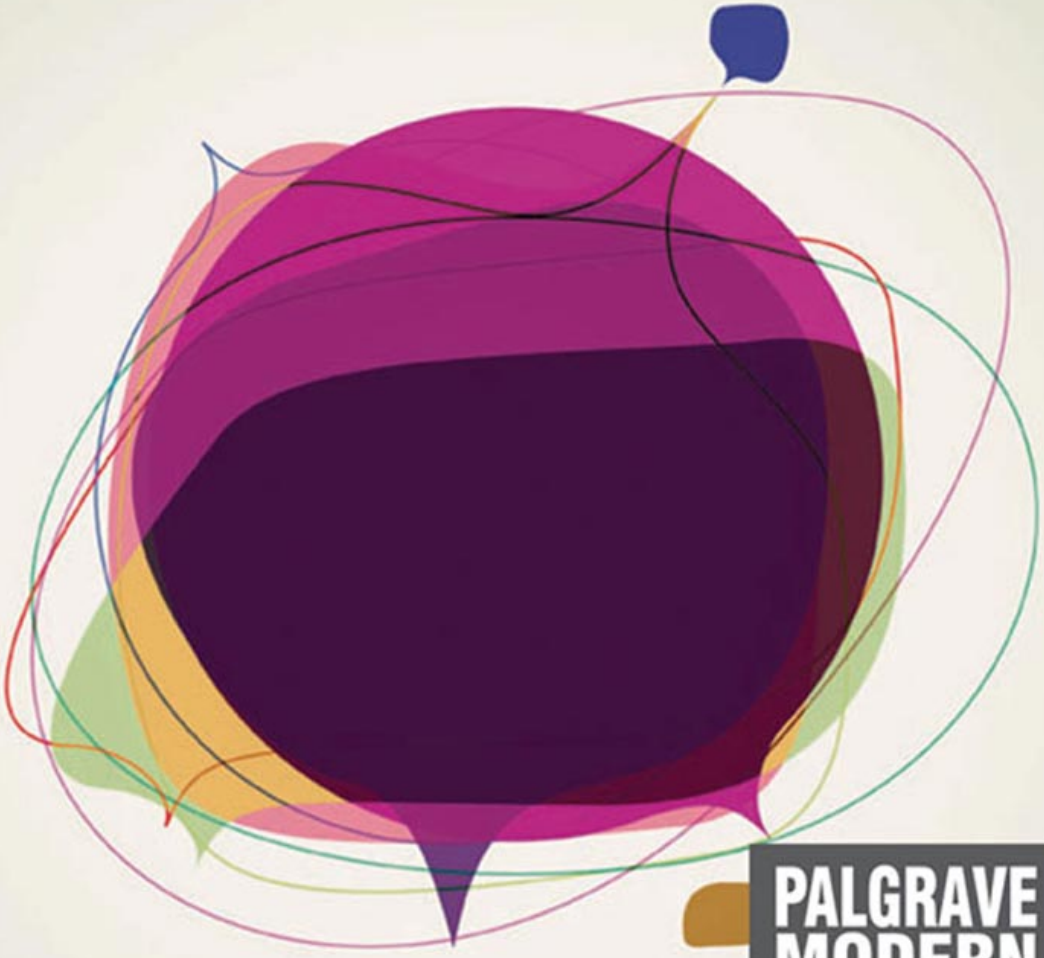


PALGRAVE MODERN LINGUISTICS

2ND EDITION

SYNTACTIC THEORY

GEOFF POOLE

SYNTACTIC THEORY

MODERN LINGUISTICS SERIES
Series Editor

Professor Maggie Tallerman
University of Newcastle

Each textbook in the Modern Linguistics series is designed to provide an introduction to a topic in contemporary linguistics and allied disciplines, presented in a manner that is accessible and attractive to readers with no previous experience of the topic. The texts are designed to engage the active participation of the reader, and include exercises and suggestions for further reading. As well as an understanding of the basic concepts and issues for each topic, readers will gain an up-to-date knowledge of current debates and questions in the field.

Titles published in the series

English Syntax and Argumentation (3rd Edition) Bas Aarts
Phonology Philip Carr
Pragmatics Siobhan Chapman
Linguistics and Second Language Acquisition Vivian Cook
Sociolinguistics: A Reader and Coursebook
Nikolas Coupland and Adam Jaworski
Morphology (2nd Edition) Francis Katamba and John Stonham
Semantics (2nd Edition) Kate Kearns
Syntactic Theory (2nd Edition) Geoffrey Poole
Contact Languages: Pidgins and Creoles Mark Sebba

Further titles in preparation

Modern Linguistics Series
Series Standing Order
ISBN 0–333–71701–5 hardback
ISBN 0–333–69344–2 paperback
(outside North America only)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in the case of difficulty, write to us at the address below with your name and address, the title of the series and one of the ISBNs quoted above.

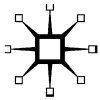
Customer Services Department, Palgrave Ltd
Houndmills, Basingstoke, Hampshire RG21 6XS, England, UK

Syntactic Theory

Second Edition

Geoffrey Poole

palgrave
macmillan



© Geoffrey Poole 2002, 2011

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6-10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted his right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

First edition 2002
Second edition 2011
Published by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN: 978–0–230–24393–4 hardback
ISBN: 978–0–230–24394–1 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
20 19 18 17 16 15 14 13 12 11

Printed and bound in Great Britain by
MPG Group, Bodmin and Kings Lynn

Still for my parents, and for Maria,
and now for Alma and Tiago too

Contents

List of Figures and Tables ix
Preface to the Second Edition x
Preface to the First Edition xii
Acknowledgements xiv

1 Introduction **1**
1.1 The scientific study of language 1
1.2 The competence/performance distinction 3
1.3 Some simple approaches to grammaticality judgements 7
1.4 Language acquisition and universal grammar 11
1.5 Description and explanation: the case of structure-dependent rules 14
1.6 Introduction to the book 18

2 Phrase Structure and Constituency **21**
2.1 A simple starting point 21
2.2 Constituency 27
2.3 The elements of a phrase-structure tree and their relations 31
2.4 Are two levels enough? 35
2.5 Constituency tests: more applications and case studies 46

3 X′-Theory and Functional Categories **61**
3.1 Introduction to X′-Theory 61
3.2 Inside the XP 62
3.3 Functional categories 66
3.4 Coordination and X′-Theory 79
3.5 Head-initial and head-final languages 81

4 θ-Theory and Case Theory **88**
4.1 An introduction to argument structure 89
4.2 θ-roles 90
4.3 Morphological case 97
4.4 Morphological case vs abstract case 99
4.5 Exceptional case marking 104
4.6 The Case Filter 108
4.7 The visibility condition: an improvement on the Case Filter? 110

5 Introduction to Binding Theory **119**
5.1 Anaphors and Principle A 120
5.2 Pronouns and Principle B 137
5.3 R-expressions and Principle C 141

6 Movement and Chains **150**
6.1 Transformations: an introduction 151
6.2 Wh-movement: subjacency and successive cyclicity 156
6.3 More evidence for successive cyclicity 165
6.4 Subjacency and the analysis of relative clauses 168
6.5 More transformations: passive and raising 172
6.6 Levels of application of principles 176
6.7 Chains 177

7 Logical Form **188**
7.1 Wh-movement in the LF component 190
7.2 Quantifier raising: another LF transformation 201
7.3 LF A-movement: expletive replacement 216

8 Empty Categories and their Constraints **227**
8.1 Binding Theory: a recap and (minor) reformulation 227
8.2 The Binding Theory status of DP-trace 228
8.3 The Binding Theory status of wh-trace 230
8.4 The search for null pronouns 232
8.5 [+a, +p] revealed! 235
8.6 Interlude: a note on Plato's Problem 239
8.7 The Empty Category Principle 240
8.8 From government to proper government 243
8.9 Subjacency vs the ECP 249
8.10 Adjuncts and antecedent government 250
8.11 Summary of the ECP 255
8.12 Relativized minimality: an introduction 256

9 Towards the Minimalist Program: Clause Structure Revisited **265**
9.1 The VP-internal subject hypothesis 266
9.2 More on the VP: double object verbs, VP-shells and light verbs 271
9.3 *v*P and Accusative Case assignment 280
9.4 From IP to TP 286
9.5 Rizzi (1997) and the 'cartographic' approach to CP 287

10 Towards the Minimalist Program: Movement Revisited **299**
10.1 The copy theory of movement 299
10.2 A movement analysis of control 307
10.3 Introduction to verb movement 319

Notes 332
References 336
Index 340

List of Figures and Tables

Figures

1.1 The object of study in linguistics 5
1.2 The relationship between languages and Universal Grammar 18
7.1 The computational system of the language faculty 218

Tables

2.1 Syntactic categories 22
8.1 DPs and their feature combinations 228

Preface to the Second Edition

Looking back on the Preface to the First Edition, written now nearly ten years ago, I suppose I should be pleased to see that I basically still agree with everything that I said there. (The alternative is that it's a consequence of total intellectual stagnation, so I'm sure that's not it!) For students, I suggest just flipping directly to the 'Preface to the First Edition' to get a sense of what the book is about and the way that it approaches the subject of syntactic theory. This preface, being the Preface to the Second Edition, is more concerned with the differences from and similarities to the First Edition and their rationale, and so is oriented more towards teachers and other specialists who will be familiar with the First Edition.

In general, I still believe very strongly that theory-building and problem-solving are the most useful general skills one can learn from the study of syntactic theory at the beginning to intermediate level. And I still believe, even now in 2011, that that is better accomplished with Government and Binding Theory (and its associated literature), as opposed to the Minimalist Program.

As such, the early chapters of the book haven't undergone any really radical changes. I've tried to add little bits and pieces throughout, new sections here and there, extra trees, etc., to improve the discussion: for example, more detailed discussion of: A′ and P′ proforms and some additional case studies in Chapter 2 (and a better 'split point' between Chapters 2 and 3); *for-to* infinitives in Chapter 4; and relative clauses in Chapter 6.

Where things have changed a bit more is in the last few chapters of the book, some of which have been almost completely rewritten. Chapter 8 is now, for example, an amalgam of Chapters 8 and 9 from the First Edition. The first part of the chapter is the first half of the previous Chapter 8, up to and including the interlude concerning Plato's Problem. This discussion is, in my opinion, extremely important, and is the point at which, in my experience, students really viscerally understand the force of Plato's Problem. (There are four different kinds of syntactic elements for which there is literally *no* evidence in the acoustic signal, because they're never pronounced, and each of those four elements has a different combination of Binding Theory properties.)

The rest of the new Chapter 8 concerns, in a very introductory way, the Empty Category Principle and Relativized Minimality. These topics, though reduced in scope from the First Edition, have been retained to provide a route into the considerable literature on these topics; literature which students can explore as part of further independent study projects. Also, for students

continuing on to study the Minimalist Program in a further, more advanced, syntax course, these issues inform (sometimes implicitly, sometimes explicitly) the development of ideas in the Minimalist Program. Certain particular paths are pursued, certain kinds of data are deemed to be particularly important, in part because of the way these topics played out in Government and Binding Theory. As such, I think the study of these topics is still very helpful for understanding aspects of the Minimalist Program.

Chapters 9 and 10 are new chapters which conceptually replace the First Edition's Chapter 10. Both are explicitly labelled 'Towards The Minimalist Program' and revisit clause structure and movement respectively, introducing more advanced ideas associated (more or less) with either late Government and Binding Theory or the Minimalist Program. (Obviously an important consideration was the extent to which the arguments can be straightforwardly adapted into the 'classic Government and Binding' model of the book where necessary.) Chapter 9, for example, considers not only the VP-internal subject hypothesis and VP shells, but also 'little vP' and the 'split CP' hypothesis. Chapter 10 introduces arguments for topics such as reconstruction/the copy theory of movement and the movement analysis of Control.

I've tried to retain the spirit of the previous edition in these new revisions. The tone of the chapters is intended to be informal but clear, and hopefully, as with the rest of the new edition, students both who continue on and who don't continue on with syntax will find something of value in them.

Geoffrey Poole

Newcastle University
(*née The University of Newcastle-upon-Tyne*)

Preface to the First Edition

This may sound like an odd thing to say about a book entitled *Syntactic Theory*, but the fact that this book is about syntactic theory is, in certain respects, incidental. I've written a book about syntax, and not quantum hydrodynamics, simply because I happen to know a little more about the former than the latter. (At least that's what my degree parchment says.) What this book is really about is theory-building: how to examine data, formulate and test hypotheses, and evaluate the results. We'll be building towards a generic version of Government and Binding Theory, which was first outlined in a 1981 book by Noam Chomsky called *Lectures on Government and Binding*. Government and Binding Theory is what I grew up with, and, crucially for my purposes, the theory is generally explicit enough and detailed enough to have deductive structure. (This is unlike a lot of more recent Chomskyian work in the so-called 'Minimalist Program'.) Because of this detail and structure, we can see clearly how the parts of the theory interrelate and test the predictions that it makes.

