An Order-Theoretic Approach to Syntactic Structure and Transformations

Frédéric Mailhot

A Thesis

in

The Special

Individualized Program

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Special Individualized Programs) at
Concordia University
Montreal, Quebec, Canada

September 12, 2005

Your file Votre référence
ISBN: 0-494-10340-X
Our file Notre référence
ISBN: 0-494-10340-X

ABSTRACT

An Order-Theoretic Approach to Syntactic Structure and Transformations

Frédéric Mailhot

Contemporary generative syntactic theory is recast in terms of the mathematical theory of partial orders and order-preserving mappings. Phrase markers are shown to be sets of syntactic objects partially ordered by set-theoretic containment, and by a formal relation analogous to traditional asymmetric c-command. Complete sentential structures are sets of phrase markers related by merger and movement transformations, both of which are order-preserving mappings.

The elaborated framework is used to provide a novel perspective on so-called phrase structure paradoxes. In conjunction with recent substantive insights concerning the nature of syntactic derivation, interpretive binding theory and VP-external merger of prepositions, a simple, unified account of the paradox is provided, avoiding problematic aspects of previous accounts.

Acknowledgements

There are so many people without whom this thesis would either not be what it is, or would simply not be. The road leading here has been unduly long and tortuous, and I don't have enough seconds left to acknowledge all those who have in some way helped me along the path. I shall instead content myself with a short list of thank yous to those whose involvement has been, either qualitatively or quantitatively, more substantial.

Premièrement (bien sûre!) mes parents, Michèle Paquin et Jean-Jacques Mailhot. Vous n'avez jamais arrêté de croire en moi et vous m'avez toujours laissé prendre mes propres décisions, même quand c'était clair qu'il y aurait des conséquences à subire. In so many ways, this thesis could not have existed without you. Merci du fond de mon coeur et avec toute mon amour.

Next, my old/new/best friend/lover/wife, Anna Mushynski. When the Rug of Certainty was so rudely yanked from under my feet 10 years (!!!) ago, you helpfully suggested an Intro to Linguistics class. You are perhaps most directly responsible for this being a Linguistics thesis, rather than (horrors!) something else, or nothing at all. You have kept me laughing throughout this degree, moved to a new city with me, showed an astonishing amount of faith in my ability to see this thing through to completion, and kept me believing on the few occasions that Doubt whispered ugly thoughts in neglected corners of my mind. Thank you so much, Anna. You make this fun.

I owe an enormous debt of gratitude —intellectual and material —to my supervisors, Charles Reiss, Mark Hale and Roberto de Almeida. Fortuity (and Ray Jackendoff) brought me to Charles and Mark in the Fall of 2002, while I was struggling through a Computer Science degree. Charles apparently saw something that made him feel that I was worth going to bat for, and invited me to be his research assistant, and then his student. He showed me that phonology could be fun (that paper **will** get finished), that I had just as much right as anyone to present at conferences, allowed me to lecture in his Cognitive Science class, and eventually let me write a thesis that had nothing to do with what we had worked on together. I met with Mark weekly for a Language Acquisition class that had little to do with acquisition but turned out to be an eye-opening introduction to philosophical issues in science and cognition that will shape the way I approach problems for a long time. Roberto, an incredibly patient man who is always ready to discuss anything related to linguistics or cognitive science, never got pushy about collecting on the work that I still owe him. It'll get done some day, I swear. Thanks to all of you for seeing me through these last two years.

Chris Wu, my erstwhile roommate, reminded me that having an age measured in geological time needn't be a bad thing. Fire escape discussions over port and cigars may or may not have contributed to this thesis, but they'll be remembered fondly.

Much love to Shane Murphy, my brother from another mother, who's been bringing the funk and making me laugh for as long as I can remember, for keeping me in check when I needed it and letting me be crazy when I didn't. The years of comic relief and moral support predate any of my academic aspirations by a long shot. Thanks just for being there through the years.

And to those of you who are reading this and haven't seen your name, rest assured you're not forgotten. So many people have in some way moved this project forward, with reassuring words, shared friendship, deep discussion. I thank you all.

Contents

Chapter 1

Introduction

This thesis is an investigation into the formal properties of phrase structure. In particular, theoretical tools from order theory, a well-studied branch of discrete mathematics, are brought to bear on the problem of answering the following question:

What is the best means of studying phrase structure?

The thesis borrows liberally from the vast body of work on phrase structure conducted within the generative tradition. Insights from current generative theory about the formal aspects of syntactic structure-building are pursued to their logical ends.

The upshot is that there is no "phrase structure" *per se*, at least not in the way that collocation is generally used. The output of the syntactic component (whether construed derivationally or representationally)[1] consists solely of partially-ordered sets of syntactic objects, which are themselves sets of lexical items.

[1] I remain agnostic on this issue; arguments on both sides have merit. (cf. Epstein *et al.* 1998 and Brody (1995,2002 *inter alia*)

On the empirical side, the order-theoretic approach to syntax that I develop suggests a new perspective on Pesetsky's (1995) *phrase-structure paradoxes*; a set of empirical facts that have been argued to pose significant challenges to conventional theories of phrase structure. The account I give has clear roots in Pesetsky's original solution to the problem, but combines independent insights from contemporary studies in generative syntax.

In the following sections I give a brief history of the study of phrase structure leading up to the present day, as well as detailing some recent work that is related, both in theoretical and empirical terms, to the proposals made here.

1.1 Background

The following is an abbreviated chronology of the modern study of phrase structure, from the beginning of the twentieth century to the present day:

- Bloomfield (1933) pioneers the use of **immediate constituent (IC) analysis**, arguing that sentences can be analyzed into an explicit hierarchical structure independently of meaning.

- Tesnière's **Dependency Grammar** (1959) decomposes sentences into *independent* and *dependent* elements, effectively introducing the notion of headedness/endocentricity into the study of natural language syntax.

- Chomsky (1956/7) formalizes Bloomfieldian IC-analysis in terms of string-rewriting rules, and develops a hierarchy of increasingly powerful **phrase structure**

grammars (PSGs).

- Chomsky (1970) invents X′-theory, a theory of allowable phrase structures, as an attempt to curb the descriptive power of PSGs. Jackendoff (1977) continues this work and undertakes the uniformization of rules of categorial projection, importing the notion of endocentricity from Dependency Grammar.

- Stowell (1981) begins a program seeking to eliminate construction-specific rules from phrase structure grammars and the reducing the X′ schema to its simplest form.

- Chametzky (1996) (via Speas 1990 and Grimshaw 1991) represents the culmination of Stowell's line of research, reducing phrase structure to a single rule of base-structure projection.

- Kayne (1994) attempts to derive properties of X′-theory from his *Linear Correspondence Axiom*, which maps c-command to linear precedence. This is the first explicit use of order theoretic notions in mainstream generative syntax.

- Chomsky (1994/5) derives core properties of X′-theory from the basic assumption of a recursive structure-building operation, *Merge*. His **bare phrase structure** makes explicit use of (quasi)set-theoretic representations of syntactic objects. Kayne's LCA is relegated to a phonological/phonetic constraint, restoring hierarchy as the primary syntactic ordering relation.

- Collins (1999) and Seely (2000) represent the current stage in Chomsky's program of reduction of theoretical baggage. There are neither category labels nor phrasal projection nodes in the structure-building aspect of syntax: Merge(α,β) $= \{\alpha, \beta\}$

1.2 Related Work/Antecedents

The final view of sentential structure as sets of partial orders related by order-preserving mappings has some antecedents as well. Epstein *et al.* (1998. ch.6) suggest that the construction of a syntactic object, viewed in derivational terms, can be analysed as a set of *Merge* events ordered strictly by the "must follow" relation.[2] Kracht (1993) offers an early look at some of the mathematical properties of syntactic command relations, and Bury (2003) suggests explicitly that phrase structure is best understood in terms of sets of lexical items with explicit intercategorial dominance relations.

[2] In fact, the authors miscategorize the relevant ordering relation, labelling it a *quasiorder* (preorder). Preorders are weaker than partial orders, having no symmetry conditions.

Chapter 2

Phrase Structure and Trees in Syntax

In this chapter I will examine one of the primary means of graphically representing phrase structure, namely syntactic trees. After a short historical overview, I will examine their principal formal properties and examine the role that these have played in recent syntactic theorizing.

2.1 Some history

The Bloomfieldian (1933) notion of *immediate constituency* has led to the dominant view in mainstream syntactic theorizing, which states that sentences are amenable to mereological analysis.[1] Chomsky (1957) introduced the notion of *phrase struc-*

[1] "Dominant" here certainly doesn't mean exclusive. There are several formalisms that do not (at least not explicitly) analyze sentences into a hierarchical structure, *e.g.* Dependency Grammar (Tesnière 1958).

ture grammar in an attempt to formalize the IC analyses carried out in structuralist linguistics.

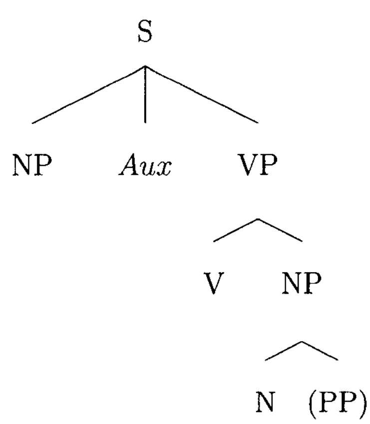
Tree diagrams were used to graphically represent the part-whole relationships between the constituents identified in an IC analysis. Thus, (a) and (b) below represent exactly the same information:

(2.1) Phrase structure rules and tree

(a)

- S → NP Aux VP
- VP → V NP
- NP → N (PP)

(b)



The trend in generative syntax (until the advent of Chomsky's (1995) Minimalist Program) was to imbue tree diagrams with increasing amounts of non-structural information. For example, the notion of *headedness*, whereby a lexical item of type X (*e.g.* nominal, verbal, *etc.*) "projects" its categorial status to a superordinate constituent, was imported from Dependency Grammar. More generally, the use of categorial information, for example having a tree specify whether a particular node is part of a nominal or verbal projection, also reflects this approach. More recently, however, the trend has been away from this and back to a purely configurational analysis of phrase structure. The following sections briefly discuss each of these approaches.

2.1.1 Making uniform trees: X′-theory

Throughout the early days of generative grammar, a bewildering array of phrase structure rules like those in (2.1a) were put forward to account for various constructions across languages. Eventually it was pointed out that the existing theory of phrase structure had no way of ruling out apparently uninstantiated structures[2] and that it failed to capture some obvious generalizations regarding left-to-right order and hierarchical properties of syntactic constituents.

Pursuing this line of inquiry, Tim Stowell (1981) attempted to unify the diverse primitives and properties of phrase structure into a simple set of constraints on syntactic structure:

(2.2) Conditions on phrase structure (Stowell 1981 p.70, cited in Webelhuth 1995)

- Every phrase is endocentric (*viz.* headed)
- Specifiers appear at the XP-level; subcategorized complements appear within X′

[2]See Hale & Reiss (2000) for comments on the question of whether or not overgeneration should be criterial for theory change.

