) a. Wer hat **VIELE** Bücher wo gekauft?
    who has many books where bought
    ‘Who bought MANY books where?’

b. Wer hat (nur) **WENIGE** Bücher wo gekauft?
    who has (only) few books where bought
    ‘Who bought (only) FEW books where?’ (German)
5.4.4 Adverb Placement

Beyond the intervention effects of the quantifiers Beck looked at (and additional ones), a similar effect can be found with adverbs: some adverbs may appear in between two Wh-phrases, others may not. Again, it can be independently shown that the illicit configurations arise with non-topicalizable adverbs.

Manner adverbs, rather low in the structure, must follow the object, unless the object is focused:\(^{40}\)

\[(66)\]

a. Peter hat das Buch kaum/komplett gelesen.

  Peter has the book barely/completely read

  ‘Peter <barely> read the book <completely>.’

b. Peter hat kaum/komplett das BUCH gelesen. \(\text{(German)}\)

In combination with multiple Wh-phrases, some of these adverbs may not appear in between the two. This might strike us as a surprise, given that focus is commonly construed with Wh-phrases (see also section 5.5). Thus, if this class of adverbs can precede a focalized object as in (66b), it should also be able to precede a (focalized) WH. (67b,c) show that this is not the case. This follows, however, if both Wh-phrases are in the \(\omega\)-domain (as opposed to a non-interrogative object as in (66)), making only room for topicalizable material to appear in between them, and such adverbs may not be topicalized.

\(^{40}\) Without further analysis of adverbial positions, we can take these elements to adjoin to vP or AspP, for example.
(67) a. Wer hat was kaum/komplett gelesen?
    who has what barely/completely read
    ‘Who barely read what?’

b. * Wer hat kaum/komplett was gelesen?
c. Was hat wer kaum/komplett gelesen?
d. * Was hat kaum/komplett wer gelesen? (German)

Indeed, these adverbs, as opposed to sentential adverbs, for example, are not topicalizable. (68) illustrates the contrast, where the object is topicalized also to force a topic-reading of the fronted adverb, rather than a high occurrence in the T-domain. In these cases the fronted adverb is presumably in SpecTopP and the verb sits on the Top-head.41

(68) a. * Kaum/Komplett hat das Buch der Peter gelesen.
    barely/completely has the book the Peter read
    ‘Barely/Completely, Peter read the book.’

b. Wahrscheinlich/Gestern hat das Buch der Peter gelesen.
    probably/yesterday has the book the Peter read
    ‘Probably/Yesterday, Peter read the book.’ (German)

Apart from other combinatorial possibilities, the same contrasts can be found in embedded contexts. Here I indicate with brackets (im)possible positions for the relevant adverbial material. Moreover, the adverbs in question do not underlie some condition that bars co-occurrence, as it is perfectly fine to use both wahrscheinlich ‘probably’ and/or gestern ‘yesterday’ in combination with kaum ‘barely’ and/or komplett ‘completely’.

41 The presupposition is that pre-subject scrambled elements in German target a position inside the ω-domain, here roughly equated with TopP, unlike Japanese or Serbocroatian, as hinted at above.
(69) …, daß *kaum/komplett* der Peter *wahrscheinlich/gestern* das Buch *wahrscheinlich/gestern* *kaum/komplett* gelesen hat.
‘… that probably/yesterday/*barely/*completely, Peter read the book’ (German)

5.4.5 Summary

In sum, certain elements are inherently not topicalizable—for reasons that we did not discuss here—and as such not predicted to appear in between two Wh-phrases under the assumption that both Wh-phrases have moved very high, into the $\omega$-domain, making room for topics to occur in between. This prediction is empirically borne out. These facts not only support our hypothesis that German moves all WHs into the $\omega$-domain (and thus allows us to classify multiple Wh-question formation into a clean tripartition), it also captures (one part) of an interesting intervention effect noted by Beck as well as additional effects which are not captured under a barrier-style intervention analysis.\(^{42}\)

5.5 German = Italian: Categorical Absence of Multiple Wh-Questions?

Now that we have supported the proposal that German is a multiple Wh-movement language, I would like to stretch comparisons. Essentially German could now be argued to form multiple Wh-questions very similar to Bulgarian in that all WHs target a left-peripheral position, tentatively identified as FocP and FP. I will now suggest that

\(^{42}\) Another part of the “Beck effect” concerns “intervention” of the same material in between split partitives or with *was für*-split constructions. See Grohmann (1998) for discussion. A further aspect, noted by Beck (1996a) and Beck & Kim (1997), concerns similar effects in Korean. While there are some differences between the “intervention” effects in Korean and German, much of what has been said here can also be applied. In this regard it is worth mentioning that Lee & Tomioka (2000) find the same effects in Japanese also, first noted in Hoji 1985, and account for both Japanese and Korean “intervention” effects in roughly the same way I suggest for German: non-topicalizable material cannot occur between two WHs. Their emphasis is semantic (and prosodic), while the present one is syntactic. Moreover, Boeckx (in progress) applies the approach in terms of non-topicalizability for similar effects found in French Wh-in situ structures.
German (also) behaves like Italian, a language that cannot form multiple Wh-questions at all with the purpose of asking for new information.

5.5.1 Discourse-Restricted Quantification

One striking difference in the formation of a multiple Wh-question in German, as opposed to English, for example, is what I dubbed “Discourse-Restricted Quantification” (DRQ; Grohmann 1998). DRQ describes a felicity condition posed on the use of more than one WH by the context. Take a WH-phrase to represent a set of individuals. DRQ requires not only that this set is a non-empty set for both WHs, but also that the discourse facilitates possible reference to all such sets.

To exemplify, consider the following two contexts:

(70) **Context I:** A customer comes to a newsstand and just sees three customers leaving in different directions, each fiddling with their purchases. After asking the salesman whether he sold anything interesting, the salesman replies: “I can’t believe it. I just sold a Dutch porno magazine, a Japanese comic and a Bosnian newspaper.”

a. Who bought what?

b. Wer hat was gekauft?

    who has what bought

    ‘Who bought what?’ (German)

(71) **Context II:** A jeweler comes home to his wife and says excited: “I had a great day. I sold a platinum watch, a gold necklace and a brilliant ring.”

a. Who bought what?

b. # Wer hat was gekauft?

    who has what bought

    ‘Who bought what?’ (German)
It appears to be the case that German can only allow for a felicitous binary Wh-question if the set of individuals is part of the common ground between speaker and hearer, introduced in the discourse. Thus, in (70) speaker and hearer can assume that the three previous customers are the buyers of the three items sold. In (71), however, the set of possible buyers is too big to facilitate a possible pairing of buyers and items sold. This description is very intuitive and can remain as such. What I capitalize on is the difference with English, where (71a) is a perfectly reasonable question and a potential answer would be something like (72):

(72) A businessman bought the watch, an old lady the necklace and a young bride the ring.

A tentative definition of DRQ would run as follows:\(^{43}\)

(73) *Discourse-Restricted Quantification (DRQ)*

Questions involving two Wh-expressions are well-formed if the value of both Wh-expressions is determined by the context; determination of values is satisfied by providing a set of at least two possible referents in the discourse.