This is important because linguistics in general, and syntactic theory in particular, offer an opportunity that you don't get in a lot of academic subjects: to learn how to think really and honestly. Even in a lot of natural sciences, like physics or chemistry, you really only learn a little about that. At the initial/undergraduate level, it's mostly about solving ready-made problems. You memorize various tools and methods that are pretty well fixed in stone and the only real challenge involves figuring out which equation is the one you need to solve the problem. There's no real suspense or sense of discovery because you're looking at problems that have had known solutions for 2,000 years.

That's not the case in linguistics. It's still very much a process of feeling our way forward. Sure, some things are more secure than others, but there's still a lot of unexplored territory and lots of open questions. It's perfectly possible as an undergraduate to make a real discovery because you've had that flash of inspiration which nobody's had before, or because you got interested in looking at some aspect of a more exotic language that nobody's ever investigated. For that reason, I can't (and wouldn't want to) claim that the information and analyses in this book are The Truth, perfectly secure and immutable for all time. If that's what you're looking for, I'm sorry. I can't help you. (Do get in touch if you find it, however!) Even the most elegant or most sensible answers sometimes have problems, and I'll try not just to brush them under the rug or to trick you into not seeing them. Also, for people who are interested in doing this, they can be interesting topics for independent study projects under your teacher's supervision.

For those people who will be going on to do syntactic theory in more detail, I've tried to cover most of the 'big' issues that underlie a lot of work that's done and that get talked about a lot. That way, when you go to read a real article on Chomskyian syntax, you should at least have some basic idea what the article's about and know some of the background. However, even for those people who never think about syntax ever again, I hope you'll be a better thinker and problem-solver at the end of the book than you were at the beginning. That'll be a useful skill no matter what you end up doing.

Geoffrey Poole

University of Newcastle-upon-Tyne

Acknowledgements

I must first and foremost thank my editor and colleague Noel Burton-Roberts. I can't imagine what he must have thought when, less than six months after my arrival in Newcastle to take up my first job, I said to him 'I wanna to write a syntax textbook. How do I go about it?' To my amazement, he said 'Well, here's how to write a proposal.' And, as a result, here we are. His help over the last couple of years in banging this into shape has been absolutely invaluable and I can assure you with 100 per cent certainty that any errors or problems herein are a result of my not taking his advice about something. Chalk it up to a combination of piss and vinegar and the folly of youth.

I also must convey the now traditional thanks to Samuel D. Epstein, my old PhD supervisor, 'for trying to teach me to do linguistics right'. His training and supervision are reflected in just about every sentence. I also want to thank all my other teachers at Harvard and MIT, especially Höski Thráinsson, Jill Carrier and Mark Hale.

It's amazing how the linguistics literature doesn't ever provide exactly the example that you need in order to make a point clearly, and so thanks to John Stonham, Ikuyo Kawanishi and Fiona Lockett for help with data from Chinese and Japanese. And thanks also to all the students in my Spring 2001 class in Topics in the Structure of English for listening patiently to my early morning attempts to make this stuff sound plausible. Finally, a salty salute to absent friends (and former linguists), especially Erich, John, Scott and Steve.

For the Second Edition, I should first thank all my students over the years together with everyone, both students and teachers, who used the First Edition – particularly those who got in touch with me to discuss some aspect or other of it. I especially in this context would like to thank Claude Boisson, Ken Ramshøj Christensen, Sung-Hyuk Park, Jim Stanford and Wim van der Wurff (now, happily, a colleague here at Newcastle).

Thanks too to Maggie Tallerman, the current series editor (also happily a colleague), for much helpful discussion and suggestions concerning the Second Edition, and also to Kitty van Boxel and Felicity Noble at Palgrave Macmillan for editorial support (and patience!)

Of course, none of the above persons should be held responsible for any problems or shortcomings with this new edition, which, like the First Edition, I'm sure derive from my not taking somebody's advice about something. Whether now at my age I can successfully appeal to youth, piss and/or vinegar to explain it, I'm not quite sure, so I guess it'll have to be just plain cussedness and contrariness. Whether that's an improvement or not, I'll have to leave it to the reader to decide.

Introduction 1

Language has been an object of interest and investigation for thousands of years. The dialogues of Plato, which form the starting point for most of Western philosophy and science almost 2,500 years ago, contain speculations about language, particularly in the dialogue *The Cratylus*. One of the most interesting things about language is that, despite what you may have been led to believe by the popular literature, linguistic capacities of the sort that humans have don't seem to be possessed by other animals. Furthermore, every human being exposed to a language as an infant will grow into a competent speaker of a human language. In other words, language appears to be a species-specific property – something possessed by all human beings and only human beings. Our language plays a central role in human thought, society and action. Therefore, in studying language, we're exploring in a fundamental way what it means to be human.

1.1 The scientific study of language

The first question we need to ask ourselves is how we want to study language. The approach that we'll be taking in this book is most strongly influenced by the American linguist Noam Chomsky, one of the most important figures in twentieth-century linguistics. We'll see constant references to his work both in the broad outlines of the theory and in the details of particular analyses.

At the most general level, Chomsky suggests that we should investigate language by using the methods of the natural sciences – formulating hypotheses, testing those hypotheses against data, making our theory as simple as possible (as long as we make the right predictions), etc. That's not to say that this approach is guaranteed to succeed. It could turn out that using the methods of the natural sciences doesn't get us anywhere, in which case we'd need to try something else. But it seems right to say that this is where we should start.

On Chomsky's view, it's just a fact about the world that people speak and understand one another, on a par with any other fact about the natural world. I expel air from my lungs which sets air molecules in motion. Those molecules

hit your eardrum and that causes thoughts to arise in your mind which roughly correspond to the content of what I said. This is just another fact about the world, like the fact that things fall down and not up. We want to try and formulate theories which explain why the world is the way it is, and not some other way.

Implicit in this approach is also the assumption that there's a right answer and there's a wrong answer, what philosophers call a 'fact of the matter'. Human beings have some way of speaking and understanding their language. We're constructing a theory about that ability, and that theory is either right or wrong. It either correctly describes the way that human beings *actually* speak and understand their language, or it doesn't. For this reason, it's important for us to be as explicit and clear as possible in constructing our theory. That way, we can more easily check to see whether our theory makes the right predictions or not.

This isn't the only possible approach to language, though. For example, the linguist Zelig Harris (who was Chomsky's PhD supervisor) believed that the study of language was more akin to the study of literature. It was a question of illuminating a subject from different points of view, with no one view being 'right' or 'better' than another. However, I'm inclined to agree with Chomsky that we should begin our investigation of language using the methods of the natural sciences, and at least see how far we can get with it.

What are we studying?

So if we're agreed that we want to approach language using the methods of the natural sciences, where should we begin? First, our ability to speak and understand our language must reside somewhere in the body. But where? The brain seems like a good bet. Certain kinds of brain injuries, such as strokes, can impair one's ability to speak and understand one's language. Similarly, if electrical activity in the brain is temporarily disrupted, this can also create language impairments. (This is commonly done during brain operations to map out which exact areas of a person's brain are used for language, in order to ensure that the surgeon minimizes the removal of tissue from those areas.) By contrast, if we look at the foot or the heart, there doesn't seem to be any injury to those areas which can impair our linguistic abilities in the same way. Therefore, at its most basic level, in doing linguistics we're studying the human brain. Whatever our linguistic capacities are and however they're organized, they're part of the human brain.

Having established that we're studying the brain, and that we want to study the brain scientifically, you might wonder why we don't start talking in this book about various kinds of complicated brain-imaging scanners. Whenever you see a science documentary about the brain, the focus is almost exclusively on CAT scans, PET scans, MRI scans, EEGs and the like. Society seems to create the impression that, in order to do 'Science' (with a capital S), you need to be wearing a white lab coat and use lots of expensive equipment. However, as we mentioned, doing science is really just about having a certain attitude and approach towards what you're studying. As long as you're

constructing and testing theories with the goal of discovering something about the natural world, you're doing science.

In linguistics, it's not that we *couldn't* do brain scans. The problem is that not really enough is understood about the brain at that level for it to be a useful thing to do at this stage. At the low, 'nitty-gritty' level, we're not 100 per cent certain what's responsible for our ability to speak and understand a language. Many people argue that neurons are relevant, but there could also be a role to play for electrical configurations, chemical activity, some combination of these things or maybe even something else that we have yet to discover. Nobody is really sure what to look at. So, while complicated experiments with brain scans are part of the popular idea of 'studying the brain', there's no point in doing them if you don't really understand what the results might mean.

If we want to study the linguistic system of the brain scientifically, and we're not going to be doing brain scans, how should we approach it? What we have to do is treat the linguistic system of the brain like a 'black box'. We'll try to put things into one side of the black box and see what comes out the other side. From careful observation about what goes in and what comes out, we can begin to theorize about what has to be inside the black box in order to explain the results that we're getting. Eventually, we'll get to the stage where we can make predictions: if X goes into the black box, then Y should come out. If Y doesn't come out then our theory is wrong somewhere, and we'll need to change it. In fact, many philosophers see this process of 'falsification' as the litmus test for science. If you're being explicit enough to know that what you're saying is false, and then you try to correct your theory, you're doing science.

1.2 The competence/performance distinction

In this book, we'll be focusing on a particular aspect of the linguistic system, namely syntax. That is, we'll be focusing on sentence structure and 'grammar'. What sort of experiments should we run in order to investigate our grammatical capacities? How do we investigate this grammatical black box? As Chomsky has noted on many occasions, there seems to be a distinction between what people *know* and what they *do*. Put another way, what we actually say during the course of ordinary conversation isn't always an accurate reflection of our linguistic capacities. We can make mistakes, such as slips of the tongue, for example. Chomsky refers to this distinction as the *competence/performance* distinction. Therefore, we don't want to focus our attention on what comes out of our mouths. We want to focus on the underlying system.

Studying behaviour

In the early twentieth century, though, it was common for linguists to study 'linguistic behaviour'. They would record as many speakers and as much natural conversation as they could, and then analyse these recordings and construct a theory based on what was found. This approach would seem to be perfectly in keeping with our goal of investigating language using the methods of the

natural sciences. If we were studying, say, dominance hierarchies in ape groups, we would probably spend a lot of time observing and recording their behaviour in different situations and then constructing a theory from that data.

However, there a number of reasons why that approach isn't the best way to investigate the syntactic capacities of human beings. When you're actually producing speech in real time, there are all sorts of factors which come into play – how tired you are, whether you're paying attention, etc. As a result, *actual* speech is full of 'mistakes', like false starts, slips of the tongue and times when you start a sentence and then don't know how to finish it. I have myself on occasion produced sentences like (1) in conversation:

(1) This is the film that you said that was good.

Now despite having produced this sentence in conversation, I would admit, if it were pointed out to me, that the sentence in (1) is *not* part of my language. Intuitively, I know there's something wrong with (1), and, more specifically, the problem occurs in the phrase 'said that was good'. In order for the sentence to be well-formed, the *that* must be deleted, creating (2). (This is true for most dialects of British and American English, although not all.)

(2) This is the film that you said ( ) was good.

Thus, there is a competence/performance distinction. What I *do* with my language doesn't always accurately reflect what I *know* about my language because of the additional burdens involved in actually producing speech in real time.

This means that we could run into problems if we were to simply record a large amount of natural speech. It would presumably contain 'errors' of the sort we've just seen. But we would have no principled way of deciding what was an error and what wasn't without reference to the speaker's *knowledge* of what is an error and what isn't. In real life, you know whether a person has been interrupted, or broken off a sentence and started a new one, but that's because you're filtering the data through your own knowledge of the language system.

Another problem is that there are sentences which are perfectly grammatical, but which come up rarely, if at all, in ordinary conversation. In illustrating this point, Gregg (1989) discusses the following sentence:

(3) Mary ate an apple, and Sue a pear.

He says of this sentence (p. 22): 'I do not produce [sentences of this type] because they sound incredibly affected, and I have enough problems of that sort already'. Nonetheless, if you're a native speaker of English, you know that (3) forms a part of your language. We would want our theory to explain why English speakers accept (3). However, it's not clear that we'd be able to do that if we just go out into the world and record people, since it might not come up very often in our sample. In fact, from a purely statistical point of view, sentences like this probably occur even less frequently than 'errors'.