- The head always appears adjacent to one boundary of X′
- The head is one bar level lower than the immediately dominating phrasal node
- Only maximal projections may appear as non-head terms within a phrase

This set of conditions specifies a highly restricted set of possible structures for any given syntactic phrase. The possible structures can be concisely represented with the following universal phrase structure schema, in which the language particular linear ordering of consituents is left unspecified, presumably to be derived on the basis of the primary linguistic data:

(2.3) X′ schema

- XP → ZP X′
- X′ → X YP

At this point a question arises: why should this particular set of structural properties be innately specified, rather than some other one? The next section discusses one attempt to answer this question.

2.1.2 Building trees one step at a time: *bare phrase structure*

Chomsky (1995, ch. 4) seeks to derive the structural properties of X′-theory from independent properties of algorithmic structure-building. He argues that a recursive operation that concatenates pairs of objects —termed *Merge* —is necessary in any finite system that can generate sentences of indefinite length. In his formulation, *Merge* puts its arguments into a set, then projects the categorial properties of one of them.

(2.4) *Merge(α,β)* $= \{\gamma, \{\alpha, \beta\}\}$

where $\gamma \in \{\alpha, \beta\}$[3]

From this assumption, Chomsky is able to derive basic facts about complements and specifiers as ordered applications of *Merge*.

2.1.3 21^{st} century phrase structure: eliminating labels & projections

Recent developments in syntactic theory have led to the conclusion that syntactic operations (structure-building and transformation) reduce to the application of a single operation, *Merge*, which takes two syntactic objects and forms a single object out of them. In particular Collins (1999) and Seely (2000) suggest that *Merge* does not "project" either of its arguments (and hence that there is no labelling of the resulting output), and that the simplest representation of the output of *Merge* is the minimal set containing both arguments:

(2.5) Syntactic objects and *Merge*

i. Lexical items are syntactic objects

ii. Given two syntactic objects, α and β, $Merge(\alpha, \beta) = \{\alpha, \beta\}$ is a syntactic object.

[3]This is nearly the mathematical definition of an ordered pair: $\{\{a\}, \{a, b\}\}$. Presumably this is not want Chomsky wants.

Note how this differs from the *bare phrase structure* formulation. The extra level of structure in $Merge(\alpha, \beta) = \{\gamma, \{\alpha, \beta\}\}$ (with γ equal to α or β) is meant to capture the endocentricity of phrasal structure; a concept that Collins and Seely claim is either unnecessary or redundant given the information available in a lexical entry (*e.g.* subcategorization frames).

If we take *bare phrase structure* and these more recent developments literally and pursue the logic to its end, we are led to the conclusion that there is no such thing as phrase structure. This view is explicitly advocated in Epstein *et al.* (1998), where it is argued that *Merge*(α,β) is best understood as meaning something like "put α and β into a relation".[4] Phrase markers —the outputs of the syntactic component that are subject to interpretation by the interface systems —are thus nothing more than sets of syntactic objects (themselves sets of lexical items) with some formal relations defined on them.

In the following section we will look more closely at these relations in the context of the formal properties of syntactic tree diagrams.

2.2 Formal properties of syntactic trees

Trees have been the object of much study within both mathematics and theoretical computer science since the introduction (Chomsky 1956) of formal language theory, and their formal properties are thus thoroughly understood. This section is an overview of the main formal properties of syntax trees. As the canonical definition

[4]This view is reminiscent of Dependency Grammar, which decomposes phrases into a set of relations between heads and "dependent elements".

of syntax trees takes them to be graph-theoretic objects, the reader is referred to Appendix A for an overview of the relevant technicalia.

2.2.1 Trees as graphs

Formally, a **tree** is a rooted, labelled, directed acyclic graph (Partee *et al.* 1990). I will unpack these terms in reverse order:

- **DAG:** each of these terms is defined in the preceding section. Essentially, this part of the definition specifies that the edges that connect nodes in a tree are directed[5] and that a tree has no loops.
- **labelled:** the vertices (nodes) of a tree have categorial information associated with them, generally represented with written labels, like *VP*.
- **rooted:** a tree has a privileged node r, the *root*, from which there is an oriented path to any other node, *i.e.* the root has in-degree 0, while all other nodes have in-degree 1.

In addition to the properties listed above, syntactic trees (*i.e.* trees *qua* linguistic *vs.* graph-theoretic objects) typically have two ordering relations explicitly defined on them; *dominance* and *precedence.* The first mirrors the orientation of the edges and captures hierarchical information, while the other is a left-right linear order on the set of nodes that are not ordered by dominance. Generally, there is a "non-tangling"

[5]This is typically expressed only implicitly in physical diagrams of trees using page-orientation

condition on the vertices of a tree, so that if two internal nodes are in a particular left-right order, then none of the nodes they dominate can violate that order.[6]

[6]See Coleman & Local (1991) for a discussion of graphs as abstract objects *vs.* diagrams and a discussion of whether it is possible for something like the non-tangling condition to do any theoretical work.

Chapter 3

Order-theoretic Syntax

In this chapter I sketch an approach to syntax, and to phrase structure in particular, that explicitly takes the mathematical theory of ordered sets as its formal foundation. This branch of discrete mathematics will prove to be useful in giving a formally rigorous theory of syntactic structure.

In addition to rendering more explicit the idea that phrase structure consists of partially-ordered sets of lexical items, I show that syntactic transformations may be viewed as **order-preserving maps** between phrase-structural posets. Moreover, I show that this approach to transformations accounts straightforwardly for the fact that movement is always to a c-commanding position. As the remainder of this thesis relies on the mathematical theory of ordered sets, the reader is referred to Appendix B for an overview of the relevant formal machinery.

3.1 Partial orders in syntax: Structure

In the most current versions of generative syntax, a phrase marker is viewed as a particular kind of formal structure; a set of syntactic objects with some formal structural relations defined on it. Of these relations, the one that is most relevant to the organization of phrase structure is *dominance*, which I have already discussed in the context of traditional syntactic trees.

As it happens, the view of dominance that bare-phrasal considerations —especially the label/projection-free theories of Collins and Seely —lead us to is isomorphic to the set-theoretic relation of *containment*. Containment, in turn, is the canonical example of a mathmetical relation known as a **partial order**; a relation that is transitive, antisymmetric and reflexive (or asymmetric and irreflexive, see Appendix B for details).

In conjunction with the discussion in (2.1.3), we are led to a view of syntax in which phrase markers are simply partially-ordered sets of syntactic objects. We begin by repeating the definition of *syntactic object*:

(3.1) Syntactic objects (repeated from (2.5))

i. Lexical items are syntactic objects[1]

ii. Given two syntactic objects, α and β,$\{\alpha, \beta\}$ is a syntactic object.

As mentioned above, the relevant ordering relation in this case is that of set-theoretic containment. We shall see below that sentences are in turn composed of

[1]For the purposes of this thesis, I am taking "lexical item" to include things like functional heads.

sets of phrase markers.

3.2 Syntactic trees *vs.* diagrams of partial orders

As previously discussed, phrase markers can be formally characterized as sets of syntactic objects (partially) hierarchically ordered by containment. An advantage of viewing syntactic structure in terms of trees is that tree diagrams allow for easy deciphering of hierarchical relations between different pieces of syntactic structure and are easy to construct given *e.g.* a bracketed string.[2] In short, tree diagrams are a straightforward and intuitive graphical representation of some of the key properties of the objects of study, namely the sentences that are the output of the syntactic component of the grammar.

Ordered sets also have an associated system of graphical representation, called **Hasse diagrams** (see Appendix B for details). This system retains the benefits of standard syntactic trees (*e.g.* easy recovery of non-linear relations) while making it easier to avoid some of the undesirable —or at least unnecessary —artifacts that accompany the current view of trees (*e.g.* the inclusion of non-structural information such as categorial labels/projection). The following example shows a simple ordered set and its associated diagram.

(3.2) A simple ordered set

$$\mathcal{L} = \{(A, B), (A, C), (A, D), (C, D)\}$$

[2]Of course, a bracketed string is itself a constituent analysis. The actual determination of constituents involves a variety of tests that will figure prominently later in this work.

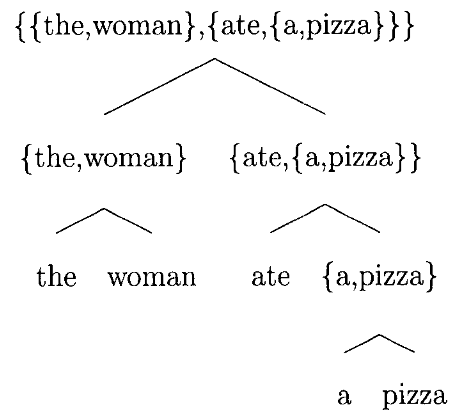
(3.3) Hasse diagram of $\mathcal{L}$

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, dash] \arrow[dr, dash] & \\ B & & C \arrow[d, dash] \\ & & D \end{tikzcd}$$

Note firstly that $\mathcal{L}$ meets the definition of partial order. It is transitive (for every pair (x, y) and (y, z) in $\mathcal{L}$, (x, z) is also in $\mathcal{L}$), asymmetric (for every pair (x, y) in $\mathcal{L}$, (y, x) is not in $\mathcal{L}$) and irreflexive (there are no pairs of the form (x, x) in $\mathcal{L}$). The diagram captures all of this information simply and straightforwardly.

Recall that the containment relation is isomorphic to the traditional syntactic relation of *dominance*; one of the two relations defined on the nodes in syntactic trees. In fact, the covering relation (see Appendix B) in a phrase marker corresponds to the traditional phrase-structural relation of *immediate dominance*, represented by the branches of a traditional syntactic tree. This effectively reduces label/projection-free trees to Hasse diagrams of phrase markers (*qua* partially-ordered sets of syntactic objects). The following example shows a simplified transitive verb phrase in this system of representation, where traditional non-leaf labels (*e.g.* VP, T', ...) have been replaced with explicit representations of the relevant syntactic objects.

(3.4) Transitive VP structure ordered by containment



As in (3.3), this diagram shows the hierarchical relations entered into by the different pieces of syntactic structure. In fact, (3.4) is quite similar to a traditional syntactic tree. In spite of this, it is clear that the two systems of representation are not the same. Traditional syntactic trees generally bear more than structural information and are augmented with *e.g.* categorial information and indices, neither of which are part of Hasse diagrams. There is an additional difference relating to the second ordering relation on syntactic trees that was mentioned earlier.

3.2.1 Precedence in phrase structure

As discussed in chapter (2), syntactic trees are rooted, directed, acyclic graphs with two ordering relations specified on their nodes: dominance and precedence. Dominance is the relation discussed above, which maps directly onto set-theoretic containment. Precedence on the other hand is a left-to-right linear ordering of the nodes in a tree that are not in a dominance relation.