In previous work, I employed this observation to further motivate movement of all WHs into a high, discourse-related projection. In a sense, the two WHs in a multiple Wh-question could be construed as “Wh-topics” of sorts. As mentioned briefly in chapter 2, Wh-topics are not contradictory per se. Wu (1996, 1999), and Tang (1988) also, argue that fronted WHs in strict, non-scrambling Wh-in situ languages like Chinese are actually topics.

\(^{43}\) This definition is slightly changed from the original in Grohmann (1998: 19). See that work for ample illustration of different contexts, and with a variety of Wh-expressions (arguments and adjuncts alike).
But note what DRQ essentially says: all WHs in a German multiple Wh-question must be D-linked. While a WH in an information question represents a novel set of individuals (“new information”), the set of individuals represented by D-linked WHs is to some degree known (“old information”). I will now try to make some sense out of this, apparently contradictory, state of affairs.

5.5.2 Italian Wh-Questions

In a discussion of Subjacency violations, Rizzi (1978) mentions in passing that Italian does not allow multiple Wh-questions at all. Calabrese (1984), further explored in Calabrese (1992), offers an explanation for the lack of multiple constituent questions in Italian. A multiple Wh-question, if not used as an echo or quiz question, asks for “new information” regarding all WHs; the set of individuals contained in each WH is new. New information in a declarative sentence is construed with non-default sentential stress—or focus. The type of focus relevant for a felicitous reply to a question is information focus, not contrastive focus (see discussion in section 5.4.3 above). In Italian, this position for information focus is unique. If every WH in an interrogative clause must correspond to a focus in the declarative answer, and if the relevant focus in Italian is unique, it follows that multiple Wh-questions are ill-formed: the multiple information requested cannot be realized in the answer.

Consider the following data (from Calabrese 1984: 67; glosses added—KKG):

(74) a. * Chi ha scritto che cosa?
   who has written what
   ‘Who wrote what?’

b. * Chi è partito quando?
   who is left when
   ‘Who left when?’

(Italian)
(75) a. * MARIO ha scritto una LETTERA.
Mario has written a letter
‘MARIO wrote a LETTER.’ 
b. * FRANCO è partito alle CINQUE.
Franco is left at five
‘FRANCO left at FIVE O’CLOCK.’

The statements in (75) should be possible answers to the questions in (74). As the translations show, this is indeed the case in English. As the stars indicate, however, in Italian neither one is well-formed: a multiple constituent question is as ungrammatical as a multiple focus structure (where focus is marked by capital letters).

This state of affairs is independent of a SP or PL reading, as Calabrese shows:

(76) * MARIO ha scritto una LETTERA, FRANCO un ARTICOLO, CARLO un LIBRO…
‘MARIO has written a LETTER, FRANCO an ARTICLE, CARLO a BOOK…’

Again, the two languages differ. Notice that the intended focus interpretation of (76) is not contrastive, which is fine in Italian, but information. Leaving aside a deeper discussion of prosodic properties, focus and information structure (see e.g. Rooth 1985, Vallduvi 1992, Winkler 1997 and references cited), we can detect a similar effect in German. That is, the analogue of (76) can only be used contrastively. It seems that German and Italian pattern alike in this respect.

If this turns out to be the case, a multiple Wh-question should not be able to be formed in German. However, German has the option of marking contrastiveness in two positions: by fronting the relevant element or by assigning it heavy stress in low position. Thus, if a question is formed that asks for contrastive information, and if it does so in a syntactically well-formed way, a possible answer is also well-formed syntactically.
What D-linking a WH does is pick out a set of individuals restricted out of all possible individuals to those specified by the discourse—the DRQ condition I suggested earlier. We could then hypothesize that the fact that even bare WHs in German multiple questions must be basically D-linked is related to the absence of multiple Wh-questions in Italian. German is like Italian in that neither allow multiple information focus and hence do not allow multiple information questions.\footnote{I would like to thank Željko Bože ković for pointing out the obvious and discussing this idea with me.} This still leaves room for DRQ: on the one hand, we can use this term to distinguish a condition of D-linking on all WHs (in German), while employing D-linking to those elements commonly identified as such (e.g. \textit{which}-phrases in English; see Pesetsky 1987). On the other hand, German and Italian apparently differ in this respect. Calabrese notes also that a multiple constituent question in Italian does not get better if the WHs are D-linked.\footnote{It is a well-known fact that there is a large degree of variation among Italian dialects. Not all speakers agree with Calabrese’s judgements, and I simply report the data as presented by Calabrese (1984). Even for those speakers, though, an interesting contrast arises with respect to contextual information. Andrea Gualmini (p.c.) notes that in the situations depicted in (70) and (71), the most appropriate way to inquire both constituents (the pairing of buyers and items sold) in Italian is to use a WH for one and a pronoun for the other, the equivalent of \textit{Who bought it all/them}?—however, only in Context I of (70) can a D-linked multiple Wh-question be asked felicitously. This is very reminiscent of the German DRQ condition. I am aware that for a satisfactory account, other matters need to be resolved first and thus have to leave many interesting issues for further research.}

(77) * Quale ragazza ha dato un bacio a quale ragazzo?
which girl has given a kiss to which boy
‘Which girl kissed which boy?’ \textit{(Italian; Calabrese 1984: 67)}

I leave this discrepancy to future research and concentrate on the common properties of German and Italian multiple questions.

If then even a multiple question with two bare WHs in German is a D-linked question, this might provide us with a clue as to why Syntactic Superiority may be violated in German: the two Wh-elements are actually D-linked, which—for reasons that are...
beyond the scope of this discussion—may circumvent the Superiority Condition. On the other hand, German requires all D-linked WHs to move into the \( \omega \)-domain, unlike English. As a result, German multiple Wh-question always force a PL reading, where the individuals of each set are contrasted with one another.

The clue to all this comes from pairing the order of WHs and the order of the values for each WH in the answer. Consider the following:

(78)  
\begin{align*}
\text{a.} & \quad \text{Wer hat was gekauft?} \\
\text{b.} & \quad \text{Was hat wer gekauft?} \\
& \quad \text{‘Who bought what?’} \quad \text{\textit{(German)}}
\end{align*}

(79)  
\begin{align*}
\text{a.} & \quad \text{Der Peter hat das Buch gekauft, der Martin hat das Fahrrad gekauft…} \\
\text{b.} & \quad \text{Das Buch hat der Peter gekauft, das Fahrrad hat der Martin gekauft…} \\
& \quad \text{‘Peter bought the book, Martin bought the bicycle…’} \quad \text{\textit{(German)}}
\end{align*}

If the order of WHs in the question is WH1 > WH2, the appropriate answer follows the same order for the values of each WH. However, if the order in the question is WH2 > WH1, the order in the answer is analogous (see also É. Kiss 1993: 103f., fn. 6 for Hungarian). As we have established above that the first position in a declarative matrix clause can be a contrastive position, and as we know that contrastive focus can also be assigned in situ, it follows that not only a PL reading is available, but also the two positions are made use of in the same way.

Here we find a contrast with English. Consider the following:

(80)  
\begin{align*}
\text{a.} & \quad \text{Which man killed which Kennedy?} \quad \text{\textit{(PL)}} \\
\text{b.} & \quad \# \quad \text{Which Kennedy did which man kill?} \quad \text{\textit{(*PL)}}
\end{align*}

Following Barss’ (1990) observation that D-linked Wh-questions in English which front WH2 over WH1 can only have an SP reading, (80b) is predicted to be ill-
formed. Incidentally, (80) is a good example to show that the PL reading is not available in these contexts (cf. fn. 13 above).