Related to some of the points we've been making above, you also know that certain strings of words do *not* form a part of your language. When we were discussing 'errors' just a moment ago, I mentioned that (1) was ungrammatical in my dialect of English. However, this aspect of your linguistic knowledge isn't something that you'd discover if you were to just record people. People make mistakes, but that's different from getting systematic evidence that certain strings of words are ungrammatical.

For all of these reasons, it seems as though simply recording speakers has problems as a method for discovering how the black box that constitutes our linguistic capacities works.

An alternative to studying behaviour: studying knowledge

Recall above the distinction between *competence* and *performance* which Chomsky draws our attention to. If studying performance (linguistic behaviour) has the problems that we've seen, perhaps there's a way that we can focus on competence (linguistic knowledge) directly; that is to say, focusing on the cognitive system in the brain that *underlies* our linguistic behaviour, rather than looking at the behaviour itself (see Figure 1.1).

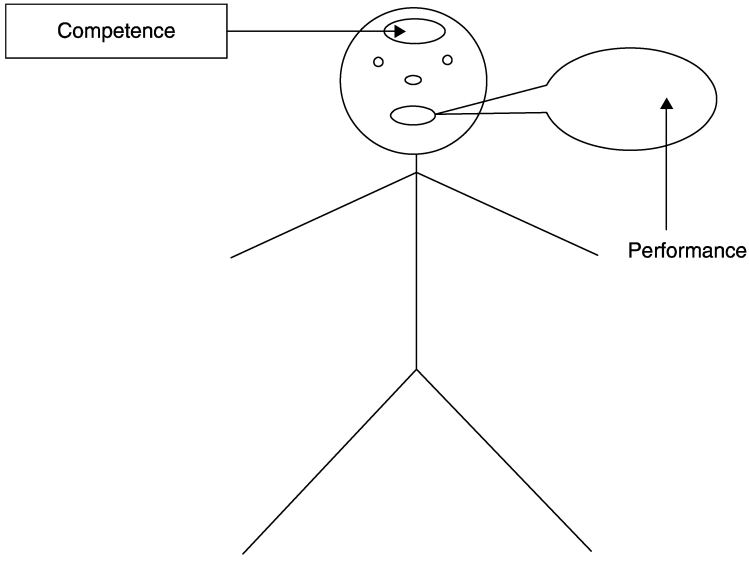


Figure 1.1 The object of study in linguistics

To do this, the standard experimental task is to solicit judgements from native speakers about the sentences of their language. Part of what you know about your language is that certain strings of words are part of that language and certain other strings are not. For example, the following string of words forms part of my language, a fairly standard dialect of American English:

(4) I haven't read the book yet, but I will tomorrow.

However, the words in a different order create a sentence which is not a part of my language:

(5) *I yet the read book haven't, but will tomorrow I.

For a native speaker of many dialects of British English, on the other hand, neither (4) nor (5) forms a part of their language. They are both ungrammatical. The grammatical version of (4) in their dialects is (6), which to me sounds very strange.

(6) I haven't read the book yet, but I will do tomorrow.

By feeding test sentences to native speakers of a language, and asking for their intuitive judgements/reactions, we can begin to understand what's inside the black box. We can begin to understand how their linguistic system is set up.

Consulting native-speaker intuitions in this way would seem to avoid, or at least minimize, the problems that we saw above when we considered simply recording speakers. There is still the problem of error due to interfering factors (which we'll discuss below). However, the burdens associated with what you're asking them to do are at least reduced. If you just have to provide your intuitive judgements, that's easier than having a 'real' conversation, with all of the distractions that that entails. You also are able to test both less frequently occurring structures and structures which you expect to be ungrammatical. Although people may not use sentences like (3) in ordinary conversation, you can still get judgements regarding them.

Grammaticality vs acceptability

As we mentioned just a moment ago, the grammaticality judgement task, which focuses on linguistic competence rather than linguistic performance, would seem to minimize errors due to interfering factors. However, it's important to see that, even when you're using grammaticality judgements, you must still consider whether the data is 'reliable'. There are things which can influence native-speaker intuitions which are not due to the linguistic system proper, but which are intimately related, including affectedness (which we mentioned above in discussing example (3)) and stylistics. For this reason, in addition to a competence/performance distinction, we also have a distinction between grammaticality and acceptability.

For example, sentences which themselves have sentences as their subjects can often be somewhat awkward stylistically.

(7) ?Once [that John was leaving] became clear, we abandoned the project.

In (7), the underlined adverbial clause contains a sentence *that John was leaving* which is itself in the subject position of the sentence *that John was leaving became clear*. This can result in a sentence which doesn't flow and which is

difficult to understand. To make the sentence easier to understand, the sentential subject is often put at the end and a dummy pronoun *it* placed in subject position:

(8) Once *it* became clear *that John was leaving*, we abandoned the project.

Unlike the somewhat awkward (7), (8) is perfect. However, despite the fact that (7) and (8) differ in *acceptability*, they are both *grammatical*, in the sense that our grammar (theory) should generate them. Our judgements in (7) are being influenced by extra-grammatical factors. In this case, stylistics and difficulty of processing.

This point is also illustrated by a story that Professor Jill Carrier, one of my old teachers, related to me about some fieldwork that she was doing with a native speaker of Tagalog, a language of the Philippines. She had a list of sentences that she was checking with the informant, and they were all confirmed as grammatical in Tagalog as expected. However, there was one particular sentence where the informant said 'Oh, no. Nobody would ever say that. That's impossible'. This was deeply puzzling, as from a structural point of view it seemed to be identical to various other sentences which were all reported to be perfectly fine. After much thought and further questioning, she discovered that the problem was that the example sentence happened to involve children swimming in the ocean. The reason that the informant had objected to the sentence was because the ocean was full of sharks and no children would ever be swimming in the ocean. It was too dangerous. Clearly, in that instance, it wasn't that the informant's *grammar* didn't generate the sentence in question. Instead, non-linguistic factors – knowledge of the real world – were influencing his judgement. What we are trying to do is isolate purely linguistic capacities.

Thus, whenever we're discussing grammaticality judgements or consulting with informants, we need to make sure that we keep this distinction between acceptability and grammaticality in mind. There may be plenty of sentences that our mental grammar generates. There are some, like (3), which are said quite infrequently, but there are also sentences which we would *never* say for one reason or another. If we want to correctly characterize how the black box of our linguistic capacities works, we want to be sure that our theory captures *all* of the sentences that it allows.

1.3 Some simple approaches to grammaticality judgements

In the previous sections, we've argued that we should focus in on a particular aspect of our linguistic knowledge: grammaticality judgements. You know instantly and without prior instruction that certain sentences of your language are grammatical and that certain other ones are ungrammatical. But how do you know what you know? Where do these grammaticality judgements come from?

To start with, let's take some first guesses at what might be going on. Just to be methodical, we'll start with some simple possibilities and work through to more complicated ones. Hopefully, you'll be convinced that these various simplified accounts aren't plausible.

The 'parrot' theory

The first, most naïve account of how you come to have grammaticality judgements would be that you judge a sentence grammatical if you've heard it before, and judge it ungrammatical if you haven't heard it before. We could call this the 'parrot theory'. This is pretty obviously wrong, I think. One of the most striking features of language is its unbounded capacity for innovation. It's quite likely that a large number of the sentences in this paragraph, not to mention the book so far, will be sentences that you've personally never come across before (unless you've heard me lecture). However, none of these sentences strikes you as strange or unfamiliar. I can even make up a sentence which I'm 100 per cent sure is new in the history of the English language (at least at the time of writing):

(9) Although Huxley's *The Doors of Perception* was never intended to be taken internally, there were many people in the late 1960s who failed to heed the warnings of prominent scholars and critics.

I'd bet a lot of money that (9) has never been produced ever in the entire history of the English language, yet you see it as a well-formed sentence of English without any effort whatsoever. I take that as fairly conclusive proof that your grammaticality judgements aren't based on simply whether you've heard the sentence before. Otherwise, you ought to judge (9) as ungrammatical, and you don't.

Other seemingly plausible suggestions

It also appears as though various other seemingly simple and plausible ideas don't play a role in our grammaticality judgements. For example, judgements about grammaticality are totally independent of what, if anything, the sentence means. Here is a famous example sentence from Chomsky (1957):

(10) Colorless green ideas sleep furiously.

Personally, I don't have the first clue what this sentence means. Ideas don't normally have colours, but even if they could, they certainly couldn't be colourless and green simultaneously. I'm also hard pressed to understand what sleeping furiously might entail. However, one thing I do know is that (10) is different from (11), where the words are in the reverse order:

(11) *Furiously sleep ideas green colorless.

The difference is that (10), despite its bizarre non-meaning, is syntactically well-formed. (11), on the other hand, is not syntactically well-formed.

The reverse situation can be found as well – that is, sentences which *do* make sense, but which we instantly judge to be ungrammatical. Consider (12):

(12) *The child seems sleeping.

Unlike (10) or (11), we have no problem assigning a meaning to the sentence in (12). If a foreigner had produced it, we would have no difficulty understanding what he or she intended to convey. However, we all know instantly that (12) is not a grammatical sentence of English, even though its meaning is clear.

In a similar way, ease of use doesn't seem to have any implications for grammaticality. There are plenty of sentences which are grammatically well-formed but which are not easy to use for one reason or another. Consider tongue-twisters, for example:

(13) She sells seashells by the seashore.

The fact that they're difficult to use (difficult to pronounce, in this case) is precisely why they're of interest. But crucially, they're not ungrammatical.

Sentences can also be difficult to use because they're difficult to parse. So-called 'garden path' sentences are ones in which your parser is 'tricked' into assuming that a sentence has a structure which needs to be revised when you get to the end of the sentence. A famous example is:

(14) The horse raced past the barn fell.

The effect when reading a garden-path sentence is usually lessened, but they can be quite startling when you hear one spoken to you. That's because when you're parsing the spoken sentence in real time, your parser is usually fooled into thinking that the past participle *raced* is in fact the main verb of the sentence. You don't realize you've made a mistake until you hear *fell*, and you're suddenly left with a word that you don't know what to do with. You've then frantically got to go back and figure out where you went wrong, which is harder to do when the sentence is spoken as opposed to read. You then realize that *raced* is supposed to be a past participle rather than the main verb. It's the horse *which was raced past the barn*. In spite of all this though, when the intended reading is pointed out to you, you realize that the sentence is in fact grammatical.

Analogy

A slightly more sophisticated version of the 'heard it before' theory is less obviously stupid. It could be that you judge a sentence as grammatical if it's 'similar to' or 'analogous to' a sentence that you've heard before or that you know is grammatical. Call this the 'analogy' theory. The essential problem is that, as

an explanation of your linguistic abilities, the whole concept just looks wrong. Consider the following pair of sentences:

(15) a. It is likely that John is here.
b. It is probable that John is here.

Likely and *probable* mean almost exactly the same thing, and so it's perhaps not surprising that they can both appear in the 'it is X that John is here' frame, as we see in (15). However, consider (16):

(16) a. John is likely to be here.
b. *John is probable to be here.

If analogy or similarity were the main factor in explaining why we have the grammaticality judgements that we have, it's not clear at all how or why (16b) is ungrammatical. *Likely* and *probable* are synonyms, they behave identically in (15), yet you know instantly and without prior instruction that (15b) is grammatical and (16b) is ungrammatical. Why does analogy fail here?

An even more telling example, I think, can be seen in the sentences that we were looking at earlier in the context of the competence/performance distinction above, having to do with the presence or absence of *that*. In virtually all situations, *that* can be either present or omitted with no change in grammaticality:

(17) a. I know *that* Mary is here.
b. I know Mary is here.

This optionality seems to remain when there is a question, rather than a statement:

(18) a. Who knows that Mary is here?
b. Who knows Mary is here?

(19) a. Who do you think that Mary saw?
b. Who do you think Mary saw?

However, there is one exception to this otherwise entirely regular pattern. When the *subject* of the embedded clause is questioned, the presence of the complementizer *that* is no longer optional. It is obligatorily absent. (Again, this is true for *almost* all dialects of British and American English.)

(20) a. *Who did you say that left the room?
b. Who did you say left the room?

Immediately, and without any prior instruction, you know that there's something wrong with (20a), even if you can't put your finger on what it is.

This is totally unexpected if something like analogy or similarity is relevant for understanding our intuitions about grammaticality. What we have is a paradigm which is absolutely and utterly regular, except for the one exception in (20a). If analogy were relevant, then (20a) should be grammatical, *on analogy* with (17a), (18a) and (19a). This really makes it look as though analogy or similarity isn't relevant for grammaticality.