The fact that substantive syntactic relations (*e.g.* agreement and anaphoric dependencies *inter alia*) are licensed under c-command, as well as the structure-dependent nature of syntactic transformations seem to indicate that linear precedence plays no role in the determination of syntactic relations, whereas hierarchical information is crucial.[3] It is here that the advantages of diagrams begin to appear. I have claimed that syntactic objects are sets ordered exclusively by containment (*i.e.* there is no linear ordering of syntactic objects). If this is the case, then the same diagram can be drawn in several ways. Consider the following graphical representations of an object $\{\alpha, \beta\}$, containing α and β:

(3.5) Isomorphic representations of $Merge(\alpha, \beta) = \{\alpha, \beta\}$

a. $\{\alpha, \beta\}$ b. $\{\alpha, \beta\}$ c. $\{\beta, \alpha\}$ d. $\{\beta, \alpha\}$

$\alpha \quad \beta \qquad \beta \quad \alpha \qquad \alpha \quad \beta \qquad \beta \quad \alpha$

Each of the above diagrams represents exactly the same object —the one that contains syntactic objects α and β.[4] The fact that diagrams allow us to explicitly abstract away from the linear order of the terminal elements lets us avoid committing a common error: the attribution of accidental properties of a representation to the represented object. In this case, the fact that the medium of representation (*i.e.* the paper on which the diagram is displayed) is two-dimensional forces a (spurious) linear ordering of the terminal nodes.[5]

Notwithstanding the previous points, it is clear that utterances (sentences *qua*

[3]It has been pointed out to me (Alan Bale, p.c.) that syntactic theory retains a small vestige of linearity . The so-called *Leftness Principle* states that a variable cannot be the antecedent of a pronoun to its left. There have been some attempts to reformulate this condition in hierarchical terms (Reinhart 1983, Higginbotham 1980), with varying degrees of empirical success.

[4]This single object is perhaps best represented as something like a 3-dimensional mobile, free to rotate and thus inherently unordered.

[5]See Coleman 1998 for extensive discussion of this problem.

outputs of human beings) are in some sense "linearized". A hearer perceives a linearly ordered sequence of words, and each word is perceived to be a (more or less) linearly ordered sequence of sounds. It is common in contemporary syntactic research (cf. Chomsky 1995 *inter alia*) to see the claim that this linearization reflects properties of the phonological or phonetic systems (although Kayne (1994) argues explicitly that the syntax produces a linearly ordered object). Note, however, that linearization cannot truly be a property of the interface to physical (articulatory) output systems. Firstly, it is well-known that in natural speech it is the norm for articulations to overlap temporally. Moreover, articulators have different masses (thus move at different speeds), must move different distances, and they have neural pathways of different lengths. This means that the inititation of the muscular movements involved in *e.g.* coarticulation or simultaneous must involve a highly complex and nonlinear ordering of timing events.[6]

3.3 Partial orders in syntax: Relations

Given the conception of syntactic structure elaborated thus far, it is possible to define other relations on the syntactic objects that compose a phrase marker, one of which in particular will be of interest. I begin by defining and describing this relation, which I term $\mathcal{ACC}$, then show why it is of relevance.

(3.6) Definition of $\mathcal{ACC}(x, y)$

[6]Thanks to Mark Hale for explaining this to me.

Given a phrase marker $\langle \mathcal{L}, \subseteq \rangle$[7]

$\forall x, y \in \mathcal{L}$, (x, y) is in $\mathcal{ACC}$ iff :

(i) $x \not\subseteq y$ & $y \not\subseteq x$

(ii) $\uparrow x - \{x\} \subset \uparrow y - \{y\}$

The first clause specifies that x and y are not ordered with respect to one another (*i.e.* neither of them contains the other) and the second clause specifies that the set of objects that strictly contain x is a proper subset of the set of objects that strictly contain y. A moment's reflection shows that the $\mathcal{ACC}$ relation is the order-theoretic analogue of the traditional syntactic relation of *asymmetric c-command*—a relation central to most substantive syntactic theorizing. As it happens, $\mathcal{ACC}$ is also a partial order over the elements of a phrase marker. The following is a proof of this claim.

(3.7) **Claim:** $\mathcal{ACC}$ is a strict partial order

To show this, I need to demonstrate that $\mathcal{ACC}$ is transitive, irreflexive and asymmetric. I will tackle these in reverse order.

1. **Asymmetry:** $\mathcal{ACC}(a, b) \Rightarrow \neg\, \mathcal{ACC}(b, a)$. This follows from the use of the proper subset relation in the definition. Suppose that $\mathcal{ACC}(a, b)$. Then everything that properly contains a also properly contains b. Since $a \neq b$ (by the non-comparability clause), $\uparrow b - \{b\}$ must contain at least one element not in $\uparrow a - \{a\}$. Therefore it cannot be the case that $\uparrow b - \{b\} \subset \uparrow a - \{a\}$, a condition on $\mathcal{ACC}(b, a)$.

[7]Recall that I am taking a phrase marker to be a set of syntactic objects ordered by the containment relation.

2. **Irreflexivity:** $\neg\mathcal{ACC}(a,a)$. This follows from asymmetry. Consider an arbitrary asymmetric relation, denoted $\mathcal{R}$. Then $\mathcal{R}(a,b)$ implies $\neg\mathcal{R}(b,a)$. Assuming that a and b are the same object leads directly to a contradiction. Therefore $\mathcal{R}$ cannot be reflexive.

3. **Transitivity:** Suppose $\mathcal{ACC}(a,b)$ and $\mathcal{ACC}(b,c)$. Then I want to show that $\mathcal{ACC}(a,c)$, which in turn means I need to prove that each component of the definition holds, namely that (i) a and c are not in a containment relation, and (ii) $\uparrow a - \{a\} \subset \uparrow c - \{c\}$. Once again, I do these in reverse order.

    - $\uparrow a - \{a\} \subset \uparrow c - \{c\}$: This follows from the fact that the proper subset relation is transitive.

    - $a \not\preceq c$ & $c \not\preceq a$: (Proof by contradiction) Suppose that a is comparable to c. Then either $a \preceq c$ or $c \prec a$. I already know that $\uparrow a - \{a\} \subset \uparrow c - \{c\}$, so $a \not\preceq c$. Then $c \prec a$. Then $a \in \uparrow c$. I also know that $\uparrow b \subset \uparrow c$ (by $\mathcal{ACC}(b,c)$) and that $a \notin \uparrow b$ (by $\mathcal{ACC}(a,b)$ which implies non-comparability of a and b). Moreover, I know that b does not c-command a (by asymmetry of $\mathcal{ACC}(a,b)$). This implies that a and b must be **covered** by the same element (that is, in the structural relation traditionally referred to as "sisterhood"). But that in turn implies that $\uparrow a - \{a\} = \uparrow b - \{b\}$, which is false by $\mathcal{ACC}(a,b)$. Therefore, a and c are not comparable.

We see then, that $\mathcal{ACC}$ —the order-theoretic analogue of a primary structural relation in syntactic theory —is a partial ordering on the elements in a phrase marker, induced from the containment order.

3.4 Syntactic operations

The structural and relational syntax I have been elaborating leads naturally to a particular account of syntactic transformations which in turn yields a straightforward explanation for the fact that there is no downward or sideward movement in syntax (*pace* Nunes 2001 and Bobaljik & Brown 1997).

3.4.1 Transformations

Following the observation that phrase markers are ordered sets of syntactic objects, I suggest that syntactic transformations are **order-preserving mappings** between phrase markers. *Merge* in its structure-building guise (Chomsky's (1999) "external Merge") clearly adds monotonically to the set of ordering relations that obtain in a given syntactic object, that is, existing containment or $\mathcal{ACC}$ relations are not destroyed by the addition of material to a phrase marker. What is perhaps less clear is that movement has this property as well. In order to simplify the discussion of this claim, I will adopt without further argument the view that syntactic movement involves exactly the same operation as structure building, namely *Merge*. In particular, movement does not involve the creation of copies of lexical items, nor of traces of moved elements. A particular instance of movement simply involves a second (third, fourth ... n^{th}) instance of merger of a single syntactic object (see Starke 2001, Epstein *et al.* 1998 and Chomsky 1999 for arguments that this view of movement is correct).

In current syntactic frameworks (Chomsky 1998, 1999, 2001), movement/remerger occurs only to satisfy an "EPP-property" of some head.[8] On the view of movement I am proposing (*viz.* movement as an order-preserving mapping relating phrase markers), movement for EPP purposes looks something like the following:

(3.8) $\{a,\{b,c\}\} \longrightarrow_{epp} \{\{b,c\},\{a,\{b,c\}\}\}$

$\{a,\{b,c\}\}$: a $\{b,c\}$; $\{b,c\}$: b c

$\{\{b,c\},\{a,\{b,c\}\}\}$: $\{b,c\}$ $\{a,\{b,c\}\}$; $\{a,\{b,c\}\}$: a $\{b,c\}$; $\{b,c\}$: b c

In this example a is a head with an EPP-feature and thus requires a specifier. The mapping labelled *epp* here creates a single new $\mathcal{ACC}$ relation (and, derivatively, new dominance relations) with an element in a's complement, holding between $\{b, c\}$ and a. All other apparently new c-command relations (*e.g.* $\mathcal{ACC}(\{b, c\}, b)$, $\mathcal{ACC}(\{a, \{b, c\}\}, \{b, c\})$, *etc.*) turn out to be ruled out according to the definition of $\mathcal{ACC}$. Interestingly, the single newly-created relation is exactly the one that has been suggested as being relevant for the purposes of feature-checking (see *e.g.* Chomsky 1995). Below I will show that this account of syntactic transformations offers a new perspective on a frequently observed but as yet unexplained property of movement.

[8] This is variously interpreted as a requirement that the head have a specifier, or as an instruction to pronounce the head of a chain, rather than the tail.

3.4.2 Sidebar: the categorical view

In mathematics, the study of a set of objects with additional structure (*e.g.* an ordering relation) nearly always involves an associated set of operations that preserves the structure in some manner. The recurrent pairing in mathematics of structured objects and structure-preserving mappings relating them has led to the development of **category theory**; the generalized study of structured objects and their associated structure-preserving mappings. The view of syntax emerging here, with phrase structure elaborated in terms of ordered sets of syntactic items and (Re)*Merge* as an order-preserving mapping suggests that a category-theoretic view of linguistic theory might be worth pursuing. I do not have space to do so in this thesis, but simply point it out as a direction for future research.

3.4.3 Whence c-command?

Over the course of the last decade, there have been some attempts to "explain" c-command, that is, to derive the fact that c-command, rather than any of the other possible relations definable on a tree, is the relevant structural relation for syntactic theory.[9] In my discussion of the $\mathcal{ACC}$ relation I skirted this question entirely, simply defining the relation and showing that it was a partial ordering on phrase markers. This fact in itself is of interest, as it is clear that the vast majority of the definable

[9]A terminological note: I use "$\mathcal{ACC}$" and "asymmetric c-command" more or less interchangeably in the remainder of this thesis. In discussions of analyses from the literature I am referring to the traditional notion of c-command defined on traditional syntactic trees. In my own analyses or in the context of explicitly order-theoretic discussion, the terminology refers to the $\mathcal{ACC}$ relation. A related point holds for trees and associated terminology; traditional trees will always be drawn with categorial labels, whereas Hasse diagrams will either be labelless or have set-theoretic node labels.

relations on pairs of elements in a phrase markers are not orderings. Nonetheless, a deeper explanation would be welcome. Of the few attempts at explaining c-command in the literature there is one which fits particularly well with the view of syntax I espouse in this thesis.