Not surprisingly, the analogues of (80) are both well-formed in German:

\[ (81) \]

a. Welcher Mann hat welchen Kennedy ermordet? \((PL)\)
   which man has which Kennedy killed
   ‘Which man killed which Kennedy?’

b. Welchen Kennedy hat welcher Mann ermordet? \((*PL)\)
   which Kennedy has which man killed
   #‘Which Kennedy did which man kill?’ \((German)\)

However, the felicitous replies in each case are as follows, where (82a) corresponds to (81a) and (82b) to (81b):

\[ (82) \]

   ‘Lee Harvey Oswald killed John and Sirhan Bishara Sirhan Robert.’

   \(lit.\) ‘John, Lee Harvey Oswald killed and Robert, Sirhan Bishara Sirhan.’

In sum, while German seems to pattern with Bulgarian in multiple Wh-question formation in that both move all WHs into the \(\omega\)-domain, it also resembles Italian in not forming an information multiple question, but necessarily D-links all WHs. Unlike English, however, D-linked WHs must be licensed in the \(\omega\)-domain, a fact we might want to explain by the prolific nature of the left periphery in German, in that it allows much more discourse-related movement into the articulated C-layer.

5.5.3 Speculations: Discussion, Consequences and Predictions

The analysis of multiple Wh-fronting languages such as Serbocroatian, Polish and Russian as essentially Wh-in situ languages (Bošković 1998b) allows for a possible ac-
count for why these languages do not show Syntactic Superiority effects—unlike other multiple Wh-fronting languages, such as Bulgarian and Romanian. Syntactic Superiority, the ban on moving a non-D-linked WH over a higher one, arises basically when two elements bear a Wh-feature. If the WHs in zero Wh-movement languages do not move to check [Wh], but some other property, such as [focus], the Superiority Condition arguably does not apply. Discourse-driven movement (focalization, topicalization, scrambling) has different properties from Wh-movement. However, as the moved WH in English does check [Wh], and as only one WH moves, this ban applies to English Wh-movement.46

If German were like English, we would expect it to be sensitive to Syntactic Superiority as well. It is not. If German were like Bulgarian, we would expect the same. As we have seen that German patterns like Bulgarian with respect to interpretation, and not to Serbocroatian, we might face a problem. However, as the last section suggests, the WHs move into the ω-domain not to check [Wh], but something else. If all Wh-phrases in German are obligatorily D-linked, they are “topic-like.” As such, the feature driving Wh-movement in German is not the same as in Bulgarian, but the Prolific Domain targeted is.

As a final stretch of comparison, we have seen arguments that German is like Bulgarian in moving all WHs into the ω-domain. German is also like Italian in not forming multiple information questions. It now looks like German is also like Serbocroatian or Japanese in that Wh-movement is driven by a discourse feature, rather than [Wh].

The details need to be worked out, but the course of investigation is clear. If Wh-phrases in German do not move to check a Wh-feature, there is no reason why they should be sensitive to the Superiority Condition. The lack of Syntactic Superiority effects in mono-clausal multiple questions is then accounted for in the same way that Japanese or

46 We do not need to specify the technical implementation of Superiority in locality terms any further. Naturally, many questions arise, one of them being why Superiority is only sensitive to Wh-features and not others. I refer the reader to Chomsky (1973)…
Serbocroatian allow for moving any WH over any other. These languages differ, though, in the landing site: Japanese and Serbocroatian WHs stay within the $\phi$-domain, while German ones are very left-peripheral and target the $\omega$-domain. This difference could be tied to scrambling. A possible direction to apply Prolific Domains and anti-locality to scrambling could run along these lines, especially if we take pre-subject scrambled constituents in German always target a topic-like projection inside the $\omega$-domain (Grohmann 1996a and subsequent work), whereas Japanese or Serbocroatian employ the $\phi$-domain.

This approach raises an additional question. If German WHs are topic-like (as D-linked Wh-phrases in general are, employing the simple diagnostic of “known” vs. “new” information in this intuitive sense), which position do they actually target? Is it FocP and FP, as in Bulgarian, or does it involve the recursive TopP, or even some other position?

In previous work I suggested a derivation that moves both WHs in German to a topic projection, before further moving the higher WH into FocP (Grohmann 1998). However, this followed the assumption that all WHs move to check [Wh] at one point, and it held fast to the belief that even multiple Wh-questions in German are information questions. This lack of understanding of DRQ has now been improved: biting the bullet and taking all WHs to be D-linked in the formal sense removes the additional movement step to FocP. As such, it is perfectly feasible that all WHs actually target TopP. This would not change anything said so far. The strongest argument for multiple $\omega$-movement might come from the “Beck-effect.” But whether all WHs target TopP or some other projections, the predictions laid out in section 5.4.2 can still be maintained. Moreover, FocP, now not occupied by a WH, might serve the purpose of encoding the contrastive element that can appear in between two fronted Wh-phrases (cf. section 5.4.3).

If we take the $\omega$-domain to be really more as envisioned originally by Rizzi (1997), FocP is couched between two topic phrases, each recursive. Assume the two WHs to target the higher and the lower TopP, respectively. In this case, there is room for one contrastively stressed element in between—under the assumption that fronted con-
contrastive elements are unique (Horvath 1986, Brody 1990, Rizzi 1997 and others). This would predict that only one contrastively marked constituent can appear in between the two WHs. This prediction is borne out also:

(83) * Wer hat MOBY DICK den STUDENTEN wann zu lesen aufgegeben?
     who has Moby Dick to-the students when to read given
     ‘Who gave when the STUDENTS MOBY DICK to read?’  (German; cf. (65))

There is nothing inherently wrong with the information expressed here. We could imagine a context in which several books were assigned to several groups of people to read. In fact, all that seems to be wrong with (83) is that two contrastive elements are fronted. Compare with some grammatical alternatives (and additional information):

(84) a.  Wer hat MOBY DICK wann den STUDENTEN (und WAR AND PEACE den LEHRERN) zu lesen aufgegeben?
     ‘Who gave when MOBY DICK to the STUDENTS to read (and WAR AND PEACE to the TEACHERS)?’

b.  Wer hat den STUDENTEN wann MOBY DICK (und den LEHRERN WAR AND PEACE) zu lesen aufgegeben?
     ‘Who gave when the STUDENTS to read MOBY DICK (and the TEACHERS WAR AND PEACE)?’  (German)

I take (85) to be an appropriate answer in both English and German:

(85) Peter gave the students to read Moby Dick yesterday (and John gave the teachers to read War and Peace the day before).

Another difference between German and Bulgarian is that topicalized material can intervene between the two WHs in the former, but not the latter. Bulgarian does, however, allow topics to precede both WHs (Rudin 1985). Naturally, many variables are involved, such as a satisfactory understanding of V2 in German, the (alleged) adjacency
between Bulgarian WHs etc. But a first pass involves the finer structure of the ω-domain. If WH1 targets FocP in Bulgarian, there is a topic position available above it; but if the higher WH in German targets TopP, there is no such position left.