Conclusion

What seems to be going on is that our intuitions about grammaticality stem from the fact that the brain contains a system for analysing sentences. When presented with a sentence of English, it's analysed by the cognitive system that you possess, providing you with a judgement about its acceptability. You have internalized a 'grammar' of your language.

What we're really doing in syntax, then, is something that you won't find in any other subject. In exploring this mental computational system and how it works, what we're doing is, in a sense, nothing more than telling you things that you already know. We are attempting to develop a theory about what's already in your head. However, that doesn't mean that you can put down the book and run away. Although you do, in an important sense, already know everything that we'll be talking about, the problem is that you don't *know* that you know it. It's knowledge, but unconscious knowledge. Syntactic theory is about making this knowledge explicit.

1.4 Language acquisition and universal grammar

We've come to a pretty radical conclusion. Part of your linguistic abilities consists of having judgements about whether sentences are grammatical in your language or not. We concluded that this was because your brain contains a cognitive system which analyses (or 'generates') sentences. In this section, we're going to ask a simple question, which turns out to have a complicated answer: how did this cognitive system get there? How did you acquire this ability to judge the grammaticality of sentences? The answer to this acquisition question will turn out to have a huge influence on the way we conduct our inquiry into syntax.

With respect to this issue, the linguist Noam Chomsky has observed what he calls 'the logical problem of language acquisition'. In a nutshell, what adults come to know about their language goes far beyond anything they were actually exposed to as children. Examples (15)–(20) are prime examples. *Likely* and *probable* are about as close synonyms as you can get. Yet you know instantly that *John is likely to be here* is grammatical but *John is probable to be here* isn't. However possible it might be that *one particular* speaker's experience enabled him or her to figure these facts out, it can't possibly be that *everyone's* experience was enough. Yet this is something that *every* speaker of English knows.

Here's another example, which we'll be looking at in the next chapter. (21) is a perfectly well-formed sentence of English:

(21) Jack met the student from England and I met the one from France.

The element *one* here seems to be filling in for *student*. (21) means *Jack met the student from England and I met the student from France*. However, if we change the prepositional phrase, things change radically:

(22) *Jack met the student of physics and I met the one of chemistry.

For some reason, although *one* can substitute for *student* in (21), it can't in (22).

So the question is: how did you come to know these facts? It certainly couldn't have come from any explicit instruction by teachers. It's unlikely that any of them would have noticed these facts. It's equally unlikely that the data you were exposed to when you were acquiring your native language was itself detailed enough for you to be able to figure out these facts. And, even more to the point, these facts are again something that *every* speaker of English knows. No matter how likely it is that you somehow managed to get exactly the right data, it's unlikely that *every* speaker of English got exactly the right data. Our experiences as children were all radically different. It looks like what we have is a case of knowledge being manufactured from nothing, not just in you, but in every speaker of English. How is this possible? This, as Chomsky sees it, is the logical problem of language acquisition.

Chomsky has a logical yet radical suggestion in order to solve this problem: if we have knowledge that couldn't have come from the data we were exposed to, it must have come from somewhere else. Chomsky proposes that this 'somewhere else' is an inborn language acquisition device specific to the human species. (This is what we mean by 'species-specific'.) Just as it is genetically predetermined that a human being will grow arms and not wings, it is genetically predetermined that a human being will acquire a human language. Children come pre-equipped to the task of language acquisition with a device that is designed to take certain kinds of acoustic disturbances ('sounds') as its input, and as its output creates a cognitive system for understanding and using language. If we assume that being exposed to a language acts as a trigger for the unfolding of a genetic program, on a par with the development of limbs and organs, this would seem to solve the questions about language that we've just raised. You end up with a linguistic system that goes well beyond the particular input data you were exposed to, in precisely the same way that organ development 'goes beyond' the particular 'nutritional inputs' (i.e. food) that you were exposed to as a child.

Chomsky's hypothesis is that this 'language acquisition device', more commonly known as Universal Grammar (UG), has two central parts to it which aid in the task of language acquisition: *principles* and *parameters*. In fact, the general approach to syntax that we'll be arguing for in this book gets

its name from these parts of UG. It's called the 'Principles and Parameters' approach.

An introduction to principles and parameters

The idea is that the *principles* of Universal Grammar account for the respects in which all human languages are the same. Most of the topics that we'll be discussing in this book concern principles of UG, and we'll be drawing data from English and many other languages in order to argue that principles of UG are at work.

Parameters, on the other hand, are intended to capture certain ways in which languages differ. Now, obviously, one way that languages differ is that they have different words for different concepts. So dog is *dog* in English, *hund* in German, *perro* in Spanish, etc. These are differences that simply have to be memorized by the child, and it's only when the child is exposed to the language that he or she will know what the words are going to be.

However, there are other ways in which languages differ. For example, the languages of the world seem to divide up into *null subject languages* and *non-null subject languages*. English is a non-null subject language. That is to say, English sentences must *always* have a subject, even when that subject appears to contribute nothing to the meaning of the sentence. Consider (23):

(23) a. It is raining.
b. It seems that John is here.

The subject of the sentences in (23) is *it*, but this pronoun is totally meaningless. You're not pointing to something when you say (23a) or (23b). The pronoun doesn't refer to anything. However, because English requires a pronounced subject at all times, these 'dummy' pronouns must be inserted. By contrast, in a null subject language like Spanish, these dummy pronouns are omitted. The subject is simply not pronounced.

(24) a. Llueve.
rains
'It's raining.'
b. Parece que Juan está aquí.
seems that Juan is here
'It seems that Juan is here.'

The hypothesis is that UG contains a parameter called the Null Subject Parameter. Therefore, in some sense, the child already knows from birth that the language that he or she will be exposed to will either be a null subject language or a non-null subject language. Once the child begins to hear data, it is a relatively simple matter to set the parameter one way or the other. A parameter is like a question. Setting the parameter is a matter of answering that question in light of the evidence from the particular language.

1.5 Description and explanation: the case of structure-dependent rules

One important benefit that our postulation of an inborn Universal Grammar also gives us is the ability to go beyond merely describing individual languages. It puts us in a position to *explain* why particular languages are the way they are. For example, one of the things which seems to characterize all languages that have been studied so far is that they form yes/no questions (and, in fact, all questions) in a *structure-dependent* fashion.

By way of illustrating the structure-dependence of yes/no questions in English, consider the following examples:

(25) a. The agent has left the building.
b. Has the agent left the building?

In (25a) we have a declarative statement. In (25b), we have the corresponding yes/no question; that is to say, a question which asks for a yes or no answer. Imagine you're a Martian scientist trying to discover what the rule is for English yes/no question formation. (Since you're a Martian, you're not familiar with English or any other human language.) On the basis of the pair in (25), a simple rule that looks like it accounts for the data is (26):

(26) Move the third word to the front of the sentence.

This rule not only accounts for the pair in (25), but also an infinite number of other sentences, including pairs like those in (27) and (28):

(27) a. The man is here.
b. Is the man here?

(28) a. My friend will go to the store.
b. Will my friend go to the store?

The rule in (26) does not require analysing the sentence in terms of its structure. You just go along the words one by one until you hit the third word, then you move that word to the front of the sentence. It is therefore a *structure-independent* rule.

Interestingly, although the structure-independent rule in (26) is very simple to apply, you quickly discover that it isn't right. Consider the (declarative) sentence in (29):

(29) The man from the bank will remember nothing.

If we were to apply the structure-independent rule in (26), we would scan along from left to right until we encountered the third word, which in this case is *from*. We move *from* to the front of the sentence, and we get (30):

(30) *From the man the bank will remember nothing?

This is clearly not a grammatical sentence of English. If we want to make the declarative sentence in (29) into a question which asks for a yes or no answer, we need to move *will* to the front of the sentence:

(31) Will the man from the bank remember nothing?

So clearly our rule for yes/no question formation in (26) isn't right. But what is?

One thing to notice about (30) is that our rule in (26), which just counted words from left to right, ended up moving a preposition (*from*) to the front of the sentence. However, we know that in order to form a yes/no question, what we need to do is move an *auxiliary verb* to the front of the sentence. (Let's put aside the cases where we have no auxiliary, in which case you'd have to insert the 'dummy' auxiliary *do*.)

(32) a. *Should* John have gone to the store?
b. *Will* the book arrive on time?
c. *Has* Mary defeated her opponents?

It's possible that we can save a structure-independent approach by modifying our rule in (26) to something like (33):

(33) Move the first auxiliary verb to the front of the sentence.

That rule *does* correctly get us (31), and all of the other cases that we've seen so far, but it's still structure-independent. It doesn't require that we know anything about the *structure* of the sentence. We just scan from left to right until we hit the first auxiliary verb, and then move it to the front of the sentence.

However, (33) isn't the English rule either. Consider (34):

(34) The man who has written the book will be followed.

Using that rule, we'd predict that the yes/no question corresponding to (34) should be (35):

(35) *Has the man who written the book will be followed?

If this isn't the right rule either, then what is it?

Since you're a speaker of English, a human language, and not really a Martian scientist, it'll probably seem obvious what the 'right' answer is. But it's important to see how unexpected that answer is. The 'right' answer is something like (36):

(36) Move the first auxiliary verb *in the predicate* to the front of the sentence.

You don't just move along and pick the third word or the first auxiliary in a *structure-independent* way. The rule that you've internalized requires treating the sentence not just as a string of words coming one after the other. Instead, you treat the sentence as if it has structure. The English rule for yes/no question formation is structure-dependent. In this case, at the very least you have to divide the sentence up into a subject and a predicate:

(37) a. Subject: The man who has seen the book
b. Predicate: will be followed

Having identified the predicate, you can then find the first auxiliary in the predicate and move it to the front of the sentence. This rule will also account for the other sentences that we've seen:

(38) a. Subject: The agent
b. Predicate: has left the building

(39) a. Subject: The man
b. Predicate: is here

(40) a. Subject: My friend
b. Predicate: will go to the store

(41) a. Subject: The man from the bank
b. Predicate: will remember nothing

So it seems that even a simple thing like turning a declarative sentence into a yes/no question requires an analysis that is, from a computational point of view, somewhat complicated. You don't just scan across and pick the first auxiliary or the tenth word or something. You need to analyse the sentence into a subject and a predicate in order to find the main verb phrase.

Now the reason for this long discussion about structure-dependent vs structure-independent rules is to make a point about UG. If we look at all the languages of the world that have so far been studied, the number of languages which use structure-independent rules is exactly none. There seems to be no language in the world which uses structure-independent rules for anything. No language, for example, forms questions by moving the third word to the front of the sentence or by taking the words in the declarative and reading them in reverse order.

As we mentioned above, if we were just looking at language after language in a descriptive way, we'd have no explanation for this fact. It would just be a coincidence. We'd look at English and say, 'Wow, no structure-independent rules'. We'd then move on to Swahili and say, 'Wow, no structure-independent rules'. And so on and so forth. On the other hand, if we as humans possess a genetically determined UG, then we can explain why we find what we find. There's something about UG which doesn't allow for structure-independent rules.

We can even go further than this, however. If we're right to say that UG doesn't allow for structure-independent rules, and that's what accounts for the fact that no language has them, then we make a further prediction that we can check. Assuming that the language acquisition process consists of language data interacting with UG, then we should predict that no child ever hypothesizes a structure-independent rule during the course of language acquisition. If UG is always there in the background, helping the child process the input data, then the child should never use structure-independent rules at any stage. UG doesn't allow them.

Interestingly, this prediction also seems to be verified. No child is ever exposed to sentence pairs like (27) and (28) above and hypothesizes a computationally simple structure-independent rule like (26) or (33). They always and automatically make a structure-dependent choice. It might be the wrong structure-dependent choice, but it's still a structure-dependent choice. Thus, it would seem that they never hypothesize a structure-independent rule for anything because it's never an option. Children come pre-equipped to the task of language learning knowing that all grammatical rules have to be structure-dependent.

Other sources of evidence

There's some interesting corroborating evidence as well from studies of an autistic patient, who's called Christopher in the literature. Christopher has some fairly severe cognitive deficits and has to be looked after in an institution, but, like some autistic patients, he also has a talent. That is, there is one area where he seems to perform much better than ordinary people. Christopher's particular talent is languages. He enjoys learning new ones and does so effortlessly and at a tremendously rapid pace. He can translate back and forth between them perfectly, getting all the grammatical endings right, etc. If UG is implicated in Christopher's talent, then we would expect to see that reflected in the way he acquires his languages.