Robert Chametzky (1996) suggests that we "invert" our view of c-command. That is, rather than asking which nodes are c-commanded by a given node α, we should instead examine the set of nodes that c-command it. Taking this perspective leads him to the discovery that the c-commandee and its set of c-commanders provide what Chametzky calls a "minimal factorization" of the tree; the smallest set of elements including the c-commandee from which the entire tree can be recovered. Although he does not recognize it, the fact that this subset exists is a consequence of the fact that a phrase marker is not just a partial order, but a (finite) join semilattice (see Appendix B), in which every pair of elements has a least upper bound and there is a unique maximal element. I am largely sympathetic to Chametzky's view, as it translates more or less directly into order-theoretic notions. Pursuing this is outside the scope of this thesis, so I turn now to a related observation.

It has long been claimed that syntactic movement only occurs "upward" in a tree, to a c-commanding position (see Nunes 2001 *inter alia* for a different view). To date this restriction on movement has eluded explanation and in fact seems unusual given the current Minimalist trend towards an increasingly unrestricted syntax. The suggestion put forward above that transformations are best viewed as order-preserving mappings between syntactic objects straightforwardly derives the observation that movement is to a c-commanding position. In a join semilattice, the simplest (nontrivial) order-preserving mapping that does not introduce new structure (as opposed to *external Merge*) must have the properties of the EPP-movement shown in (3.8).[10]

In the following chapter I move away from formal/theoretical argumentation and consider a particular empirical problem that has been presented as a challenge to traditional theories of phrase structure: so-called **phrase structure paradoxes**. I show that the order-theoretic view I have elaborated thus far, in conjunction with some independent analyses, provides a new perspective on the problem.

[10]As noted above, the moved element technically is not in the $\mathcal{ACC}$ relation with its base position, in virtue of the non-comparability clause. If however, a new piece of structure were to be merged in the associated position, the new c-command relation would be created.

Chapter 4

Phrase Structure Paradoxes

Generative syntax has traditionally assumed that sentences are organized into a hierarchical constituent-based structure and that only constituents (*viz.* only nodes of trees) can be the structural descriptions of transformational rules (archetypically movement). Against this theoretical backdrop, Pesetsky (1995:230) presents the following *VP*-fronting data as evidence for a particular constituent analysis of verb phrases:

(4.1) *John said he would give the book to them in the garden on each other's birthdays ...*

1. ...and give the book to them in the garden on each other's birthdays he did.
2. ...and give the book to them in the garden he did on each other's birthdays.
3. ...and give the book to them he did in the garden on each other's

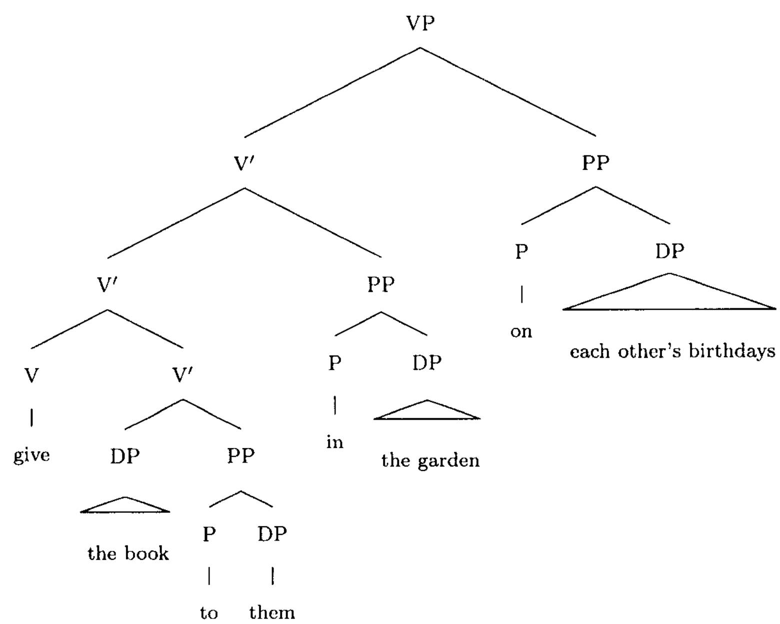
birthdays.

4. *...and give the book he did to them in the garden on each other's birthdays.

5. *...and give he did the book to them in the garden on each other's birthdays.

The point of the above lies in the possibility of stranding increasingly large rightward (following the dummy verb *did*) prepositional phrases. Each of the successfully fronted pieces of structure in the example must, by hypothesis, be a constituent and therefore the structure of verb phrases is such that nonargument PPs are adjoined above and to the right of the verb and its arguments (both DPs and PPs). This type of constituency is essentially what has been assumed in earlier *Principles & Parameters* analyses. Pesetsky refers to this structure as a *Layered* VP:

(4.2) Layered VP



4.1 The Descriptive Insufficiency of Layered *VP*s

Although the evidence for a layered VP structure is compelling, Pesetsky notes that it cannot be the whole story with respect to the constituency of the verb phrase. He presents the following data to show that the traditional analysis is problematic on several counts:

1. Coordination, which is typically considered to be a standard test for constituency, appears to be able to target structures that are not constituents in a layered analysis:

   (a) Mary gave [a book to Jim on Monday] and [a CD to Fred on Wedneday].

   (b) The Canadiens played [the Canucks in Vancouver on the 23rd] and [the Rangers in Montreal on the 30th].

2. The c-command relations in layered VPs seem to make exactly the wrong predictions about anaphor binding, bound-variable pronoun readings and negative polarity licensing. In particular, in example (4.2), it seems as though items on the right c-command those on the left,[1] but the following data show that this must be wrong:

   (a) *Anaphor binding:*

      L→R Mary danced with these people$_i$ in each other$_i$'s hometowns.

      R→L *Mary danced in each other$_i$'s hometowns with these people$_i$.

   (b) *Bound variable pronoun readings:*

      L→R Sue spoke to Mary about each employee$_i$ in his$_i$ house.

      R→L *Sue spoke to his$_i$ friends about each employee$_i$.

   (c) *Negative polarity item licensing:*

      L→R Sue spoke to no linguist about any conference.

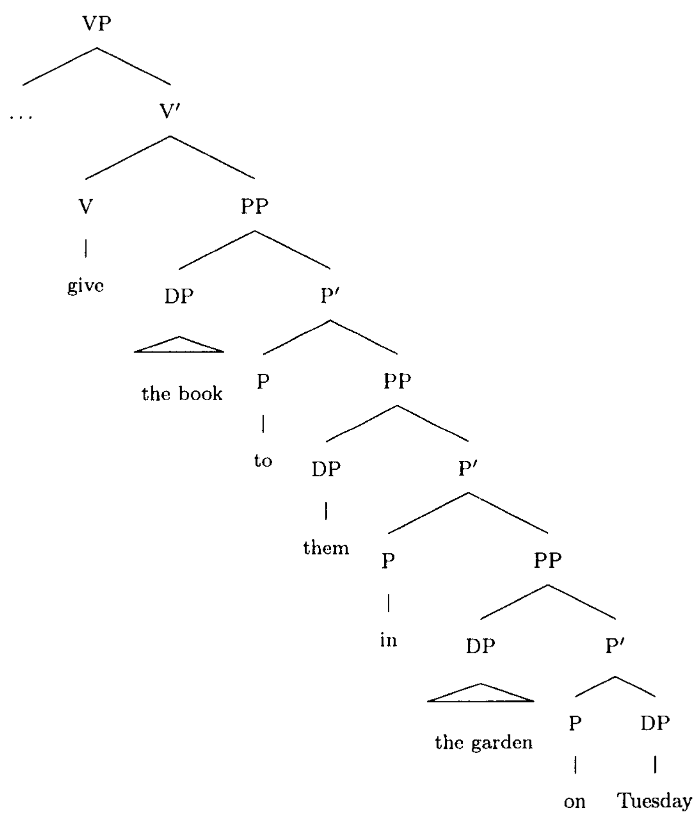
      R→L *Sue spoke to any linguist about no conference.

In each of the above pairs of sentences, there is an item whose interpretation requires that it be in the proper c-command configuration with some antecedent. In the binding cases, the reflexive *each other* must be c-commanded by a suitable local DP, in the bound variable sentences, the pronoun *his* must be c-commanded by the QP, and in the negative polarity examples, the NPI *any* must be c-commanded by a suitable quantificational element.

[1]Note that this ignores questions about whether and how complements of PPs are able to c-command out of their phrases.

Based on these facts, as well as the fact that prepositional complements in the previous sentences appear to be able to c-command out of their containing phrases, Pesetsky proposes a radically different structure for the verb phrase (inspired by Larson's (1988) VP-shell theory), which he terms a *Cascade* VP:

(4.3) Cascade VP



Note that in this structure there are prepositions and prepositional "complements" that are not in the sisterhood relation, *e.g. in* and *the garden.*

The Cascade VP structure solves the problems faced by Layered VPs. The c-command relations that can be read off the structure exactly fit the observed binding facts and NPI licensing data. Note in particular that "objects" of prepositions c-command everything to their right and there is no longer a need to worry about whether or not prepositions are invisible for the computation of c-command relations. In addition, all of the pieces of structure that (problematically) showed up as potential conjuncts are in fact constituents in a cascaded *VP*.

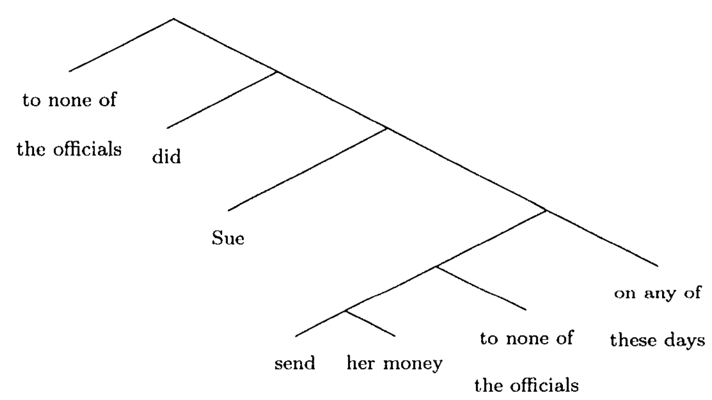
4.2 The Paradox

Notwithstanding the successes of Pesetsky's Cascade VPs, some problems remain. Consider the following sentence:

(4.4) *To none of the officials* did Sue send her money ⟨*to none of the officials*⟩ on any of these days.