(86) a. Ivan vāera kakvo kupi?
   b. Vāera Ivan kakvo kupi?
yesterday Ivan what bought
   ‘What did Ivan buy yesterday?’ (Bulgarian; Richards 1997: 111)

(87) a. * Koj vāera kakvo kupi?
   b. Vāera koj kakvo kupi?
yesterday who what bought
   ‘Who bought what yesterday?’ (Bulgarian)

The same effects can be found in Romanian (Ileana Comorovski, p.c.). This shall conclude our discussion of some properties of multiple Wh-questions across languages, with special attention paid to German. Naturally, questions regarding the exact positions suggested here remain open, and so do others. The goal of this chapter was, however, to take the Hagstrom-Bošković approach seriously, build on their analysis and see where it leads us when we consider German—the language that seemed rather odd in the initial tripartition of strategies for Wh-question formation under the assumption that only one Wh-phrase movers overtly and the other stays “in situ.” Our discussion of in situ positions, the relevance to Prolific Domains and Exclusivity and others suggest an alternative approach without changing the crucial ingredients of the Hagstrom-Bošković approach.

5.6 Conclusion
In this chapter, I argued primarily for overt, syntactic movement of both Wh-elements in a binary Wh-question in German. I suggested that one WH targets FocP and the other a lower projection FP. In this sense, German patterns more like Bulgarian than English.

Following recent attempt to tie the interpretation of single-pair vs. pair-list in multiple questions to the syntax, such an analysis accounts for why German behaves like other multiple Wh-movement languages in the relevant aspects. We have seen arguments that Serbocroatian, another language that fronts all WHs, actually does not move Wh-phrases to the same positions as Bulgarian does. Bulgarian involves ω-movement of all WHs, while the Serbocroatian WHs remain below, somewhere inside the φ-domain. Further comparison allowed us to classify Serbocroatian as a Wh-in situ, or better, zero Wh-movement languages, on a par with Japanese. Other factors are involved, relating to scrambling, but following the Hagstrom-Bošković approach, we were able to classify the three strategies of forming multiple Wh-questions across languages in a clean typological tripartition: some languages move all WHs into the ω-domain (multiple Wh-movement), other only one (singular Wh-movement) and a third class non (zero Wh-movement).

A multiple Wh-movement analysis of German questions makes, of course, certain predictions, especially for the syntax. By adopting an articulated structure for the ω-domain from recent research, we were able to investigate one prediction: if the position between the two fronted WHs is a topic projection, we expect only possible topics to appear in between the two Wh-phrases. This prediction turned out to be borne out. By identifying (non)topicalizable material, independent of Wh-questions, we were able to account for the so-called “Beck-effect” of intervention in a straightforward manner.

Once we established that German is like Bulgarian in ω-moving all WHs, we looked at other properties of German multiple Wh-questions and found a striking restriction, namely that all WHs, regardless of their internal structure, must be D-linked. In this
sense, all Wh-phrases represent old information, in the sense that the set of individuals expressed by all WHs must be known to speaker and hearer. This allows for an additional comparison: I suggested that German is like Italian in not allowing multiple constituent questions. This we tentatively linked to the lack of more than one position for information focus, the type of focus expected in a felicitous, non-contrastive answer to a constituent question.
CHAPTER 6

PICKING UP THE PIECES:
A RADICAL VIEW FROM THE LEFT

This final chapter—partly as a conclusion of this work, partly as a road map for further research—summarizes the main properties of the framework of Prolific Domains, clausal tripartition and Exclusivity (the theory of anti-locality) which I have proposed and presented in the preceding chapters. In so doing, I elaborate on two concepts introduced earlier, address and context, and lay out a possible route for the interaction between syntactic computation and interaction with the interfaces. In this context, similarities with and differences to recent approaches will be discussed.

6.1 Prolific Domains: The Framework

In this dissertation I have proposed a theory of anti-locality effects, the fact that while subject to locality conditions, movement must not proceed too locally. The clue to define anti-locality lies in an appropriate domain of evaluation relevant for closeness. Minimalist inquiries into syntactic computation offer a natural cut: the part of the clause where theta relations are created, the part that licenses agreement properties and the part responsible for those types of discourse information that are syntactically relevant. This
tripartition is natural in the sense that it is defined over parts of a phrase marker which are characterized by checking configurations of formal features pertaining to each of these three “super-features,” and there seems to be evidence, all other things being equal, that no element checks more than one feature of these three types. We called each of these three areas a Prolific Domain, an area within the phrase marker for a particular clausal computation ( “domain”) which itself consists of a number of projections (“prolific”). We labeled the first Prolific Domain $\theta$-domain, the second $\phi$-domain and the third $\omega$-domain. This is repeated here from (26) in chapter 2:

(1) *The concept of Prolific Domains (\(\Pi\Delta\))*
   i. $\theta$-domain: the part of the derivation where theta relations are created
   ii. $\phi$-domain: the part of the derivation where agreement properties are licensed
   iii. $\omega$-domain: the part of the derivation where discourse information is established

We considered to think of the three “super-features” as context values, a notion to be refined momentarily. This helped us formulate the theory of anti-locality in terms of Exclusivity, with the definition of (30) from chapter 2 repeated here for convenience:

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

Before addressing the notions “address” and “context” further, occurrences of $O$ within a Prolific Domain was evaluated over non-distinct copies. As a rough guide, we can thus take Exclusivity to be the ban on XP-movement within a Prolific Domain. The notion of Prolific Domain, in turn, was understood in terms of the following working hypothesis, which will be fleshed out further in this chapter ((27) from chapter 2):
(3)  *Prolific Domain* (working hypothesis)

A Prolific Domain ΠΔ is a contextually defined part of the computational system

i. which provides the interfaces with the information relevant to the context and

ii. which consists of internal structure, interacting with derivational operations.

These are the cornerstones of the framework sketched in the course of this work. They allowed us to ban movement of a given element from one position (launching site) to another position (target site), where the target site would be too close to the launching site. All in all, this makes the right predictions: lexical items do not move from one θ-position to another within the same clause, they do not move several times within the same agreement layer to check several sets of φ- or Case-features, they do not move within an articulated COMP to check off a set of topic-, focus- and/or Wh-features, and so on. Movement is understood in the usual way (Chomsky 1993), in terms of Copy and Merge, where the lower copy is subsequently deleted on principled grounds (Nunes 1995, 1999). Nunes in fact shows that deletion of non-distinct copies is forced by the phonological component: in order to linearize the objects of a phrase marker at PF, where each object must be asymmetrically c-commanded by the one preceding it. If an object is part of a dependency consisting of two or more non-distinct copies, the lower copies must be deleted, or else the non-distinct copies would at the same time precede and follow each other. Assume this much to be correct, roughly as outlined here.