Structure-dependency is actually one of the things that Christopher has been tested on, and the results seem to support the existence of UG. Christopher and a control group of normal people were presented with two languages to learn. One was Berber, a North African language, which has incredibly complex rules and constructions. The other was a made-up language which crucially contained *structure-independent* rules. Although Christopher mastered Berber without difficulty just as he had other languages, he struggled with the made-up language, and was never able to master it. The control group struggled with Berber, and also struggled initially with the made-up language, although they eventually did better than Christopher on the made-up language.

This would seem to provide strong corroborating evidence for the claim that UG exists and that it prohibits structure-independent rules. By hypothesis, Christopher's considerable linguistic resources come from UG. His general problem-solving abilities are not very good. By contrast, the control group may not have the access to UG that Christopher has (being adults), but they do have good general problem-solving skills. This would explain why the control group had difficulties with Berber which Christopher did not, and why they had more

success with the made-up language. Since the made-up language contained rules which were prohibited by UG, the control group could switch gears and treat it as though it were a puzzle, calling on their general problem-solving abilities. Since Christopher has very little resources in this area, he was unable to succeed with it.

Description and explanation

So, ultimately, while we're interested in individual languages, we're also interested in discovering what we can about UG, the initial state of the language faculty in the mind/brain. As a result, we have a constant flow of ideas back and forth between individual languages and UG. In order to account for some particular fact about English, we might hypothesize that UG is implicated. But a hypothesis about the make-up of UG makes a prediction about other languages. We therefore then need to check those other languages to see if the prediction is verified. This process is illustrated in Figure 1.2.

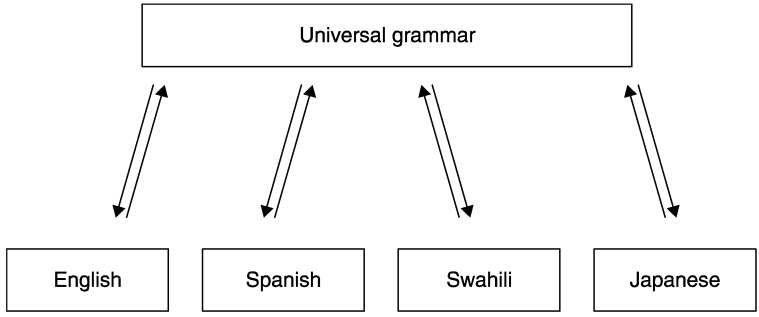


Figure 1.2 The relationship between languages and Universal Grammar

In this way, we can make a distinction between two kinds of arguments: *conceptual* and *empirical*. Roughly speaking, when we're making an argument that goes *from* UG *to* a particular language, we're dealing with a conceptual argument. Our theory about how UG works says X, and so therefore we expect that we'll find X when we look at individual languages. When we're going *from* individual languages *to* UG, then we're making an *empirical* argument. We've got some fact X about a language, and so we need to work that fact X into our theory about UG.

We'll try as much as possible to make arguments of both types. Rather like a pair of scissors, we'll work best when we come from both directions at once. In that way, we can be as secure about our conclusions as possible.

1.6 Introduction to the book

Now that we've discussed some general things about the study of language, I want to turn to a specific introduction to the subsequent chapters. Each chapter covers a specific topic, from sentence structure to transformations and beyond.

Several of the later chapters, however, return at a more advanced level to issues raised in earlier chapters. This way, you can get an introduction to a topic in an earlier chapter, and then come back in a later chapter to more recent proposals. If you work all the way through to the end of the book, you should be well prepared, with a little bit of coaching, to take the next step – reading actual articles in the linguistics literature.

In the individual chapters, one of the things you'll find is that there are in-text exercises scattered throughout. There are two kinds of these. The first kind is labelled 'Exercise'. These are just like the exercises at the end of the chapter, but placed right after the relevant point of the discussion in the chapter. These are there to give you a chance to practise using the theory while the relevant concepts are fresh in your mind; answers to these can be found at the end of the chapter. The second kind of exercise is labelled 'Before Continuing'. These exercises will usually come at a point where we've encountered a problem with the theory. The reader is encouraged to stop for a minute before reading on and give some thought to how to solve the problem or address the question. The immediately following discussion will then provide an answer to the question. Just like anything else, the easiest way to get comfortable with the theory is to practise it. Pretty soon the whole thing will be second nature. As I mentioned, the end of each chapter also contains 'Further Exercises' for you to practise the things that you've learned.

Each chapter also has a 'Further Reading' section. In many cases, the references will be to articles that we've actually discussed in the chapter. Sometimes in the text I will have omitted complexities that aren't relevant to the argument at hand, or I have translated an argument from a slightly different framework into the one we're using. Where this occurs I will mention it in the 'Further Reading'. One of the things that's slightly different about this book is that each chapter will conclude with a section entitled 'Open Issue'. These sections are designed to take something from the chapter a step further or in a controversial direction. I thought it was important to include things like this to show you that we're not dealing with a theory that's fixed in stone, with all problems solved and nothing new to think about. They're meant to be starting points for your own thinking and to provide suggestions for research projects. You might try to find evidence from a different language for an Open Issue proposal or perhaps look for counter-examples and discuss whether or not they represent real problems.

Summary

In this chapter, I've tried to sketch out some of the basic assumptions which underlie the more technical parts of the rest of the book. As a native speaker of a language, one aspect of your linguistic abilities is the ability to judge whether a given string of words forms a part of the language you speak or not. In other words, you have judgements about grammaticality and ungrammaticality. This ability stems from the grammar of your language which you unconsciously 'constructed', based on the language acquisition data that you were exposed to

in childhood. However, what your grammar contains seems to go well beyond the data you were exposed to. Additionally, speakers of a language all converge more or less on the same grammar despite the wide variation in particular data they happened to have been exposed to. For these reasons, it's hypothesized that there is a language acquisition device, or Universal Grammar, which is part of our common genetic endowment as humans. In trying to develop a theory about 'English', by which we mean the internalized grammar of a speaker of English, we also need to construct a theory of Universal Grammar.

FURTHER READING

Probably the most accessible work by Chomsky himself on these general topics is Chomsky (1988). It's a transcription of a series of lectures given at the University of Managua for a general audience; that is, for people who didn't know anything particularly about linguistics. One caveat, though: since the audience were native Spanish speakers, the examples he picks, though not complex, are taken from Spanish rather than English. There are translations of everything, though, and it's actually good practice for getting used to non-English examples. Most standard Chomskyian linguistics textbooks don't devote much discussion to general, 'nature of linguistics' questions, but a notable exception is Haegeman (1994). Chapter 1 of that book has some very good discussions of general issues and the setting out of the Chomskyian linguistic programme. For people interested in reading more about the autistic patient, Christopher, and what he can tell us about the Chomskyian approach to language, you can check out Smith and Tsimpli (1995).

Phrase Structure and Constituency

2

Now that you have some idea what our general approach to 'language' is going to be and seen some reasons why studying language and languages in the manner of the natural sciences might be an interesting way to proceed, it's now time to get on with the theory-building part of things. We'll start by looking at some very simple sentences and exploring what you already know about them in virtue of being a native speaker of a language, English in this case.

One of the things that you know about the sentences of your language is that they're made up of words and that certain of those words go together to create larger units. Since you don't always know *consciously* what you know about your language, we'll look at various traditional tests which can tease out what you know about the structure of sentences.

We'll then introduce *phrase-structure rules* and *phrase-structure trees*. Phrase-structure trees are used to illustrate graphically the structure of a given sentence. We'll then refine our phrase-structure rules and trees in order to better account for what you know about the different ways words are organized in the structure of sentences.

2.1 A simple starting point

Let's start by taking a look at an ordinary English sentence like (1), with a view towards exploring some of the things that you know about it:

(1) The student will meet her friend at the station.

The first observation to make with respect to (1) will strike you as the most blindingly obvious thing I could possibly say. I claim that part of what you know when you know your language is that a sentence like (1) is made up of words.

I know what you're thinking: 'Of course (1) is made up of words! How could it possibly be otherwise?' However, if we were to look at a sound spectrograph of (1), a visual representation of the sound energy created when (1) is produced,

you would see that there are no breaks in the sound energy corresponding to where the word breaks are. It's all one continuous flow. Therefore, even something as simple as the fact that you perceive (1) as having words is due to the cognitive system that you have for processing language. You might be inclined to think that you know about word breaks just because the breaks are there in the writing. However, this is just an accident of English. Plenty of other writing systems (like Japanese *katakana*) put breaks after every syllable rather than every word. Word breaks are also easily perceived by speakers who are illiterate, nor do you have any problem 'hearing' word breaks in sentences with new words or words that you've never seen written.

However, you know more than just that the sentence in (1) is made up of words. You also know intuitively that certain words in (1) belong, in traditional terminology, to the same part of speech. That is, *student*, *friend* and *station* are all nouns, *at* is a preposition, *meet* is a verb, and so on. You might again think that this is something that you've been taught explicitly at school, but in fact all you're doing at school is being given labels for things you already know about.

We'll take those intuitions over fairly directly, although we'll refer not to 'parts of speech', but rather to a word's *syntactic category*. Here are some examples of elements from various syntactic categories. As you can see, the correspondence between syntactic category and traditional 'part of speech' is pretty straight forward:

Table 2.1 Syntactic categories

Category	*Example*
Noun	table, sincerity, station
Preposition	at, on, in, by, over
Adjective	green, intelligent, generous
Verb	give, put, run, sleep, disturb
Auxiliary	can, should, must, will
Determiner	the, a, this, those, several, my

The only one that might be slightly unfamiliar is *determiner*. This is a sort of catch-all category which includes what are traditionally called articles (*the*, *a*), demonstrative adjectives (*this*, *that*, *many*, *several*), and the like. However, don't worry now about memorizing the syntactic category of every word. It'll become second nature in no time.

Now, not only do you know that each of the words in (1) belongs to a certain syntactic category. You also know that certain strings of words in (1) go together to form larger units and that other strings of words don't. For example, the two words at the beginning of the sentence *the student* go together to form a *noun phrase* (NP). On the other hand, the string of words *student will meet* does not form any kind of unit. *Her friend* and *the station* are also examples of noun phrases in (1). The nouns *student*, *friend* and *station* are the heads of these three noun phrases respectively. They are the most important elements – the pivots around which the entire phrase turns. Phrases can also be parts of larger phrases. For example, the noun

phrase *the station* combines with the preposition *at* to form the *prepositional phrase* (PP) *at the station*. In this case, the preposition *at* is the head of the prepositional phrase.

EXERCISE 1

Give the syntactic category of each word in the following phrases. What is the syntactic category of the entire phrase (e.g. noun phrase, adjective phrase, etc.)? Which word is the head of the phrase?

(a) the book about linguistics
(b) at the end of the platform
(c) see the answer
(d) aware of the problem

Phrase-structure trees and phrase-structure rules

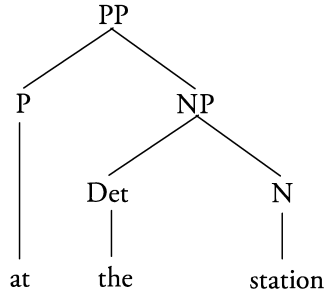
So it seems as though part of what you know when you know English is that there are two 'levels' of organization. There is the word or 'head' level and the phrase level. Each phrase has a head, although not every word in the sentence will be the head of a phrase.

In order to express your knowledge about the structure of phrases, we could simply list the information:

(2) *at the station* (PP)
  a. at: preposition (P) (head of phrase)
  b. the station: NP
    i. the: determiner (Det)
    ii. station: noun (N) (head of phrase)

However, a more compact way of illustrating this organization is to use a phrase-structure tree:

(3)



(3) indicates precisely the same information as (2). The entire phrase *at the station* is a PP, which is itself made up of a preposition combining with an NP. This particular NP consists of a determiner and a noun. Therefore, the phrase-structure tree makes a specific claim about what you know about your language. To draw the tree in (3) for the phrase *at the station* is to claim that your cognitive system assigns that structure to it.

However, you know more about your language than this. It's not just that this particular PP is made up of a P and an NP (and perhaps there are others which aren't). Rather, you know that *all* PPs are made up of a P and an NP. What your cognitive system seems to possess is the following phrase-structure *rule*:

(4) PP → P NP

(4) doesn't just make a claim about a particular PP, it says that all well-formed PPs are composed of a preposition followed by an NP. If we claim that this rule forms part of what you know when you know your language, then we can account for not just the PP in (3), but an infinite number of other PPs as well:

(5) to the store, at many banks, from a friend, towards the future, etc.