On the one hand, the moved element clearly corresponds to a constituent in a Layered VP, but to a non-constituent in a Cascaded VP, since there is rightward stranded material.

(4.5)

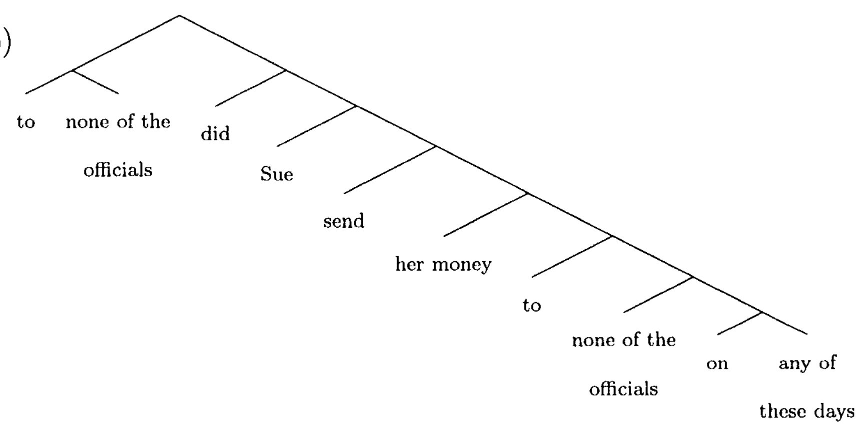


Thus, we have good evidence for layered structure.

On the other hand, *any* in the rightward stranded PP gets a negative polarity reading. On standard assumptions this requires c-command by a negative element, presumably *none* since it is the only clear candidate. Given the presence of *Aux*-inversion, it is standard to assume that movement of the fronted element is to a specifier position high in the left periphery. Given this, the required c-command relation cannot hold post-movement, since *none* is embedded inside the fronted object, as can be clearly seen from the tree in (4.5).[2]

This means that the c-command relation must hold from the reconstructed position of the PP. But we can see from the tree above that a Layered VP yields the wrong c-command relations. What is required is for the VP to have a Cascaded structure.

(4.6)



We thus have simultaneous evidence for both Layered and Cascaded verb phrase constituency. It is this clash of phrase structures that has come to be known as *Pesetsky's Paradox.*

[2]One may wonder how this situation differs from that in *None of the boys ate any cake.* A plausible explanation is that in sentences like the latter one, the subject constituent is arguably a QP that bears negative force on its own, whereas in the example in the text, the NQP is further embedded inside a PP. This alone is enough to ensure that the requisite c-command relation fails to hold.

4.2.1 Dual-System Syntax

Pesetsky suggests that both Layered and Cascade phrase structures are real (in the sense of being part of the proper analysis of the phenomenon in question), and that at least those sentences that show evidence of paradoxical structure have both structures simultaneously associated to them.

> [...] I will suggest that Cascade Syntax and Layered Syntax are not competing proposals about the hierarchical organization of sentences, but represent aspects of this organization that are **both** relevant to syntactic phenomena. (pp.230, emphasis mine —FM)

Pesetsky, while being careful to point out that the two types of structure are not derivationally related, establishes a number of mapping relations between Layered and Cascade VPs, *e.g.* a movement such as VP-fronting in Layered syntax is followed by a restructuring of the associated Cascade.

In the next chapter I will argue that it is, in fact, possible to maintain a unitary view of syntactic structure, and that the facts that Pesetsky presents as problematic can be accounted for by means of the order-preserving mappings I claimed earlier are the heart of transformational syntax. My argument depends on some claims put forward by other researchers, so I will begin the following chapter by outlining the relevant theoretical proposals.

Chapter 5

Paradox Lost

5.1 Introduction

In this chapter I will argue that it is possible to give an account of Pesetsky's Paradox that does not rely on the construction of parallel structures. In particular, I will show that the structures that demonstrate the allegedly paradoxical properties are, in fact, related by the kind of order-preserving mapping that I have claimed underlies current theories of syntactic merger and movement. The account I will provide borrows insights from several researchers, so the chapter begins with a discussion of the relevant features of each proposal.

5.2 Phillips 2003: transient constituents and snapshot tests

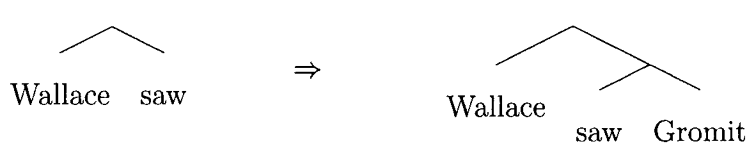
Adopting a novel view of syntactic computation, Colin Phillips (2003) suggests that it is possible to retain a unitary structural analysis of sentence structure in the face of apparent paradoxes if one adopts the following view of syntactic derivation:

(5.1) *Incrementality Hypothesis* (Phillips' (11))

Sentence structures are built incrementally from left-to-right.

In addition to the Incrementality Hypothesis, Phillips proposes two additional constraints that ensure that new objects are introduced at the bottom of the structure. The key property of this form of structure building is that constituents are created and subsequently destroyed in the course of the derivation of a sentence. In light of this, Phillips claims that constituency diagnostics that yield conflicting evidence simply apply at different points in the derivation. Consider the following example derivation:

(5.2) Simplified incremental derivation



In this simplified derivation, we see the constituent *Wallace saw* being created in the first step, and then being destroyed when the object of *see* is merged into the bottom of the tree.[1] At any stage in a derivation the structure created is a licit one and can be diagnosed as a constituent by an appropriate test applied at that point.[2] In our above example, we might imagine a similar sentence with a coordinate structure: *Wallace saw and Mary kicked Gromit.* A notable fact about this type of syntactic derivation is that it is monotonic, in the sense that it can never change an asymmetric c-command relation once it has been established. This in turn means that syntactic relations mediated by c-command are never undone in the course of a derivation. This is clearly reminiscent of the order-preserving character of the mappings between syntactic objects in the order-theoretic framework I elaborated in previous chapters. Since order-preservation is transitive, no derivationally related structures can have conflicting order relations.

The main intuitions I will be drawing from Phillips' account are that structural grouping can be temporary in the course of syntactic computation, and that the particular constituencies revealed by various structural tests are like "snapshots" of a derivation —indicative only of the phrase marker's structure at a specific point in an incremental derivation.

[1]Pesetsky only refers to the structure of the *VP* in his book, so it's unclear what predictions, if any, he would make concerning the potential constituency of *Subj-Verb* or *Subj-Infl*.

[2]An interesting question is whether it is ever possible to control the derivational point at which our tests apply, or whether it is possible to design a test that applies at a particular point in the derivation.

5.3 Epstein *et al.* 1998: interpretation in a level-free syntax

For the last decade, Sam Epstein and his colleagues have been elaborating a particular view of syntactic computation, the core tenets of which are nicely summarized by the title of their main monograph, *A Derivational Approach to Syntactic Relations* (Epstein *et al.* 1998, *DASR*). The key assumption underlying *DASR* is that syntax has no "levels of representation", but rather that syntactic computation (*viz.* structure-building and movement operations) proceeds algorithmically and incrementally and that interpretive operations (*e.g.* SpellOut) apply in lockstep with the derivation. This involves a reconceptualization of several aspects of syntax that have traditionally been framed in representational terms, chief among these being the definition of core structural relations such as c-command. I will not say anymore about the derivation/representation dichotomy here, as it is not the target of my borrowing. Instead I want to look at the *DASR* account of some facts pertaining to structurally-mediated interpretation.

5.3.1 Derivationalism and binding: quantifying over syntactic relations

In Chapter 2 of *DASR*, the authors develop a derivational account of such interpetive operations as reflexive binding and variable binding. The core of this account is the so-called *interpretive binding theory*, in which the indexing and interpretive aspects of binding are unified and coded into the binding conditions themselves. For example, Condition A is reformulated as *If **a** is an anaphor, interpret it as coreferential with some c-commanding phrase in [a relevant local domain -fm]* (*DASR*, pp. 47). In conjunction with the level-free approach to syntactic computation, this view straightforwardly accounts for many aspects of anaphoric dependencies. As an example, consider the following pair of sentences:

(5.3) A reconstruction asymmetry (Esptein *et al.* 1998, pp.63)

a. John wondered [which picture of Bill] he saw *t*

b. John wondered [which picture of himself] Bill saw *t*

In the first sentence, *he* and *Bill* cannot be coreferential. The standard account of this is as a Condition C violation, which is in turn predicated on obligatory reconstruction of the moved *wh*-phrase to its base position. On the other hand, *himself* in the second sentence can take either *Bill* or *John* as an antecedent. This is explained in terms of optional reconstruction for the moved phrase, so that the reflexive is either bound by the higher or lower DP, depending on whether or not reconstruction has taken place.

On the derivational approach, the quantificational words in the interpretive binding conditions are taken to apply literally over stages in the computation of the phrase marker. In particular, the conditions for disjoint reference must apply at every step of the derivation, while those for anaphoric dependencies must apply in at least one stage of the computation. Under a theory of movement that involves identical material at the base and target positions (*i.e.* movement as copying, or Remerge, which *DASR* assumes), it is clear that *Bill* in its base position will always be c-commanded by *he*[3], and hence the disjoint interpretation can apply. In the second case, there is at least one stage in the derivation at which *himself* is c-commanded by *Bill* (namely, when *Bill* is merged into the structure), and at least one stage at which it is c-commanded by *John.* Consequently, the anaphoric dependency is free to apply at either of these stages, giving rise to the optionality of reference.

The crucial aspect of the above for my account of phrase structure paradoxes is that interpretive procedures (*e.g.* establishment of anaphoric dependencies or disjoint reference) are framed in terms of principles that quantify over stages of a derivation. This will require a slight restating into a formulation that is less tied to the derivational framework of *DASR.*

5.4 Kayne 2004: prepositions as probes

The final piece of the puzzle comes from a direction that Richard Kayne has been exploring for a few years, but particularly from suggestions in Kayne (2004). In that article, particular facts about French causatives lead him to the following conclusion:

> Prepositions are not merged with what we think of as their objects. Rather, prepositions enter the derivation *outside VP* and subsequent to merger of a [...] licensing functional head. (Kayne 2004, pp.206)

Kayne further suggests that the higher functional head attracts the prepositional "complement" into its SPEC. In order to derive the correct word order (PP following

[3]Of course this is only true after *he* has been merged into the structure.

the VP), the preposition itself has something akin to an EPP-feature and probes for the VP, attracting it to SPEC,P. This interleaving of remnant movements has been called the "roll-up derivation" by its proponents (cf. Szabolcsi & Koopman 2000).

The notion that prepositions may be neither base-generated inside the verb phrase nor merged directly with their complements will form the core substance of the account of phrase structure paradoxes that I turn to now.