Non-distinctness of copies turned out to play a major role in our framework also. We found that while it holds in general, anti-locality seems to be violated in some cases—namely those that involve a dependency between a phrase and a pronoun of sorts in which both were arguably non-distinct copies of one object in the phrase marker at some point in the derivation (identity, connectivity etc.). The approach suggested here is that in these instances, the pronominal element is a spelled out copy of the phrase (“Copy Spell Out”), and that it seems to bear all characteristics of a deleted copy, except for its PF-matrix: it is c-commanded by the original copy, it is the lower of two copies, and so on.
If lower non-distinct copies need to be deleted for PF reasons in the case of Move, we might want to consider similar motivation to spell them out in other instances. The definitions of Prolific Domains and the Condition on Domain Exclusivity suggest exactly that. We have assumed this to be the case throughout, and it did some work for us. Let us now look at some relevant details a little bit closer. The next section is thus an extension of the core proposal from chapter 2. It is not the only possible conceptualization of the idea, and other ways are not excluded, but it is one which seems appropriate in the context of recent work on the syntactic computation and its connection to the interfaces. To mind comes in particular recent work by Chomsky (1998, 1999), Epstein et al. (1998), Boeckx (1999a), Uriagereka (1999b), or Platzack (in press), for example.

6.2 Syntax at the Interfaces: A Sketch

The tripartition envisioned here (and elsewhere; cf. Rizzi 1997, Platzack, in press and many others) is intuitive, and given the work it does, we can accept its existence with little further ado. What needs to be addressed are the details. Many questions have to be left open, but I will sketch a possible approach concentrating on the ingredients of the definitions repeated above and putting them into some perspective.

One question that will not be addressed is why these three and only these three “super-features” should exist and play such a crucial role for the computation.\(^1\) For what-

\(^1\) Another questions regards the number of the Prolific Domains. I do not exclude other possible Prolific Domains, possibly in form of small clauses, or on a different level than the clause (such as the nominal level.). It is likely that DP constitutes another Prolific Domain, as has often been observed that the nominal layer and the full clausal layer share a number of properties.

Returning to the clausal level, Rizzi (2000) also considers “super-features,” though from the other end of locality, the upper bound on dependencies in terms of Relativized Minimality (basically updating the framework of Rizzi 1990).
ever reason, they do. But we can ask ourselves how they come into play. As our working hypothesis says, a Prolific Domain is contextually defined. Let us take each of these contextual “super-features” to be a context value. We thus deal with three relevant context values, represented as $|\theta|$, $|\phi|$ and $|\omega|$. Anything confined in the domain marked by one of these context values is part of that Prolific Domain. The possibly simplest way to integrate such values into the system—without any additional devices—is by treating them as part of the lexical make-up of the relevant lexical or functional head, rather than indexicalizing relevant material with this context information, for example. One view of the initial make-up of these heads is that they contain at least the formal feature they correspond to, which can be an abstract morpheme or a phonetically realized element. The particular details depend on the theoretical assumptions regarding the initial structure, the exact content of the lexical array LA, the role of morphology, and possibly more. Under this view, whether a particular head belongs to one or the other Prolific Domain is universally determined. Thus, V and v are always part of the $\theta$-domain because both contain $|\theta|$, Agr or T always belong to the $\phi$-domain and FocP or CP to the $\omega$-domain. This is desirable insofar as the observations on anti-locality laid out in chapter 2 arguably hold across languages (that is, no element in a simple, monoclausal construction checks two $\theta$-features, several Cases, and so on).

We now have accounted for the base of the definition of a Prolific Domain (3) and the specification (3ii), given more or less standard assumptions about the lexicon LEX, the computational system C_{HL} and the interaction between the two. But we need to say something about (3i), namely the notion of information to the interfaces—at least in order to specify the concept Exclusivity further. I will thus discuss the notion of address (identification) and how it fits in, before comparing this to other, similar approaches.

---

2 See for example the work done in Distributed Morphology on some of these problems (Halle & Marantz 1993 and much subsequent work).
A simple of view of the interaction between $C_{HL}$ and the interfaces (qua the intermediaries LF and PF) maps all information accumulated in the course of the derivation and maps it onto the interface levels LF and PF at a particular point in the derivation. This point is often referred to as Spell Out, the point of moving from overt to covert syntax. In the original minimalist model of Chomsky (1995b), Spell Out is a unique operation that targets the derivation at one point—roughly the same that was previously held to constitute the boundary of S-structure, but without forcing any conditions on the derivation up to that point. In order to keep the complexity of $C_{HL}$ as simple as possible, it applies only once, apparently because this is the most economical way. However, if economy is really taken to mean “up to convergence,” Spell Out should be applied more than once if it can be shown that some derivation would not converge had Spell Out not applied more than once (Uriagereka 1998: 322). Uriagereka (1999b) does this and suggests that Spell Out should be treated like any other operation of $C_{HL}$, apply as often as needed to converge. In this context, the question arises how often Spell Out should apply. The answer Uriagereka gives is once per command unit, basically, every time a left branch is formed. Chomsky (1998, 1999) picks up on the idea of Multiple Spell Out and partitions the clause into phases, parts of the derivation that correspond to propositional objects which converge, where convergence can be taken to denote an application of Spell Out, a point in the derivation at which syntactic information is sent to the interfaces.

We will come back to both approaches soon, but for now take the idea seriously that the operation Spell Out can be applied more than once and that the interfaces are thus not computed over the entire structure—in other words, at least the consequence that LF and PF are fed cyclically can be upheld, regardless of the details of “Spell Out.” Given the (technical) discrepancy in this respect between Chomsky’s and Uriagereka’s models, it is not implausible that another way of implementing the idea of Multiple Spell Out can be envisioned. Leaving a deeper discussion aside for the purpose of presentation, let us consider its relevance to Prolific Domains and the framework developed here.
I suggest that (3i) be understood as follows. Once all features inside a collection of projections of the same type are checked, the derivation up to this point is evaluated. This collection refers, of course, to a Prolific Domain, as just sketched and evaluating thus concerns each Prolific Domain in turn (or cyclically). Evaluation consists of sending the information to LF and PF. Let us consider this in some detail.

If Spell Out is indeed the operation that ships the relevant information to PF and LF, and if it can apply more than once, PF and LF should not be considered to be one-dimensional levels but, what Uriagereka (1999b) or Uriagereka & Martin (1999) call, “components”—collection of information that is computed cyclically, one set at PF and another at LF. Such a cyclic application of Spell Out then takes each Prolific Domain as the relevant sub-part of the derivation that needs to be spelled out, or as I want to put it, evaluated. Evaluating a Prolific Domain consists of sending each object in the relevant part of the phrase marker to the PF- and LF-components. As many of these objects occur in more than one Prolific Domain in the course of a given derivation, it is necessary to mark them in a uniform way throughout. Call this the “address” of an object in a phrase marker. Each address of, or Spell Out point for, a given object, marks the LF- and PF-features of that object at that point. Arguably, the notion of address is important for LF as it identifies all features of one element relevant for interpretation, so that at the end of the derivation the entire set of checked features—for example, {[agent], [3rd person singular masculine], [accusative], [Wh]} for who in Who did John kiss?—can be matched to all occurrences of who. On the PF side this is presumably easier, as each occurrence looks alike, in this case [ ] spelling out the θ-domain, [ ] spelling out the φ-domain, and [ ] spelling out the ω-domain.