Similarly, the tree in (3) claims that you know something about the particular noun phrase *the station*. It's made up of a determiner plus a noun. However, we can account for more by assuming that you know the following phrase-structure rule:

(6) NP → Det N

This rule accounts for your knowledge of *the station* as well as an infinite number of other NPs:

(7) the book, that aeroplane, a review, several decisions, etc.

A well-formed NP can also have an optional preceding adjective phrase (AP), as well as an optional following PP:

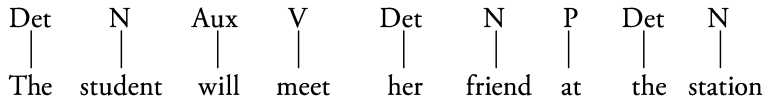
(8) The (green) books (about magic) are helpful on occasion.

Therefore, the rule in (6) should really be at least (9):

(9) NP → Det (AP) N (PP)

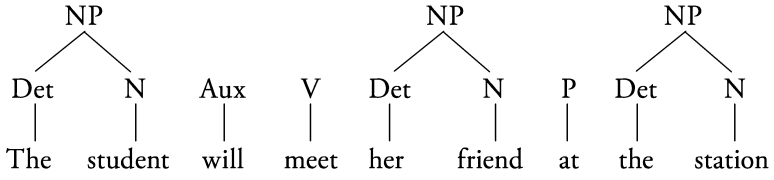
Having looked a little bit at the individual phrases from the sentence in (1), let's now see if we can draw the phrase-structure tree for the entire sentence. My personal feeling is that it's easier to work from the bottom up when first learning to draw phrase-structure trees, so let's start with the words and their syntactic categories, including now verbs (V) and auxiliaries (Aux):

(10)



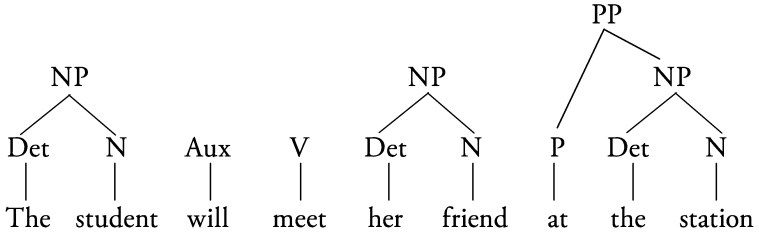
As we mentioned above, *the student*, *her friend* and *the station* are all noun phrases, so let's put those in:

(11)



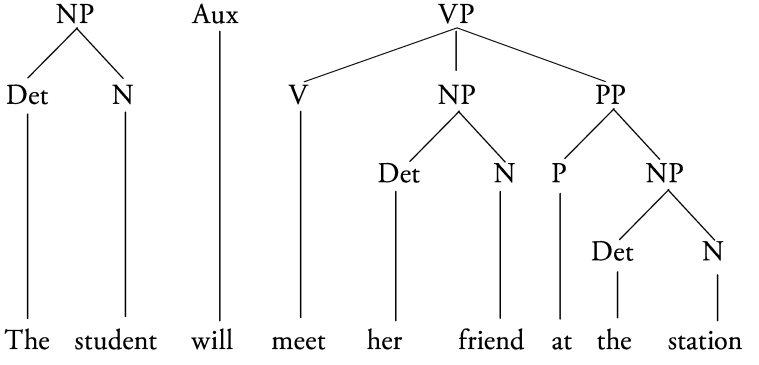
We also know that Ps go together with NPs to form PPs, so let's hook *at* and *the station* together:

(12)



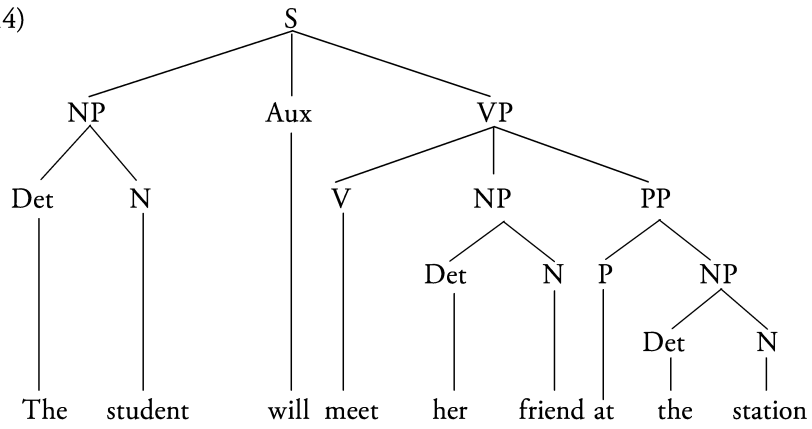
Now *meet* is a verb, and is therefore the head of a verb phrase (VP). *Her friend* is the direct object of the verb, and so therefore must be in the verb phase. The PP *at the station* gives the location where the meeting action took place, and so therefore is also in the verb phrase. You might think the auxiliary should also go in the verb phrase, but it doesn't, for reasons we'll discuss in Chapter 3.

(13)



We're now just left with the subject NP, the auxiliary and the VP. These things go together to form the sentence as a whole:

(14)



So now we have a phrase-structure tree associated with the sentence in (1). However, as before, we want to make claims about not just the sentence in (1) but an infinite number of other sentences. What we therefore want to do is take what we've done so far and generalize it. If we do that, we have the phrase-structure rules in (15):

(15) a. S → NP Aux VP
b. NP → (Det) (AP) N (PP)
c. VP → V NP PP
d. PP → P NP

Just as with our NP rule, though, we're going to need more rules for well-formed VPs. In (15), we have only one rule for VP expansion, VP → V NP PP. However, there are many more possible VPs than just this one. Consider the VPs in (16):

(16) a. John *slept* (VP → V)
b. Mary *gave Bill the book* (VP → V NP NP)
c. Ivan *went to the museum* (VP → V PP)

If we were going to try to account for a substantial fragment of English using phrase-structure rules, we'd need a lot more rules of the sort in (16). However, we won't be pursuing this at any length here because we're going to see in Chapter 3 that the phrase-structure rules (in the form that we've seen so far) need to be substantially modified anyway. However, this chapter is just about getting used to the idea that words are organized into larger groups and how phrase-structure rules work in general, so we'll keep using them for the moment.

EXERCISE 2

Using (14) as a model, draw the trees for (16a) to (16c) using the changes to the VP rule which are indicated in the parentheses.

2.2 Constituency

So far, we've been leaving things on a somewhat intuitive level when it comes to the details of sentence structure. We all share the feeling that, for example, *at the station* forms some kind of a unit, a PP in fact, whereas other strings of words don't form a unit. Let's call these units *constituents*. *Student will meet*, for example, doesn't feel like a unit of any kind.

Feelings and intuitions are all well and good, and it's true that they play an important starting role in just about any scientific project. However, it's also true that intuitions can be wrong. For example, I have the intuition that every morning the sun rises in the sky while the earth remains perfectly still. On some level, I know that that intuition is wrong, but that doesn't alter the fact that the intuition is very strong indeed. Therefore, it would be nice if we could devise some tests to back up our intuitions as to which groups of words do or do not form constituents. To that end, we'll take a look at some traditional tests which are used to indicate constituency: (1) the substitution test, (2) the coordination test and (3) the movement test. Some are a little bit more reliable than others for reasons which we'll see in later chapters.

The substitution test

The first test that we'll look at is called the *substitution test*. This constituency test relies on the fact that groups of words which function as constituents in the sentence can generally be replaced in a sentence by a single word. This confirms that the group of words forms a constituent. Depending on exactly the single substitution word you choose, you can even identify the syntactic category of the constituent you're dealing with.

Pronouns, for example, can be used to tell whether a given string of words constitutes an NP constituent. If a given group of words can be replaced by a pronoun, then that group of words is not only a constituent, but is specifically an NP. Consider the contrasts between the sentences in (17):

(17) a. The Belgian scientist who won the Nobel Prize at the age of 22 retired today.
b. She retired today.
c. *The Belgian scientist who she today.

The pronoun *she* is able to replace the entire string of words *the Belgian scientist who won the Nobel Prize at the age of 22*. On the other hand, as (17c) indicates, *she* cannot replace the string of words *won the Nobel Prize at the age of 22 retired*. This suggests that the string of words does not form an NP constituent.

The Belgian scientist who won the Nobel Prize at the age of 22 is not the only string of words in (17a) which can be replaced by a pronoun. For

example, we could take *the Nobel Prize* and replace it by the pronoun *it*, creating (18):

(18) The Belgian scientist who won it at the age of 22 retired today.

You'd have to have been talking about the Nobel Prize during a previous part of the conversation in order for (18) to be an appropriate thing to say, of course. Compare (18) with (19) for example:

(19) *The she won the Nobel Prize at the age of 22 retired today.

Unlike (18), (19) is just an unintelligible jumble. This is because *she* has replaced *Belgian scientist who* in (17a), and these words do *not* form an NP constituent.

EXERCISE 3

Using pronouns such as *he/she*, *it*, *them*, etc. identify all of the NPs in the following sentences:

(a) The new gorilla's attendant put a tasty banana on the purple table.
(b) The person responsible for security left the visiting dignitaries in a minuscule antechamber.

Just to have a general term to refer to these single words that can substitute for others, let's refer to them as 'proforms'. *Pronouns* substitute for NPs, but there are other proforms which can substitute for other syntactic categories.

For example, words such as *here*, *there* or *then* can substitute for certain kinds of PPs. Consider the contrasts in (20):

(20) a. Bill's book about economics in the eighteenth century was found in the dining room.
b. Bill's book about economics then was found in the dining room.
c. Bill's book about economics in the eighteenth century was found there.
d. *Bill's book about there century was found in the dining room.

Again, we see that the proforms *then* and *there* can only substitute for a constituent which is a PP, like *in the eighteenth century* or *in the dining room*. A group of words which is not a constituent, like *economics in the eighteenth* cannot be substituted for. Nor can *there* or *then* substitute for a constituent which is not a PP, like *Bill's book about economics*:

(21) *Then was found in the dining room

EXERCISE 4

Use PP proforms to identify the PPs in the following sentences:

(a) Mary put the book in Sue's jacket on Thursday.
(b) The agent who left the computer on the table at lunchtime was dismissed on the very day he was hired.

The coordination test

Conjunctions, elements like *and*, *or* and *but*, can serve as a very useful aid in determining the constituent structure of a sentence, since only constituents can be conjoined. Furthermore, both elements to be conjoined must be of the same syntactic category (i.e. two Ns, two APs, etc.). Sequences of words which are not constituents cannot be conjoined. Consider the following sentence:

(22) John's exposé in *The Guardian* gave new hope to the government's opponents.

We could use the substitution test from the previous section to identify several strings of words as constituents. For example, *John's exposé in The Guardian* is a constituent, as it can be replaced by the proform *it*.

(23) It gave new hope to the government's opponents.

We could, however, also use the coordination test to confirm this result by attempting to coordinate *John's exposé in The Guardian* with another NP. If the result is grammatical, than this would confirm that *John's exposé in The Guardian* is indeed a constituent:

(24) [John's exposé in *The Guardian* and the subsequent revelations] gave new hope to the government's opponents.

In (24), *John's exposé in The Guardian* has been coordinated with another NP *the subsequent revelations*, indicating that it is indeed a constituent.

We get the same results if we try to coordinate other constituents. For example, we can also have coordinations like (25a) (VP-coordination) and (25b) (PP-coordination):

(25) a. John's exposé in *The Guardian* [gave new hope to the government's opponents and put fear into ministers' hearts].
b. John's exposé in *The Guardian* gave new hope [to the government's opponents and to seekers of justice throughout the world].

On the other hand, if we were to take a string of words which does not form a constituent, the resulting coordination would be ungrammatical:

(26) *John's exposé in *The Guardian* [gave new and renewed old] hope to the government's opponents.

EXERCISE 5

Use the coordination test to identify three constituents in the following sentence:
John's predilection for original thinking placed his family in dire peril.

The movement test

In addition to substitution and conjunction, it's also possible to use movement as a test for constituency. We'll talk about movement in greater detail in Chapter 6. For the moment, all that's really important to notice is that certain groups of words seem able to be moved around as a unit in the sentence, whereas others cannot. For example, consider (27):

(27) John will leave a book on the table.

It is possible to take various constituents in (27) and move them to the front of the sentence, giving them an emphatic or contrastive interpretation. This is called *topicalization*:

(28) a. On the table John will leave a book (but on the chair he will leave some paper).
b. The book John will leave on the table (but the letter he will put in the bin).