5.5 An order-theoretic account

Having set up the necessary background theory in full, I will now proceed to a new view of phrase structure paradoxes. To that end, I begin by laying out the building blocks of the analysis. (i) Order theory, a branch of discrete mathematics, is particularly apt for the study of syntax in general, and phrase structure in particular, as it provides a unified framework from which to understand traditional tree structures and the transformations that relate them; from an order-theoretic perspective, a phrase marker is a set of syntactic objects partially ordered both by set-theoretic containment and by asymmetric c-command and a sentence is a set of phrase markers related by order-preserving mappings, (ii) there is a set of empirical facts —Pesetsky's Paradox —that appears to require either parallel, conflicting constituent analyses given a traditional approach to syntax (Pesetsky 1995), or else a radical reformulation of the architecture of syntax (Phillips 2003), (iii) three independent research directions in current syntactic theory; transient constituency (Phillips 2003), level-free interpretive binding (Epstein *et al.* 1998), and remnant-movement derived VP-external prepositional phrases (Kayne 2004) can be incorporated into the order-theoretic view of syntax to provide a novel analysis of phrase structure paradoxes.

An outline of the account is as follows:

- Underlying phrase structure involves the VP-external merger of prepositions and associated case-licensing functional heads.

- The prepositions and their licensing heads both act as probes with EPP properties, alternatingly attracting non-argument DPs (prepositional "complements") and the VP remnant.

- The stepwise attraction of the prepositional complements creates the configurations necessary to license the anaphoric dependencies that lead Pesetsky (1995) to propose Cascade structure, given (a representational view of) the interpretive level-free approach to binding.

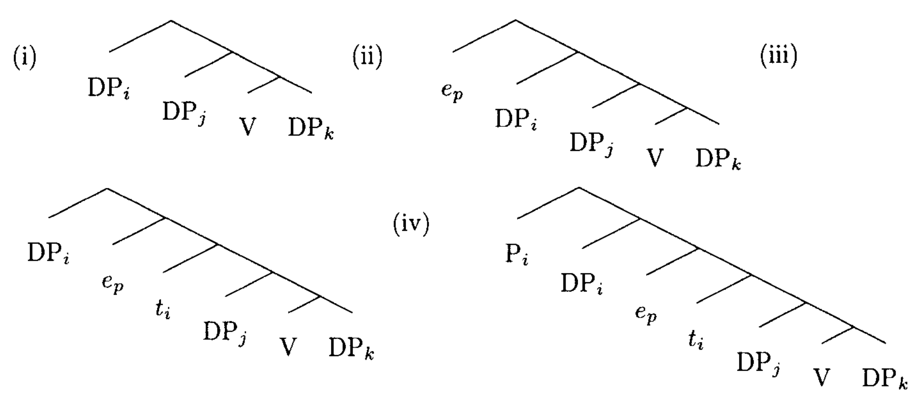
- The endpoint of the mappings described above yields structures with the same VP-fronting properties that motivated Pesetsky's Layered structure.

- A unitary analysis of phrase structure (as a set of phrase markers related by order-preserving mappings) can be preserved in the face of apparently paradoxical phrase structure.

5.5.1 Underlying and derived structures

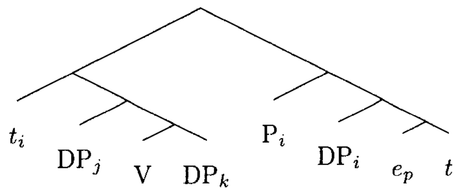
Following Kayne (2004), I assume that prepositions are introduced into a phrase marker VP-externally and subsequently to the merger of an associated phonetically null case-licensing functional head (which I designate e_p). The mergers of both prepositions and their associated heads give rise to remerger of lower structure (in current terms, both have features with EPP-properties). The interleaving of merger and movement transformations builds a series of fairly complex structures which are collectively responsible for the data that first led to the formulation of Pesetsky's Paradox. The following example is a fairly explicit sketch of the structures and mappings of a verb phrase with two adjunct prepositional phrases[4]. In the next section I will show that the derived structures yield a simple account of the data pointed out by Pesetsky, when taken in conjunction with the particular theoretical proposals discussed earlier in this chapter.

(5.4) VP structures and mappings[5]

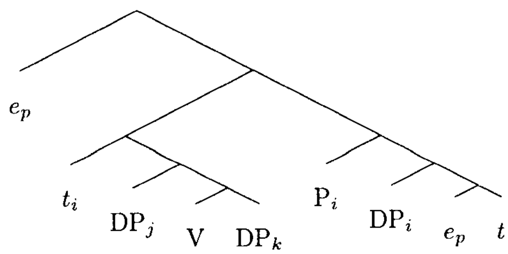


[4]I abstract away from the middle field and left peripheries, as well as from subjects and "light" verbs as these are not directly relevant to my account.

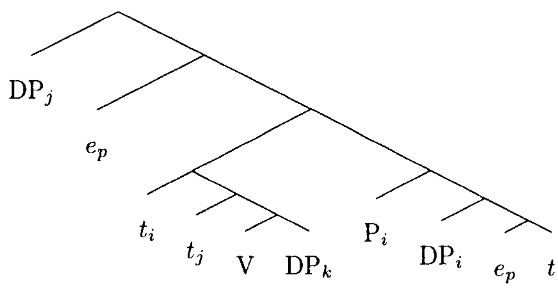
[5]I assume that theta-theoretic reasons force merger of verb arguments closer than merger of adjunct DPs. I have also indicated traces of movement for expository convenience only. Recall that I am assuming a *Remerge*-style theory of movement with identical structure at the base and target sites. See Szabolcsi & Koopman 2000 for arguments about the size of the constituent affected by remnant movement.



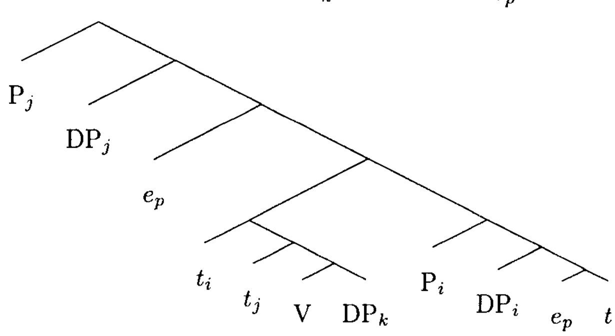
(vi)



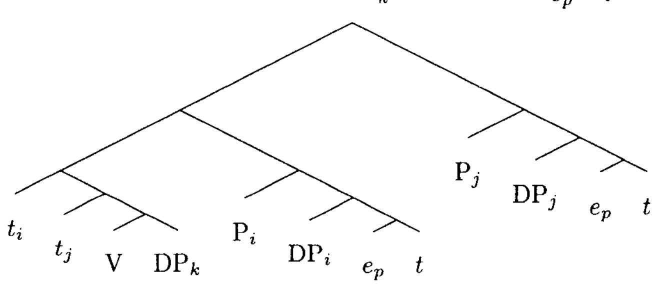
(vii)



(viii)



(ix)



As pointed out above, the alternating mergers and EPP-style movements quickly give rise to a rather complicated structure, but one which can nonetheless be shown to account simply for a variety of interpretive effects and "stranding" data.

5.5.2 Stranding "prepositional phrases"

The presentation of Pesetsky's Paradox began with a discussion of VP-fronting data which showed that in a verb phrase with adverbial PPs increasingly large pieces of rightward structure can be stranded. More specifically, it was shown that the verb and its internal argument form a constituent to the exclusion of the adverbial PPs, and that the verb, internal argument and inner/leftmost PP form a constituent to the exclusion of the rightmost PP, or that the entire group can form a constituent, but that the two PPs cannot form a constituent to the exclusion of the verb and its argument. The conclusion drawn by Pesetsky is that something like a traditional VP structure, in which nonargument PPs are adjoined above and to the right of the verb-argument constituent, must be correct. Simple inspection of the final derived structure above shows that it also accounts for the relevant empirical observations; each of the nodes along the main left branch can serve as a potential target of movement. In fact the derived structure bears a close resemblance to the Layered structure suggested in Pesetsky (1995), the main differences being that the VP already has extracted structure, and that prepositions do not form constituents with their complements.

I turn now to the empirical phenomena that led Pesetsky to posit the parallel "Cascade" structure for these phrases, namely the licensing of anaphoric dependencies between the various PP complements.

5.5.3 Anaphoric prepositional complements

I begin by repeating some of the basic data that pose a problem for the traditional analysis of verb phrase structure, as captured by Pesetsky's layered analysis:

(5.5) Anaphor binding

Mary danced with these people$_i$ in each other$_i$'s hometowns.

(5.6) Variable binding

Sue spoke to Mary about each employee$_i$ in his$_i$ house.

(5.7) NPI licensing

Sue spoke to *no* linguist about *any* conference.

In each of the above sentences, an element is dependent on a c-commanding antecedent for its interpretation; the reflexive in the first, the variable in the second, and the NPI in the last. In each case, however, the relevant c-command relation appears to go from left to right in the verb phrase, seemingly contradicting the constituency argued for above on the basis of VP-fronting facts.

Recall that I am assuming that something like the interpretive binding theory adopted by Epstein *et al.* (1998) is correct, and in particular that the binding conditions are stated in terms of quantification over possible stages of application. In the *DASR* framework, this quantification is over stages in a stepwise derivation of phrase structure. In order to dovetail with the theory of syntax I have put forward in this thesis, this view must be slightly modified.

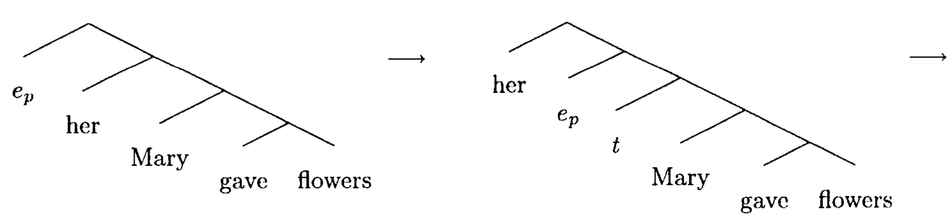
I have suggested that phrase markers are partially-ordered sets of syntactic objects and that transformations are order-preserving mappings between phrase markers.[6] On this view, a complete sentential structure consists of a set of phrase markers along with the order-preserving maps that relate them. It is a short step from here to seeing that the quantificational statements in the derivational interpretive binding theory can be made over the set of individual phrase markers that make up a complete sentential structure.