---

3 This term has been used by other scholars in slightly different ways, yet always as marking a special occurrence. Manzini (1992) and Uriagereka (1997), for example, tie this special occurrence “address” to Case-marking, relevant for a theory locality in one case and an optimal, economy-driven checking approach in the other, respectively.
This is a simplified view, of course, but one that will suffice for presentation. We can think of an occurrence of an object as a specially coloured element. All non-distinct copies are of the same and each different lexical item of a different colour. A simple derivation for John kissed Mary contains roughly the following LA:

(4) \[ \text{LA} = \{\text{John, kissed, Mary, v, Agr, T}\} \]

Let us concentrate on the maximal projections here, the DP John and the DP Mary. At the end of the derivation, both will have checked their \(\theta\)- and \(\phi\)-features, and both have different Cases. Assume that each Prolific Domain is a Spell Out or evaluation point, so that after overt syntax, we get the simplified representation in (5) or shown in tree representation in (6) below.

(5) \[ \text{TP John T [AgrP Mary Agr [vP John kissed-v [VP kissed-V Mary]]]} \]

Leaving further details aside, the two relevant maximal projections, DP Mary and DP John, are spelled out first once the \(\theta\)-domain is created, namely when vP is complete. The PF-component receives \([\ ]\) and \([\ ]\). Under a Single Spell Out model, LF would be created at the end and would only have to count all occurrences of the same item and take the relevant features checked at each occurrence. Under a Multiple Spell Out view, we have to say something else. Presumably, all that is shipped to LF at the point of \(\theta\)-Spell Out is something like [theme] for Mary and [agent] for John, the two \(\theta\)-roles. Both DPs still have at least their \(\phi\)-features to check, and there is no need to inform

---

4 Recall that under the minimalist view, insertion of lexical items from the LA to the derivation includes the entire PF-matrix of the element as it is eventually pronounced.

5 To illustrate the point, assume all DPs move overtly, whether this is English or German, under our set of assumptions, for example.
LF that they have not yet been checked. At the next point of Spell Out, when the $\phi$-domain is created, the two DPs are again shipped to the PF-component in their output forms $[\ldots]$ and $[\ldots]$. PF now has two of these and must decide which one to delete and which one to keep. LF receives the information [3rd person singular feminine, accusative] and [3rd person singular masculine, nominative] for Mary and John, respectively. However, it does not receive anything else; LF now has to match these features with the features already accumulated, such as [theme] and [agant].

This might be a crude sketch of the interaction between syntactic computation and the interface components, and it begs for more details. However, it suffices to illustrate the less ambitious goal of the current work to lay out a theory of anti-locality. We now have a way of understanding Copy Spell Out in terms of a PF-requirement.
By calling an occurrence of an object in the phrase marker “address” and encode the information shipped to LF as “address identification” (or Address Identification, capitalized by convention), LF is able to match [3rd person singular feminine, accusative] with [theme] and [3rd person singular masculine, nominative] with goal, regardless of the PF-structure of these two objects. The matching is done by Address Identification, which we could think of as different colours. The first of these two matched sets would be red, the second blue, for example. In this sense, Address Identification is nothing but a book-keeping device for the LF-component.  

Exclusivity is defined in terms of a unique Address Identification, which is a uniquely coded (set of) feature(s) or an occurrence at LF. If coloured sets of LF-features indicate identity of sorts (in the form of matching, as described above), it should not matter whether there are two different-coloured features of a particular context (such as [agent] and [theme]) or two identically coloured ones. In the latter case, the two would be matched in relation to each other, and, depending on further content of these features, John likes himself or John and Mary like each other are potential results. For PF, on the other hand, the difference between John and Mary vs. himself is quite important: in one case, there are two different items, each of which needs to be pronounced once, while in the other—under the derivational approach sketched in chapter 3—is a copy that should not be deleted as usual, but pronounced. The repair strategy is to spell out one of the two copies bearing these features and recolouring it for PF purposes. Thus, if the LA in (4) does not contain Mary, the “red” DP John would move from one θ- to another θ-position, and there would be two red occurrences at Spell Out. But by previously spelling out the lower occurrence, marking the PF-matrix blue, the two occurrences are made distinct, PF

---

6 I abstract away from further complications, in particular the distinction of interpretable and uninterpretable features, an innovation which might be dispensed with (which would then be one of the more ambitious goals of this framework). This is just a sketch, with the aim to clarify the idea behind Exclusivity. I leave further details open for future research.
does not have to delete any copy, and the LF-component can happily match the relevant features accumulated in subsequent operations (under identity of sorts, i.e. it only deals with red features throughout the derivation for both instances of *John*).

Apart from the issue raised in fn. 6, this view implies two things. First, Copy Spell Out must take place before the relevant Prolific Domain is established, otherwise LF would receive conflicting information (by (2i)). Second, a Prolific Domain is established or evaluated (Spell Out) as soon as it is complete, thus allowing an element to move beyond. This will be addressed more in the next section.

If this is on the right track, we might want to modify the model of the grammar shown in (8) in chapter 1 along the following, more colourful—if only tentative—lines (from Grohmann 2000a: 42), regardless of whether we take Spell Out to be the relevant notion of something less committal (“evaluation”):

(7)
To conclude our review of the Domain-driven framework presented here, we now have a way of conceptualizing the assumptions on the computational system underlying the definitions of Prolific Domains and Exclusivity. By evaluating locally—that is, at the point of establishing—each Prolific Domain, an occurrence of an object in a phrase marker must be unique, or exclusive, at the point of evaluation, which is within a given Prolific Domain. Spelling out the lower copy can repair a potential violation, but must take place before evaluation.

Thinking of occurrences in colour terms is one way of looking at things. Another analogy would be with the realm of quantum physics. Exclusivity as viewed here is in a way reminiscent of Pauli’s Exclusion Principle, which says that “no two electrons in an atom can have the same set of quantum numbers” (Serway, Moses & Moyer 1989: 230, emphasis omitted—KKG). What the quantum numbers do is identify the position of an electron within an atom. What feature checking does is identify the position of an object in a phrase marker, in particular within a given Prolific Domain. There are many more examples from other areas one could draw from, but the bottom line is that Exclusivity of an element within a specially contained area seems to be a natural state of affairs.

6.3 Exclusivity, (Anti-)Locality and Cyclicity

Aside from many other consequences, this sketch allows us to compare the present system with other, similar ideas on a dynamic conception of syntactic computation, such as the theory of Multiple Spell Out (Uriagereka 1999b) and the (re)introduction of phases (Chomsky 1998, 1999). The main advantage of the first is that it derives the LCA immediately, on the assumption that all command units (basically, left branches) are spelled out, thus linearizing all objects in the phrase marker. Uriagereka treats the opera-
tion Spell Out to basically “morphologize” every object, turning it into one word at the interfaces, and as a consequence prohibiting extraction out of a spelled out element. This might be the right way to understand Spell Out, but it is different from the purpose ascribed to it here.\(^7\) As such, Multiple Spell Out and Prolific Domains are not necessarily mutually exclusive. Perhaps a better way of characterizing the above sketch is not in terms of spelling out, but in terms of evaluating. The PF and LF components thus evaluate every Prolific Domain locally, at the point where it is created, leaving open further specification of the operation Spell Out, perhaps in Uriagereka’s terms.

The idea of a phase, however, is not as simple to integrate. In the following, I pick out the core properties Chomsky (1998, 1999) ascribes to a phase and the correlates found with Prolific Domains.\(^8\) The upshot will be that the framework of Prolific Domain fares no worse than a phase-based system in conception or empirical coverage.