On the other hand, if we try to take a string of words which does not form a constituent and topicalize it, the result is ungrammatical:

(29) a. *Book on, John will leave the table.
b. *On the, John will leave the book table.

Unfortunately, the movement test is the least reliable of the three tests that we've mentioned because movement in general, and topicalization in particular, are subject to a number of constraints which have nothing to do with constituent structure. For example, topicalization of a VP is rather odd in most dialects of English:

(30) Leave the book on the table, John will.

Probably the most well known 'speaker' of a dialect in which this is possible is the character Yoda from the *Star Wars* films. However, there *is* a serious point we can make. In writing dialogue for that character, the writers never topicalize non-constituents, so even if Yoda's topicalization is a little freer than most dialects of English, he would never produce sentences like those in (29).

Because of these 'outside' influences on movement, the movement test can only be used in a 'positive' way. That is to say, if something passes the movement test, then it definitely is a constituent. If it *doesn't* pass the movement test, then you can't conclude anything. It might be a constituent, or it might not. As an example of this, notice that *the table* in (27) is clearly a constituent (being an NP). It passes the substitution and coordination tests with flying colours (31):

(31) a. John left a book on it.
b. John left a book on the table and the chair.

However, if we use the movement test, (32) is rather odd:

(32) *The table, John will leave a book on, but the box, John will put a record in.

By contrast, the phrase *left a* fails the substitution and the coordination tests, and is definitely not a constituent. It also fails the movement test:

(33) a. *John it/there book on the table.
b. *John left a and threw the book on the table.
c. *Left a, John book on the table (but put a, John record in the box).

So the movement test should be used with a little caution. Only constituents pass the movement test, but both constituents and non-constituents can fail it. A quick confirmation with the other tests is often a good idea if there is any doubt.

EXERCISE 6

Use the movement test to identify two constituents in the following sentence: Mary can see her friends on Thursday in Birmingham.

2.3 The elements of a phrase-structure tree and their relations

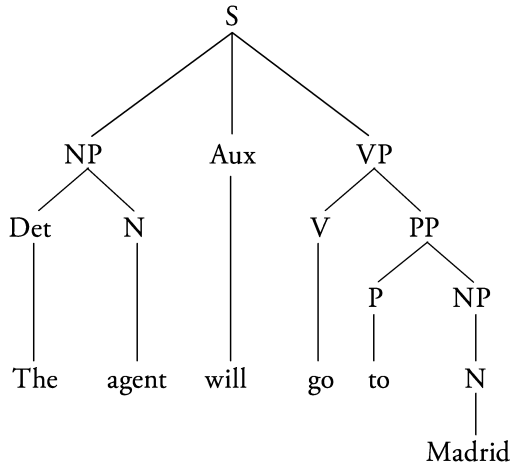
Having seen a few elementary phrase-structure rules and phrase-structure trees in the previous section, we need to introduce some terms for talking about phrase-structure trees and the relations between the elements in them, which

are called 'nodes'. We'll be using these terms so frequently throughout the book that they'll become second nature in no time.

We'll start off in a seemingly aggressive vein with the term *dominate*. One node *dominates* another when you can trace a path from the first node to the second one moving only downward through the tree. Take a look at the sentence in (34) and the phrase-structure tree representation for it in (35):

(34) The agent will go to Madrid.

(35)

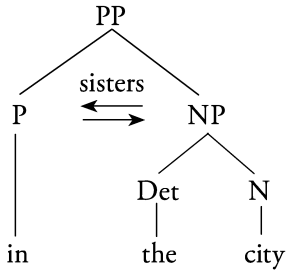


Notice that the S node dominates *every* other node in the tree. Starting from S, you can get to any of the other nodes in the tree by moving only downward. On the other hand, the PP node dominates only the P node and NP node associated with *Madrid*. The PP node does not, for example, dominate the NP node associated with *the agent*. You can't get from the PP node to the NP node associated with *the agent* by only going down. You need to go up through VP and S before going down.

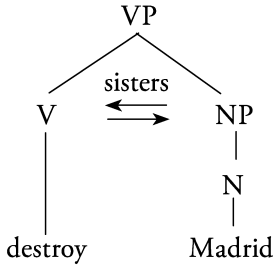
A more specific type of domination is *immediate domination*. A node *immediately dominates* every node which you can get to by moving down exactly *one* level. So while S dominates every other node in the tree, it immediately dominates only the NP *the agent*, the Aux *will* and the VP *go to Madrid*. Similarly, the VP node dominates five nodes (V, PP, P, NP and N), but it only *immediately* dominates V and PP.

As a point of cultural interest, nodes in a phrase-structure tree are always female rather than male. It's not uncommon to see the nodes which something immediately dominates referred to as its *daughter* nodes. Less often, you'll see a node which immediately dominates something referred to as its *mother* node. One place, however, where the matriarchal term is used very frequently is for *sister* nodes, that is, daughters of the same mother, as in (36) or (37):

(36)



(37)



Notice that these relational terms can give us a way of defining a constituent. Any group of heads which are *exhaustively dominated* by a given node (i.e. there is a node which dominates every one of those heads and no others) is a constituent. For example, we know that the group of words *to Madrid* is a constituent, and that's represented in the tree in (35). The PP node dominates the P *to* and the N *Madrid*, and no other heads.

The last relation between nodes that we'll look at in this section will play a major role in the chapters to come: c-command. C-command was first formulated by a linguist named Tanya Reinhart (1976), and has since come to play a fundamental role in syntactic theory. The notion of c-command seems relevant to almost every phenomenon. (We'll see most of the common ones at some point in this book.)

C-command is usually defined in the following way:

(38) A node α c-commands another node β if and only if the first branching node which dominates α dominates β.

In the tree in (35), for example, the subject NP node (dominating *the agent*) c-commands every other node in the sentence apart from S itself. The first branching node which dominates the subject NP node is the S node, and that S node dominates every other node in the sentence. However, the NP *Madrid* c-commands only the preposition *to*. The first branching node dominating *Madrid* is the PP node, and that dominates only the preposition. (Just as a side note, even though it technically satisfies the definition, nodes don't c-command themselves or any nodes that they themselves dominate.) Using the relational terms of this section, we could put (38) another way by saying

that a node c-commands (1) its sisters and (2) every node that its sisters dominate.

EXERCISE 7

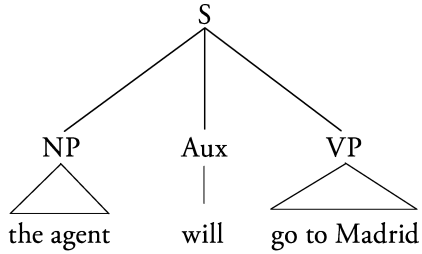
Draw the tree for 'The man will see the book'. Indicate three pairs of nodes which are sisters. What nodes does Aux c-command?

Constituents, phrase-structure trees and the 'triangle' notation

Having discussed various ways of referring to the elements of phrase-structure trees, I want to end this section by briefly introducing a notational device that we'll be starting to use in *drawing* phrase-structure trees and which is something we'll use a lot during the rest of the book. As we've been discussing, phrase-structure trees encode a lot of information. But often, when we're explaining a particular point or showing how a particular analysis works, not all of that information is equally important. We might want to focus on *certain* aspects of the structure of the sentence, but not really be interested in other details. When we don't want to worry about the details of a particular constituent because they aren't relevant for our discussion, we'll draw a triangle under the node in the phrase-structure tree.

So, let's say that we were talking about the sentence in (34) again. The complete tree is in (35), but if we were just making the point that the subject NP and the VP c-commanded each other, we might draw an abbreviated phrase-structure tree like (39):

(39)



The tree in (39) indicates that *the agent* is an NP and *go to Madrid* is a VP, but doesn't indicate the internal structure of these constituents, because that's not what we're focusing on. As we move on to more complex syntactic analyses and arguments, it'll be handy to be able to take some of the structure and 'put it off to the side' in this way. However, when drawing trees for exercise purposes, for example, it wouldn't be appropriate to use triangles, since the point of the exercise is to show exactly what the complete structure is.

2.4 Are two levels enough?

So far, we've seen that knowing your language entails not only knowing that individual words belong to a syntactic category, such as Noun or Verb, but also that certain words go together to form larger phrases, such as Noun Phrases and Verb Phrases.

However, there's some evidence to indicate that two levels, the word/head level and the phrase level, isn't enough. The cognitive system that you possess seems to analyse phrases in terms of *three* levels, not two. The idea that there are *heads* (things like nouns, prepositions, adjectives, etc.) and *phrases* (noun *phrases*, adjective *phrases*, etc.) may well be familiar to you from somewhere, whether it's traditional English grammar instruction or maybe a foreign language that you studied. However, probably nobody's ever tried to convince you that there is a level of organization *between* the head (word) level and the phrase level.

In order to show that there are constituents which are intermediate between the head level and the full phrase level, we can use exactly the same arguments that we used in the previous section to show that phrases exist.

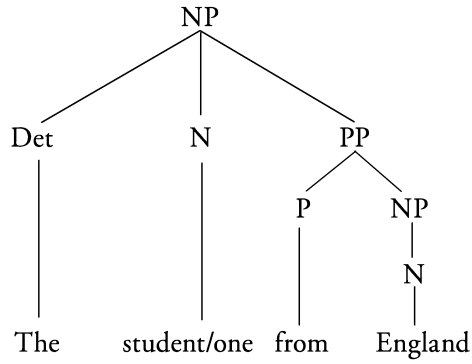
The structure of NPs

To start off with, consider (40), focusing in on *one*:

(40) Jack met the student from England, and I met the one from France.

Clearly, *one* in (40) is some kind of proform. The question is what it's substituting for. Given (40) the answer looks fairly straightforward. *One* appears to be substituting for N, so we have a structure like (41):

(41)



However, consider (42):

(42) *Jack met the student of philosophy, and I met the one of Greek.

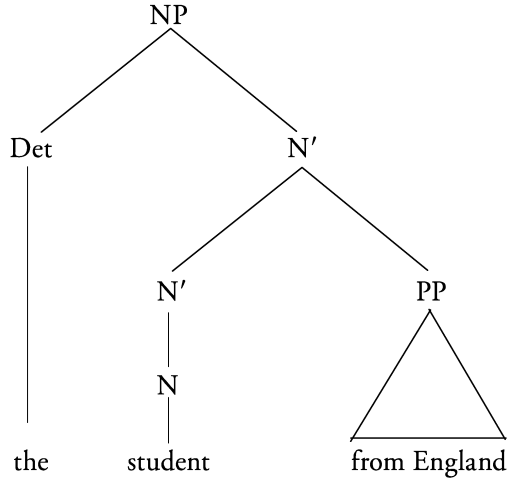
(42) presents a problem. On the basis of (40), it looks like we just established that *one* substituted for N. If *one* can substitute for the N *student* in (40), why can't it substitute for the N *student* in (42)?

One way to approach the problem is to question the assumption that there are only two levels of phrase structure (N and NP). At the moment we only have the head level and the full phrase level. Therefore, there are only two things that *one* could be substituting for: N or NP. But if we say that *one* substitutes for N, then there's no explanation for why (42) is ungrammatical. On the other hand, if *one* substitutes for NP, then (42) is predicted to be ungrammatical, but so is (41). That's also wrong.

However, if there's a *third* level, between the head level and the full phrase level, and *one* is a substitute for this *intermediate* category, we can explain the difference between (41) and (42). Following standard usage, we'll call this intermediate level the N′ level. ('N′' is read as 'N-bar', by the way – a long story.)

Let's see how postulating three levels for NPs instead of two, and assuming that *one* substitutes for N′, helps us out, starting with (41). Under this new three-level hypothesis, if *one* can substitute for *student* to the exclusion of the prepositional phrase *from England*, then *student* must be *both* an N *and* an N′ at the same time. In other words, the structure of *student from England* is really (43):

(43)



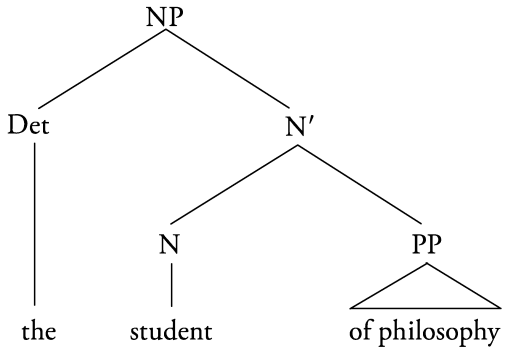
The prepositional phrase *from England* is not a sister to N, but rather is a sister to N′.