To see how this applies to our problem, consider a sentence like the following:

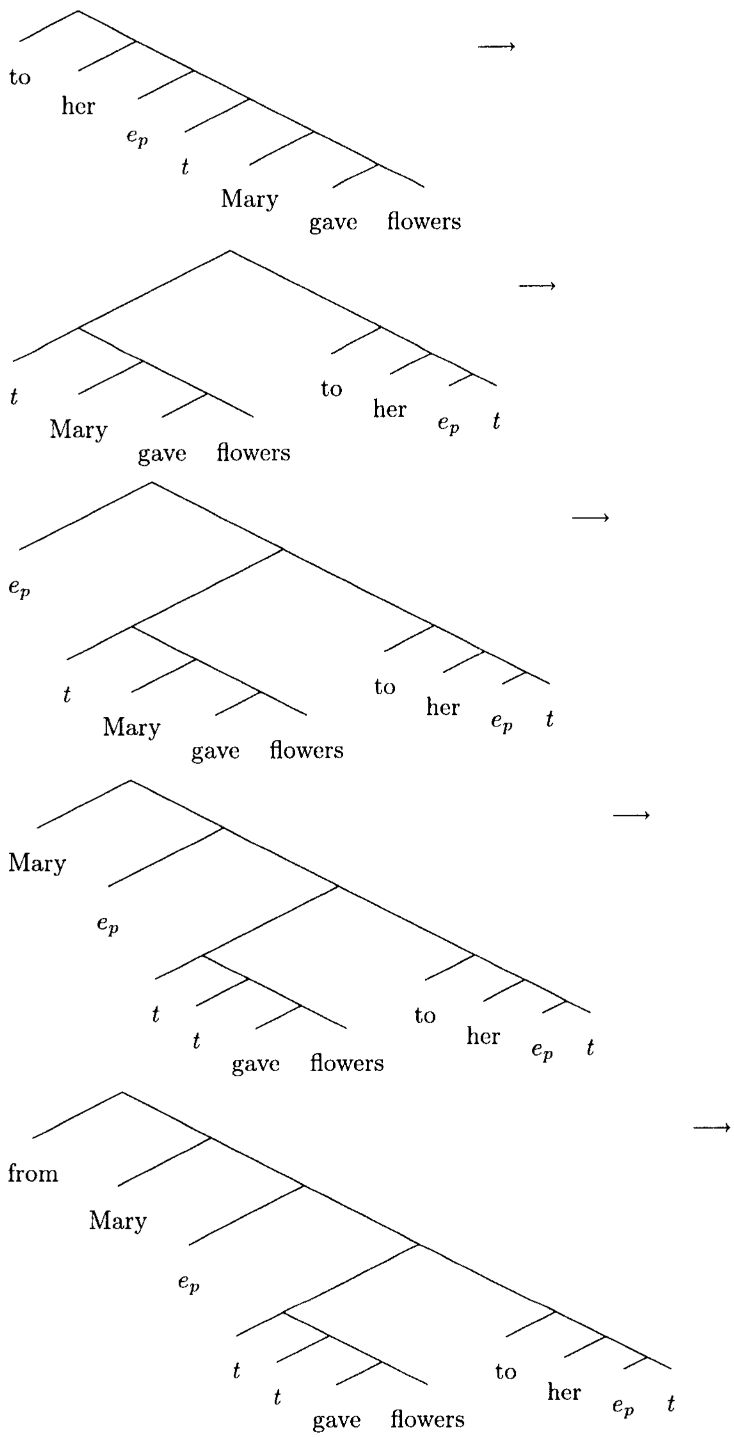
(5.8) *John gave flowers to her$_i$ from Mary$_i$.

The impossibility of coreference here will be shown to be a Condition C violation. As this is a principle that conditions disjoint reference, I am assuming (following *DASR*) that the relevant structural configuration must obtain at every possible instance in the sentential structure. Once again abstracting away from irrelevant details and using traces to simplify the diagram, the sentential structure is as follows:

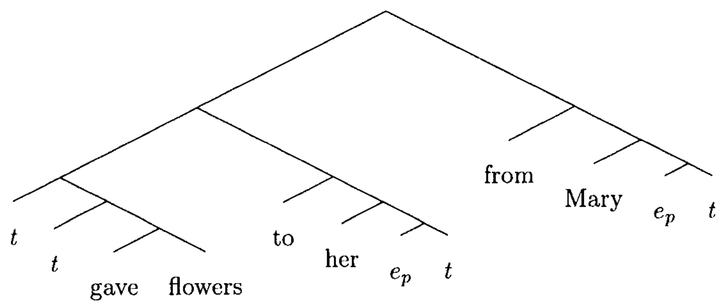
(5.9) Disjoint reference



[6] Thereby hinting at the relevance of category theory to a formally rigorous analysis of current generative syntax.



to her e_p t Mary gave flowers
t Mary gave flowers to her e_p t
e_p t Mary gave flowers to her e_p t
Mary e_p t t gave flowers to her e_p t
from Mary e_p t t gave flowers to her e_p t

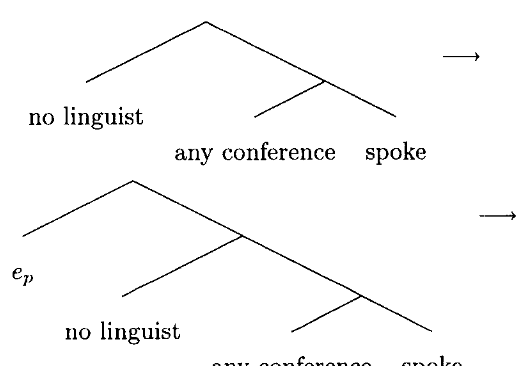


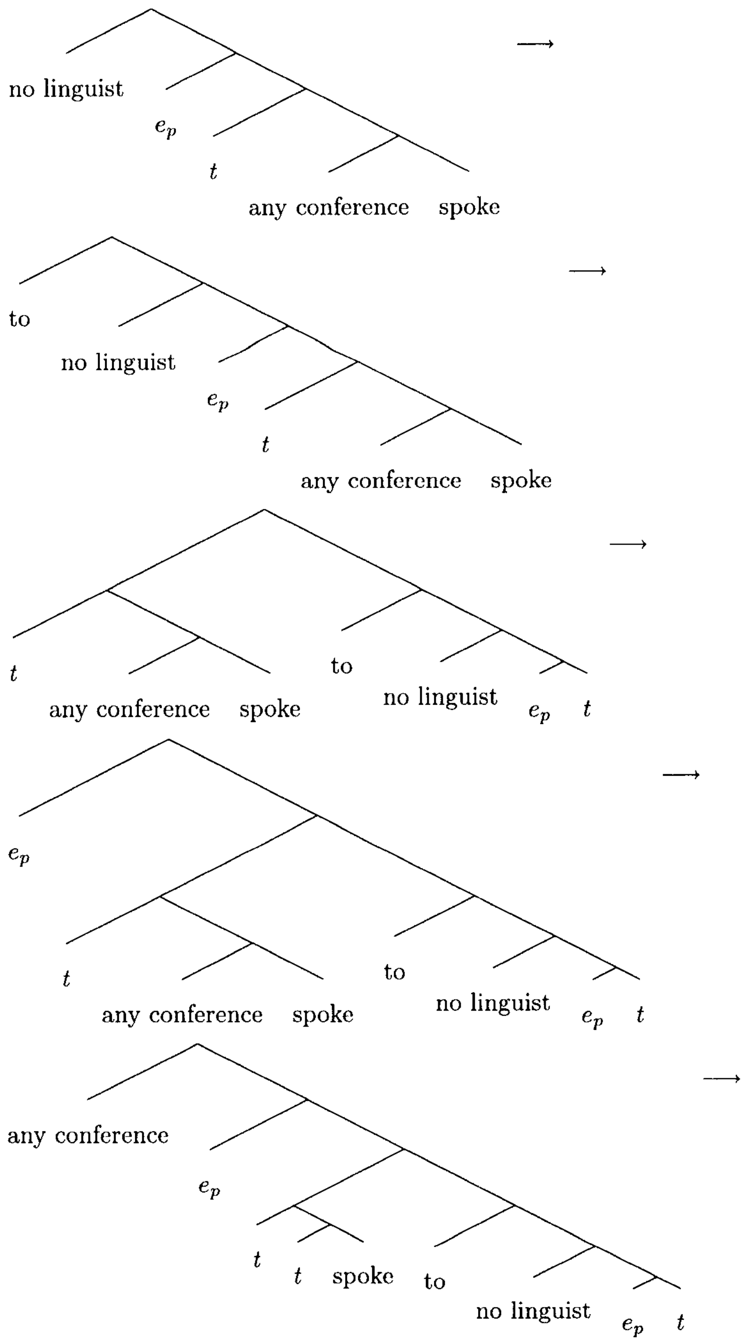
The relevant factor here is the order of merge of the nonargument DPs. Given the theory of movement I am assuming —in particular that transformations are monotonic with respect to the establishment of c-command relations —*her* c-commands *Mary* in every phrase marker in the preceding sentential structure. I will finish this chapter with an example of an anaphoric dependency, for which the relevant structural relation must obtain *at least* once in a sentential structure.

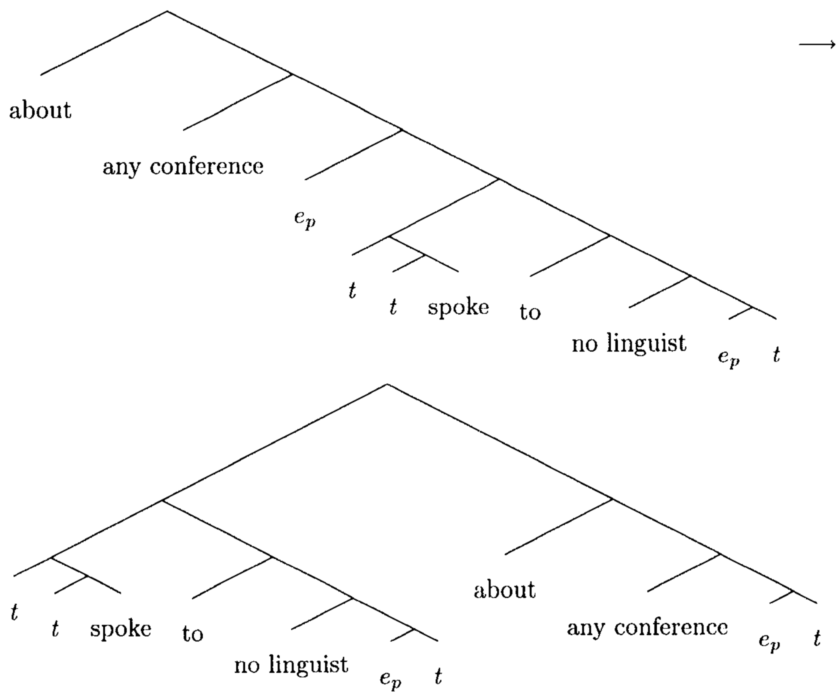
(5.10) John spoke to *no* linguist about *any* conference.

Any in this sentence receives a negative polarity reading, which arguably involves some type of anaphoricity. If this is correct, then the proper c-command relation must hold in at least one of the phrase markers that collectively characterize this sentence. Once again, the structure:

(5.11) Negative polarity




no linguist
e_p
t
any conference
spoke
to
no linguist
e_p
t
any conference
spoke
t
any conference
spoke
to
no linguist
e_p
t
e_p
t
any conference
spoke
to
no linguist
e_p
t
any conference
e_p
t
t
spoke
to
no linguist
e_p
t



In order for *any* to be licensed as a negative polarity item, it must be c-commanded in at least one of the phrase markers either by a sentential negative marker or else by a quantificational phrase that carries negative force. In this case, *no linguist* fulfills the latter role. As it is merged in a position that c-commands the NPI, the relevant structural configuration is immediately satisfied.