For the purpose of a brief summary, Chomsky suggests that \(vP\) and \(CP\) induce (strong) phases, essentially bounding nodes for subsequent movement. An element can only move out of a phase into a higher one from an edge position. It is obvious, then, that a phase cannot correspond to a Prolific Domain. First, we treat TP as a “phase” as well, and second, the structure that allows an element to get to the edge of a phase and then move further is banned from the system endorsed here. The edge of a phase is basically the relevant head (\(v, C\)) and all of its specifiers. In the current framework specifiers are

---

\(^7\) It also faces an empirical challenge in that extraction out of left branches is not prohibited in all languages (such as \(\text{was für-split}\) in German, for example, which is insensitive whether it applies to a subject or an object), but this implementation presumably depends on other factors, such as the ultimate structural analysis of the “base structures” of these strings. See also Stepanov (2000) for discussion and references.

\(^8\) Note that Chomsky (1998) introduces two phases, basically \(vP\) and \(CP\). In Chomsky (1999), he distinguishes weak from strong phases. As far as I can see, the relevant notion for our present comparison is a strong phase (\(vP, CP\)), while weak phases shall not concern us here. (One weak phase is a \(VP\), or deficient \(vP\), which is relevant for the allegedly phonological operation \(\text{THematization/EXtraction}\), for example, in which case \(\text{TH/EX}\) alone may “spell out” rather than the entire phase.)
unique, hence no element can use an outer specifier as an escape hatch. The assumptions on feature checking are also slightly different, as a consequence of multiple specifiers on the one hand (where the object checks its φ- and Case-features against v) and the option to check features without movement on the other (via Agree). Let us view these and other properties of phases in a little bit more detail.

Chomsky (1998: 21) defines a phase as “propositional” and “convergent.” The latter relates to the cyclic operation of Spell Out/evaluation and could translate as such into the Domain-driven framework: Prolific Domains and phases are convergent, they both induce Spell Out. Regarding the former property, it is not quite clear what the details of a “propositional” property of phases are (see e.g. Uriagereka & Martin 1999, Pietroski 2000, Uriagereka 2000, or Juan Uriagereka’s University of Maryland fall 2000 lectures). I suppose vP could relate to the (minimal) event structure of the clause, while CP would be propositional in that it encodes all other specifications and basically completes the clause. I would like to contrast this vague understanding of “propositional” (coupled with the question mark why TP does not induce a phase, at least a non-deficient TP) with our property of a Prolific Domain, namely that it is essentially “contextual.” I have offered suggestions in how far a Prolific Domain is “contextual” in chapter 2, and we have seen further evidence throughout. In the two core properties of definition, Prolific Domains and phases seem to fare equally well or badly. Let us now look at some technical aspects.

A phase-driven framework is subject to the Phase Impenetrability Condition (or PIC; Chomsky 1998: 22, 1999: 10), which basically states that elements within a phase are not accessible to further computation at the next higher phase, unless they are located in the edge of that phase. The Domain-driven framework takes the opposite stance, namely that an element cannot move within the same Prolific Domain (Exclusivity qua the definition of the CDE in (2)). That PIC and CDE can indeed be compared in this context follows from the following divergent assumptions on the syntactic derivation between the two frameworks.
First, we take displacement to be triggered by a property ("feature") on the displaced element itself, that is we extend the original conception of Move (Chomsky 1993). The phase-driven approach assumes the trigger to lie in the target ("Attract"). Moreover, much of the motivation that led to movement in earlier minimalism is now dispensed with and uninterpretable features are eliminated in a long-distance configuration by the operation “Agree.” This is as far as we can compare the basic background assumptions. A more detailed evaluation of Move vs. Attract/Agree is beyond the scope of this work and depends on many other assumptions (see, for example, fn. 6).

Second, we eliminated the possibility of multiple specifiers, or anything more than one maximal phrase per projection created by movement. As such, an edge can never be created in Chomsky’s sense, as an escape hatch. These two differences account for the opposite framework that the PIC and the CDE yield, on the one hand, and for the similar spirit the two imply, on the other. The comparison of differences and similarities thus extends further. Recall that we take the position that a Prolific Domain is evaluated as soon as it is created, while the relevant equivalent of Spell Out in a phase-driven framework takes place at a later stage: a lower phase PH1 is only then spelled out and made inaccessible for further computation once the next higher phase PH2 is reached. This allows elements in PH1 to move to the edge of H1 (the head of that phase) and be attracted by H2 (the head of the higher phase). Thus, late Spell Out of phases allows for an escape hatch for material that needs to be displaced, while early Spell Out of Prolific Domains freezes all Domain-internal movement.

This is a rough comparison of the theoretical underpinnings of the two theories. Let us then look at empirical coverage. In order to move out of a phase, an element must first move to the edge of that phase. This is not feasible in a Domain-driven approach and the most natural counterpart of movement to the edge is movement to the next Prolific Domain. For standard cases, this seems to make the right cut. Thus, an outer SpecvP checks those features that SpecAgrP checks here, namely $\phi$- (and Case) features.
We could reinterpret this by suggesting that the vP-phase and the θ-domain are the same, if the former is cut off after its first specifier, the thematic subject position. Regardless of numbers of specifiers, anything beyond would then fall into a different phase or Domain. But this easy transposition cannot be upheld because TP (or the φ-domain) does not constitute a phase in the relevant sense. In all simple, monoclusal structures this difference does not amount to much. Take Wh-movement, for example. In both frameworks a direct object Wh-phrase is generated in its θ-position and ends up in the surface Wh-position (specifier of FocP or CP). In one framework, it moves through AgrP, the φ-position, for by now obvious reasons, then targets the final position inside the ω-domain, that is, it moves from the θ- to the φ- to the ω-domain. In the other, it would be frozen in its base position once we reach “C” (the higher phase), unless it moves to the edge of the lower vP-phase, from where it is accessible to subsequent movement to SpecCP. I take it that both approaches embody the same necessary assumption: that all movements must target the next highest Prolific Domain for principled reasons (to compute a complete set of context features) or that all movements must use the escape hatch at the edge (which Chomsky calls “independent feature-driven movement” or IFM; Chomsky 1998: 21f.). I do not want to question the “naturalness” of either alternative, but consider some cases which can be compared, which arguably suggest that only vP and CP are phases, and which empirically support the phase-driven approach within the relevant set of assumptions. These cases involve clause boundaries, or multyclausal structures.

In Chomsky’s framework, movement from one thematic position to another is ruled out, and control structures contain a controlled empty subject in the embedded infinitival clause, PRO. In the current framework, we do not want to exclude θ-to-θ-movement a priori and dispense with PRO as an object in C_HL. Adapting Hornstein’s (1998, 1999, 2000) analysis of control constructions slightly, along the lines suggested by Castillo, Drury & Grohmann (1999), we take the controlled subject to be a copy of the phonetically realized subject. Compare the two derivations:
The controlled embedded infinitival in (8a) needs to be contained in its own phase, otherwise anything beyond the vP-phase would be accessible to the matrix vP-phase.\(^9\) This can be done by postulating CP in between the two clauses (but see fn. 9). (8b) does not need to stipulate the presence of CP, as it does not need to induce a phase. Moreover, if the controlled subject is indeed the copy of the matrix subject, we do not need to treat the embedded grammatical subject position as present. If non-finite T cannot license a phonetically realized element in these contexts, there is no reason to assume it in the first place—unless one would want to express the Extended Projection Principle (EPP) somehow. But this has come under attack in recent work.\(^{10}\) This allows for a generalization on locality conditions in a Domain-driven framework:

(9) \textit{Domain-driven Locality}

a. Intra-clausal movement targets the next higher Prolific Domain.

b. Inter-clausal movement targets the same type of Prolific Domain.

We have seen instances of (9a) throughout, and the ban on the opposite, with the caveat of spelling out lower copies (CDE). The derivation in (8b) suggests something like (9b), and while a complete demonstration lies outside the scope of the current paper, some core cases can be illustrated, suggesting that (9) is indeed on the right track.

\(^9\) The structure portrayed in (8a) is one potential derivational history for this sentence, where it does not matter for current purposes whether PRO is inserted or derived. What is relevant is the fact that CP is forced, for phase-reasons, despite evidence that control clauses involve a TP, as noted by Uriagereka & Martin (1999); see Ormazabal (1995), Bočković (1997c).

Take the “counter-case” of control structures. The generalization in (9) makes one prediction, if correct: raising structures should contain intermediate SpecTP, on the assumption that a raising subject does not move through the matrix θ-position but targets the matrix SpecTP straight away (under the traditional approach) or that it needs to target the same type of Prolific Domain if it crosses a clausal boundary (current approach). The standard analysis of raising constructions is the following (Chomsky 1981, 1993):

(10) \[[TP \text{John} T [VP \text{seems-V} [TP \text{John} to-T [vP \text{John} sleep-V-v]]]]\]

The matrix subject is the grammatical subject of both clauses (traditionally, for reasons of the EPP) and the thematic subject of the embedded predicate only. Let us assume that the matrix clause does not have a vP (see section 3.2.3). The relevant movement step of the matrix subject is the last one, from embedded SpecTP to matrix SpecTP. If this is indeed the right movement, it conforms to (9b): the move cross a clause boundary and targets the same type of Prolific Domain in the target clause as it originates from in the launching clause, here the φ-domain.

However, as we said above that the EPP should not be invoked, if possible, the intermediate landing site might be a thorn in our side. This is essentially the position Castillo, Drury & Grohmann (1999), and others, advocate: the EPP does not exist, hence there is no motivation for an intermediate touch-down in infinitival TP, hence non-finite (deficient) SpecTP does not exist. Thus the anti-EPP proponents mentioned above and elsewhere deny intermediate SpecTP in these constructions also. But these two sides, pro or anti EPP, are not necessarily the only options available. Bočković (2000b) proposes a reconciliation of both viewpoints, namely that the EPP does not exist,11 but that certain

---

11 As Lasnik (to appear) notes, the EPP has undergone considerable conceptual—and as a result, technical—changes over the past two decades. While treated as a formal feature in Chomsky (1993, 1995a), latest conceptions of Chomsky (1998, 1999) are much more in (continued...)
intermediate subject positions do. One empirical case is the so-called wager-type verb class (Pesetsky 1991, Bošković 1997c). Verbs of this type select for a non-finite clausal complement whose subject position is filled by an expletive. To cut a long story short, if there can be inserted into intermediate subject positions in certain raising structures, that position should be available (but not in control structures). Bošković (2000b) compiles a wide range of data that suggest that both the derivations sketched in (8b) for control and in (10) for raising are correct.

This approach captures the best of a movement approach to control and an EPP-less theory, yet maintains a difference between the two other than in one additional θ-role. It also allows Uriagereka & Martin (1999) to keep (non-defective) TP as a phase, without, at the same time, getting into a conflict between the notions of Prolific Domain vs. phase, all other things being equal. Regarding the latter, the different analyses for control constructions allow phases and Prolific Domains to be empirically adequate in either framework, and we can take the decision between the two approaches to follow from the appropriateness of the background assumptions sketched here. Note that raising cases are handled equally well by either theory: the phase-driven framework does not want the matrix and embedded clauses be separated by a phase, while the Domain-driven approach captures the successive-cyclic nature of raising by (9b).

The following are abstract representations illustrating (9):

(11)  *Intra-clausal movement*

\[
[\ldots \text{XP} \ldots [\ldots \text{XP} \ldots [\ldots \text{XP} \ldots]]]
\]

(12)  *Inter-clausal movement*

\[
[\ldots \text{XP} \ldots [\ldots \text{XP} \ldots [\ldots \ldots [\ldots \text{XP} \ldots [\ldots \ldots [\ldots \ldots [\ldots \text{XP} \ldots [\ldots \text{XP} \ldots]]]]]]]]]
\]

line with the original assumption that sentences need to have subjects, or basically, that certain specifiers need to be filled (Chomsky 1981, 1982).
(11) would apply to a sentence like (13a), while (12) would correspond to (13b):

(13) a. Who did Mary see who?
b. Who seems to be likely to sleep?

In a phase-driven system, the intermediate copy of who would sit at the edge of the relevant phase from which it can move to its final destination in (13a). In the current approach it sits in the next highest Prolific Domain (φ-domain) and targets another position in the next highest Prolific Domain (ω-domain). The relevant intermediate positions in (13b) can be identified with SpecTP and are driven by IFM in a phase-based theory and by the generalization on inter-clausal movement here.

6.4 Concluding Remarks

In this chapter I have reviewed the core proposal of the theory of anti-locality presented in this work. We now have a full-fledged framework that not only captures anti-locality, as we set out to do, but also offers a new look on the interaction between syntactic computation and the interaction with the PF and LF interfaces. A comparison with recent approaches aiming for the same goal yields the following picture. A cyclic application of the operation Spell Out is an important ingredient, whatever the exact details, and it needs to be seen what the correct portion of the derivation is which gets spelled out. We also saw how Prolific Domains and phases fare with respect to conceptual and empirical motivation. If nothing else, the ideas presented here make predictions like any other good theory and are thus testable (a desired result from a theoretical point of view). Independently, the analyses of empirical phenomena considered in the past chapters constitute a contribution to the analytical repertoire of minimalist research.


Boeckx, Cedric. In progress. French in situ Interrogatives as (C)overt Clefts. Manuscript, University of Connecticut, Storrs.


Bošković, Iljko. 2000a. On Multiple Wh-Fronting. Talk given at the University of Maryland, College Park. [November 7]

Bošković, Iljko. 2000b. A-Movement and the EPP. Talk given at the University of Maryland, College Park. [November 10]


Epstein, Samuel David & T. Daniel Seely. 1999. SPECifying the GF “Subject;” Eliminating A-Chains and the EPP within a Derivational Model. Manuscript, University of Michigan, Ann Arbor & Eastern Michigan University, Ypsilante.


Grewendorf, Günther. 1998. Multiple Wh-Fronting in German. Manuscript, Johann Wolfgang Goethe-Universität/Frankfurt am Main.


Richards, Norvin. 2000. Movement Dependencies in a Root First Derivation. Talk given at the University of Maryland, College Park. [November 17]
Rizzi, Luigi. 2000. Locality and the Cartography of Syntactic Structures. Talk given at the University of Maryland, College Park. [November 29]