The next chapter summarizes the findings of this thesis and examines some directions for future inquiry.

Chapter 6

Conclusions

This thesis has involved both theoretical and analytical investigations. On the theoretical side I explored a new approach to various aspects of syntactic theory; structural, relational and transformational. This involved not so much the introduction of a new system as the reinterpretation of various aspects of current syntactic theorizing from the point of view of order theory, a long-studied and theoretically rich branch of discrete mathematics. The fit seems to be rather nice. I tightened up the view of syntactic objects as partially ordered sets of lexical items, clarified the formal role that order plays in other areas of syntax, showing that core syntactic relations like c-command and basic transformations (merger and movement) could also be incorporated into the order-theoretic view. The view I have elaborated opens the door to the culmination of nearly a century of formal investigation into the structural aspects of natural language syntax.

In terms of empirical results, I took the order-theoretic perspective and applied it to a decade-old puzzle which had been claimed to be problematic for traditional theories of phrase structure. Incorporating some independent analytical insights concerning the interpretation of syntactic structure and the order of merger of particular substantive categories, I was able to show that the so-called phrase structure paradoxes could be resolved in a unified and elegant manner, without recourse to parallel structures (cf. Pesetsky 1995) or a radical reorganization of the syntactic component of the human faculty of language (cf. Phillips 2003).

Of course this work has only begun to scratch the surface of the areas in which the order-theoretic view of syntax may shed light on old problems, or clean up old solutions. Two particular avenues of exploration that may prove interesting are the category-theoretic interpretation alluded to in Chapter 3, and the question of whether or not the complete sentential structure is itself ordered by the merger and movement transformations. This latter question has begun to receive some attention in derivational frameworks.

References

Bloomfield, L. (1933) *Language.* Revised from 1914 edition. New York: Holt.

Bobaljik, J. and S. Brown. (1997) Interarboreal Operations: Head Movement and the Extension Requirement. *Linguistic Inquiry*, 28.

Brody, M. (2003) *Towards an Elegant Syntax.* Routledge. London, UK.

Brody, M. (2002) Derivation and representation in syntax. In Epstein, S. & D. Seely (2002).

Bury, D. (2003) Phrase Structure and Derived Heads. PhD diss. University College London.

Chametzky, R. (1996) *A Theory of Phrase Markers and The Extended Base.* SUNY Press. Albany, NY.

Chametzky, R. (2000) *Phrase Structure: From GB to Minimalism.* Blackwell Publishers. Malden, MA.

Chomsky, N. (1956) Three Models for the Description of Language. *IRE Transactions on Information Theory* **IT-2**

Chomsky, N. (1957) *Syntactic Structures.* Mouton. The Hague.

Chomsky, N. (1970) Remarks on Nominalization. In Studies on Semantics in Generative Grammar, The Hague: Mouton, 1972, 11-61.

Chomsky, N. (1995) *The Minimalist Program.* MIT Press. Cambridge, MA.

Chomsky, N. (1999) Beyond Explanatory Adequacy. In *MIT Occasional Papers in Linguistics*, 20.

Coleman, J. (1998) *Phonological Representations: their names forms and powers.* Camberidge University Press. Cambridge UK.

Collins, C. (2002) Eliminating Labels. In Epstein, S. & D. Seely (2002).

Davey, B.A. and H.A. Priestley (2002) *Introduction to Lattices and Order.* (2nd ed.) Cambridge University Press, Cambridge UK.

Epstein, S., E. Groat, R. Kawashima and H. Kitahara (1998) *A Derivational Approach to Syntactic Relations.* Oxford University Press, New York NY.

Epstein, S. and T.D. Seely (2002) *Derivation and Explanation in the Minimalist Program.* Blackwell Publishers. Malden, MA.

Frank, R. and K. Vijay-Shanker (2001) Primitive C-Command. *Syntax* 4:164-204.

Grimshaw, J. (1991) Extended Projections. ms. Rutgers.

Jackendoff, R. (1977) *X' Syntax: A study of phrase structure.* MIT Press. Cambridge, MA.

Kayne, R. (1994) *The Antisymmetry of Syntax.* MIT Press. Cambridge, MA.

Nunes, Jairo (2001) Sideward Movement. *Linguistic Inquiry*, 32.2.

Partee, B., A. ter Meulen and R. Wall (1990). *Mathematical Methods in Linguistics.* Kluwer Academic. Boston, MA.

Pesetsky, D. (1995) *Zero Syntax: Experiences and Cascades.* MIT Press, Cambridge MA.

Phillips, C. (2003) Linear Order and Constituency. *Linguistic Inquiry*, 34.1.

Seely, D. (2000) Projection-free Syntax. ms. Eastern Michigan University.

Speas, M. (1990) *Phrase Structure in Natural Language.* Kluwer Academic. Boston, MA.

Starke, M. (2001) Move dissolves into Merge: A theory of locality. PhD diss. University of Geneva.

Stowell, T. (1981) Origins of Phrase Structure. PhD diss. MIT.

Tesnière (1959). *Éléments de Syntaxe Structurale.* Klincksieck, Paris.

Appendix A

Basics of Graph Theory

(A.1) A **graph** G is a pair (V, E), where V is a set of *vertices* (points in a metric space) and $E \subseteq V \times V$ is a set of pairs $e_{ij} = (v_i, v_j)$ representing edges connecting vertices.

For the following definitions I will use G to denote an arbitrary graph.

- For any edge $e_{ij} \in E$, v_i and v_j are said to be *adjacent.*
- If the members of E are ordered pairs, $< v_i, v_j >$, then the edges (and by extension G) are called *directed.* For an edge $< v_i, v_j >$, v_i is called the *initial* vertex, and v_j the *final* vertex. A graph whose edges are not directed is called *undirected.*
- A *path* in G is a sequence of edges $e_0 \ldots e_n$ such that the terminal vertex of e_{i-1} is the initial vertex of e_i.

- G is *connected* if it has undirected paths[1] connecting every pair of distinct vertices, $e_i \neq e_j$.

- G is *cyclic* if it contains at least one path $e_0 \ldots e_n$ such that the final vertex of e_n is the initial vertex of e_0 (alternatively, if there is at least one edge that can be removed without disconnecting the graph). A graph that is not cyclic is called *acyclic.*

- An element of V is said to be of *degree-n* if it is adjacent to n other nodes.

- In a directed graph, the *in-degree* of vertex v_i is the number of edges terminating at v_i, while the numbers of edges beginning at v_i are its *out-degree.*

- Nodes of degree one are called *leaves.* Nodes that are not leaves are *internal.*

[1] Meaning that if G is directed, we ignore the orientation of its edges.

Appendix B

Mini Course in Order Theory

In this appendix I give a brief introduction to those concepts of order theory which are used in the thesis. The presentation here is inspired by, and draws heavily from, that of Davey & Priestley (2002).

B.1 An introduction to ordered sets

Given a set P and a relation $\preceq$ on P obeying one of the following sets of properties for all $a, b, c \in P$, then $\preceq$ is called a **partial order**, and the pair $\langle P, \preceq \rangle$ is called a *partially-ordered set*, or *poset*.

- Weak partial order:
  - Transitive: $a \preceq b$ and $b \preceq c \Rightarrow a \preceq c$
  - Reflexive: $a \preceq a$
  - Antisymmetric: $a \preceq b$ and $b \preceq a \Rightarrow a = b$
- Strict partial order:
  - Transitive: (see above)
  - Irreflexive: $a \npreceq a$
  - Asymmetric: $a \preceq b \Rightarrow b \npreceq a$
- If $a \preceq b$ or $b \preceq a \; \forall a, b \in P$, then $\preceq$ is called **total** (*e.g.* natural numbers ordered by $\leq$)
- If there is no condition on symmetry, the resulting relation is called a **preorder**, or **quasiorder**

(B.1) The **covering** relation

Given a poset $\langle P, \preceq \rangle$, and elements $a, b \in P$, b **covers** a, written $a \multimap b$, iff $a \preceq b$ and there is no element $c \in P$, $c \neq b$ and $c \neq a$, such that $a \preceq c$ and $c \preceq b$.

(B.2) An **upper bound** of a subset

Given a poset $\langle P, \preceq \rangle$ and a subset Q of P (with the induced order), an element $p \in P$ such that $q \preceq p, \forall q \in Q$ is an **upper bound** of Q. The set of all upper bounds of Q is denoted Q^u

(B.3) An **lower bound** of a subset

Given a poset $\langle P, \preceq \rangle$ and a subset Q of P (with the induced order), an element $p \in P$ such that $p \preceq q, \forall q \in Q$ is a **lower bound** of Q. The set of all lower bounds of Q is denoted Q^l.

(B.4) The **join** of two elements, a and b

Given a poset $\langle P, \preceq \rangle$, and elements $a, b \in P$, the **join** of a and b, written $a \vee b$, is the smallest element that is greater than both a and b, that is, it is the **least upper bound** of a and b. Note that $a \vee b$ may not exist for any particular pair of elements.

(B.5) The **meet** of two elements, a and b

Given a poset $\langle P, \preceq \rangle$, and elements $a, b \in P$, the **meet** of a and b, written $a \wedge b$, is the greatest element that is less than both a and b, that is, it is the **greatest lower bound** of a and b. Note that $a \wedge b$ may not exist for any particular pair of elements.

(B.6) A **lattice**

A poset $\langle P, \prec \rangle$ for which $a \wedge b$ and $a \vee b$ exist for every distinct pair of elements a and b is called a **lattice**.

(B.7) A **join-semilattice**

A poset $\langle P, \prec \rangle$ for which $a \vee b$ exists for every distinct pair of elements a and b is called a **join semilattice** (a meet semilattice is defined analogously). Note that every finite join semi-lattice has a unique maximal element.

B.2 Hasse Diagrams

Hasse diagrams are a graphical means of visualizing the structure of a poset. Their construction is simple, following only two rules:

(B.8) Constructing a Hasse diagram for a poset $\langle P, \prec \rangle$

i. if $a \preceq b \in P$, then draw a point corresponding to a **lower than** a point corresponding to b

ii. if $a \multimap b$ or $b \multimap a \in P$, then draw a line segment connecting a and b

We give a Hasse diagram for a well-known example in order theory, the powerset of a, b, c, ordered by containment:

$$
\begin{tikzcd}[arrows=dash]
 & \{a,b,c\} \arrow[dl] \arrow[d] \arrow[dr] & \\
\{a,b\} \arrow[d] \arrow[dr] & \{a,c\} \arrow[dl] \arrow[dr] & \{b,c\} \arrow[dl] \arrow[d] \\
\{a\} \arrow[dr] & \{b\} \arrow[d] & \{c\} \arrow[dl] \\
 & \emptyset &
\end{tikzcd}
$$

(B.9)