Turning now to (42), if *one* can't substitute for *student* in (42), that must mean that *student* is *not* both an N and an N′ in (42). *One* appears only able to substitute for the entire phrase *student of philosophy*:

(44) Jack met this student of philosophy, but I met that one.

That would be consistent with the following analysis of the NP *the student of philosophy*:

(45)



Here, rather than *student* being both an N and an N′, the only N′ is the phrase *student of philosophy. Student* by itself is not an N′.

If we look at what happens when we have *both* an *of* PP and a *from* PP, we can see that it looks like our analysis in (43) and (45) is on the right track. Consider the following NP:

(46) the student of philosophy from England

If we apply the *one* substitution test to (46), this seems to confirm our analysis. The whole of *student of philosophy* is an N′, but *student* by itself is not.

(47) a. Jack met the student of philosophy from England and I met the one from France.
b. *Jack met the student of philosophy from England and I met the one of Greek from France.

Notice also that the tree in (43) predicts that *student from England* is also an N′, in addition to just *student.* That also seems right, as *student from England* can also be substituted by *one.*

(48) Jack met this student from England, but I met that one.

Having introduced this third level of organization, the N′ level, as a solution to the *one* substitution problem, we now need to look around to see if there is any independent supporting evidence to back this analysis up. Do the other tests for constituent structure that we introduced in the previous section confirm the N′-analysis?

The movement test doesn't seem to. If we attempt to topicalize our hypothesized N′, the result is ungrammatical:

(49) *Student of philosophy, I met the yesterday, but student of Greek I met the today.

However, bear in mind what we said about the movement test. If a string of words can be moved as a unit, then it definitely is a constituent. If the string of

words can't be moved as a unit, then it might or might not be a constituent. So the ungrammaticality of (49) is, in the end, inconclusive.

However, we still have the coordination test. In order for two things to be coordinated, they must be the same syntactic category and the same level. That is to say, you can only coordinate two NPs, for example, not an NP and an N. Also, the combined phrase as a whole has the syntactic category of its two conjuncts. So two conjoined NPs together form an NP, and can appear anywhere a 'normal' NP can.

Looking at (45), this structure claims that *student of philosophy* is an N′ to the exclusion of the determiner *the*. Can we coordinate two N′s to the exclusion of the determiner? It appears as though we can. Consider the NP in (50):

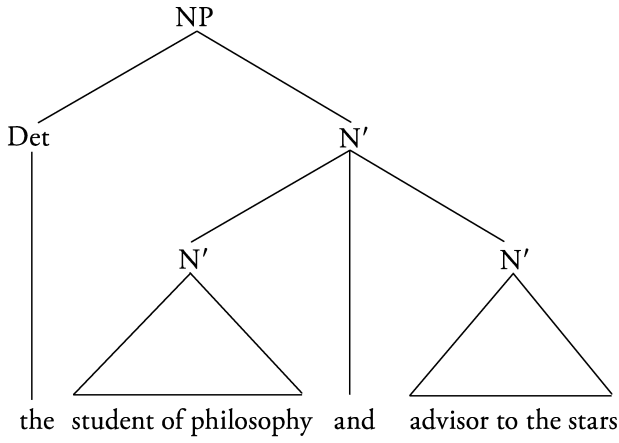
(50) The student of philosophy and advisor to the stars.

In (50), we've conjoined *student of philosophy* and *advisor to the stars*. Clearly, *advisor to the stars* is not an N. Neither is it an NP. It has no determiner and can't appear as the subject of a sentence:

(51) *Advisor to the stars predicted disaster for Brad and Jennifer.

Therefore, it seems as though *advisor to the stars* must be an N′. That would seem to show that *student of philosophy* is also an N′. Therefore, the structure of (50) must be (52):

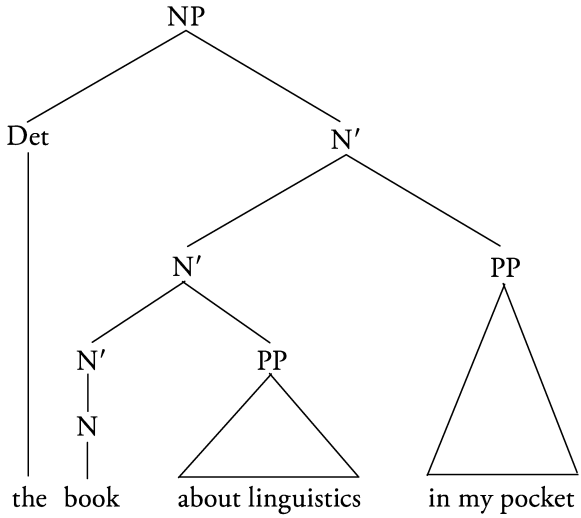
(52)



Notice that the N′ level, unlike NP or N, can be repeated or *iterated* as many times as necessary. For example, we could have an NP like (53), which would have the structure in (54), as shown by (55):

(53) The book about linguistics in my pocket.

(54)



(55) a. I like the book about linguistics in my pocket more than the one about physics on the table. (*one* = *book*)
b. I like the book about linguistics in my pocket more than the one on the table. (*one* = *book about linguistics*)
c. I like this one more than that one. (*one* = *book about linguistics in my pocket*)

Since *book*, *book about linguistics* and *book about linguistics in my pocket* are substitutable by *one*, they must all be N′s.

So, with the addition of (iterable) N′s, we'll need to modify our phrase-structure rules for NPs. In Section 2.1.1, we had the following rule for NP:

(56) NP → (Det) (AP) N (PP)

But this phrase-structure rule has only two levels, NP and N. We need to split this rule into two parts in order to introduce the N′ level:

(57) a. NP → (Det) N′
b. N′ → (AP) N (PP)

(57) claims that every NP immediately dominates at least an N′ and every N′ immediately dominates at least an N. However, because N′s can be iterated, we have to leave open the possibility of N′s immediately dominating other N′s, as in (54). So we need three rules instead of two:

(58) a. NP → (Det) N′
b. N′ → (AP) N′ (PP)
c. N′ → (AP) N (PP)

By the addition of the phrase-structure rule in (58b), we can create structures like the one in (54). Therefore, if (58) constitutes part of what you know when you know English, we can explain how you know that (54) has the structure it does. We can therefore explain how you know facts about *one* substitution in sentences like (40), (42) and (55).

The structure of VPs

We can give a parallel argument for the existence of an intermediate level between the head and the phrase level for verbs, using the proform *do so*. Consider (59):

(59) Joe will buy a book on Tuesday, and Sam will do so on Friday.

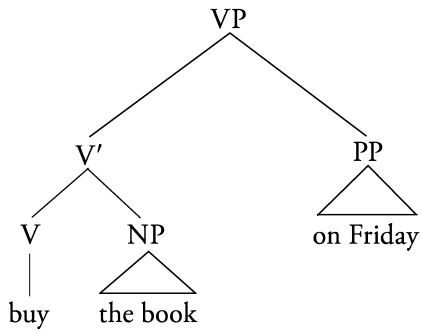
Just as with *one* above, the question is what *do so* is substituting for in (59). In (59) it is substituting for the sequence of words *buy a book*. Therefore, *do so* must be substituting for more than just V. To confirm this, if we try to use *do so* to substitute for just V, the result is ungrammatical:

(60) *Joe will buy a book, but Sam will do so a newspaper.

However, under a two-level system, our only other choice is that *do so* is substituting for VP, but that doesn't look right either. If *do so* were substituting for VP, we'd expect (59) to be ungrammatical. The prepositional phrase *on Friday* is clearly part of the VP. It's indicating the time at which the buying of the book will take place.

Do so in (59) is substituting for *part*, but not *all*, of the VP. If the substitution test is genuine, and only sequences of words which are constituents can be substituted, then there must be some intermediate level of constituency that's *between* V and VP which *do so* is substituting for. If we have a three-level system, incorporating a V′ level between VP and V, we would seem to be able to capture the contrast. *Do so* is an element which substitutes for V′, rather than V or VP. Then, if the PP *on Friday* is a sister to the V′ *buy the book*, then we explain why *do so* can replace *buy the book* to the exclusion of the PP *on Friday*:

(61)



We can also explain why (60) is ungrammatical. Just as *of* PPs must be sisters to N and cannot be sisters to N′, direct objects must be sisters to V. They cannot be sisters to V′. It is *buy* together with the direct object *the book* which is a V′.

Just as with the NPs that we saw in the previous section, the intermediate level in VPs is iterable as well. So, for instance, we could have a sentence like (62):

(62) John gave a speech at 3 o'clock on the 30th of June in Madrid.

BEFORE CONTINUING

Use the *do so* test to show that there must be 3 V′s in (62).

Do so can replace *gave a speech*, *gave a speech at 3 o'clock* or *gave a speech at 3 o'clock on the 30th of June*.

(63) John gave a speech at 3 o'clock on the 30th of June in Madrid ...
a. ... and Mary did so at 5 o'clock on the 27th of September in Valencia.
b. ... and Mary did so on the 27th of September in Valencia.
c. ... and Mary did so in Valencia.

Notice too that, as we'd predict, *do so* cannot substitute for *gave* to the exclusion of *the speech*, since *gave* is not by itself a V′. It is only a V.

(64) *John gave a speech at 3 o'clock on the 30th of June in Madrid, and Bill did so a statement later that day.

Before we leave this topic, though, I want to mention a sentence which may have occurred to you:

(65) Mary asked John to give a speech at 3 o'clock on the 30th of June in Madrid, and he did so.

Here *do so* seems to be substituting for the entire VP *give a speech at 3 o'clock on the 30th of June in Madrid*.

Sometimes when you have just *do so* by itself, *do* and *so* flip around the subject. So (66) sounds pretty unnatural:

(66) John bought a new book yesterday and Bill did so.

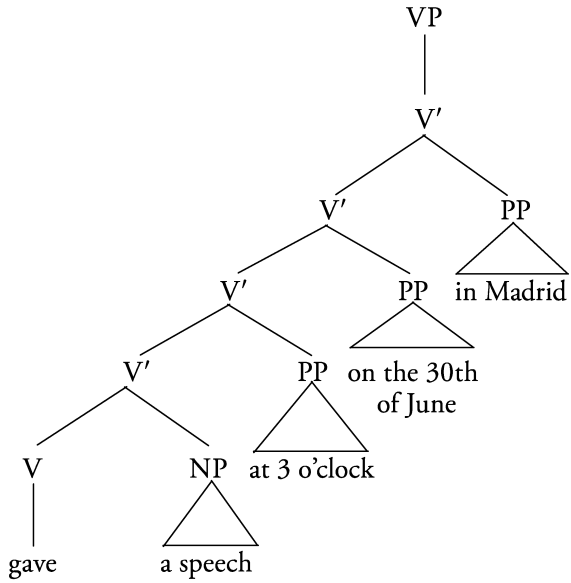
You need to have 'as well' or something in there. However, (67) is fine.

(67) John bought a book yesterday and so did Bill.

In either case, *do so* again seems to be substituting for the whole VP.

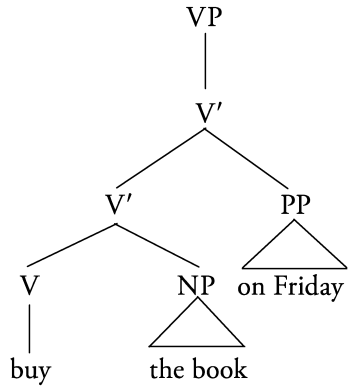
There are a number of ways of dealing with this problem which are all essentially identical. What we'll do here is maintain that *do so* substitutes for V′, so anything that can be replaced by *do so* must be a V′. That means that the VP *gave a speech at 3 o'clock on the 30th of June in Madrid* in (62) is actually a VP which immediately dominates a V′, which in turn immediately dominates the V′ *gave a speech at 3 o'clock on the 30th of June* and a PP *in Madrid*, as illustrated in (68).

(68)



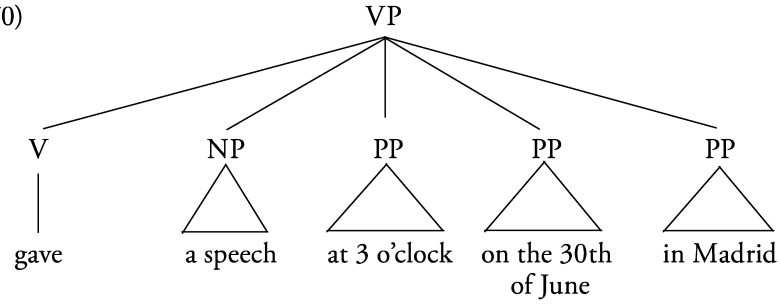
And that means that the tree in (61) above should actually have been (69).

(69)



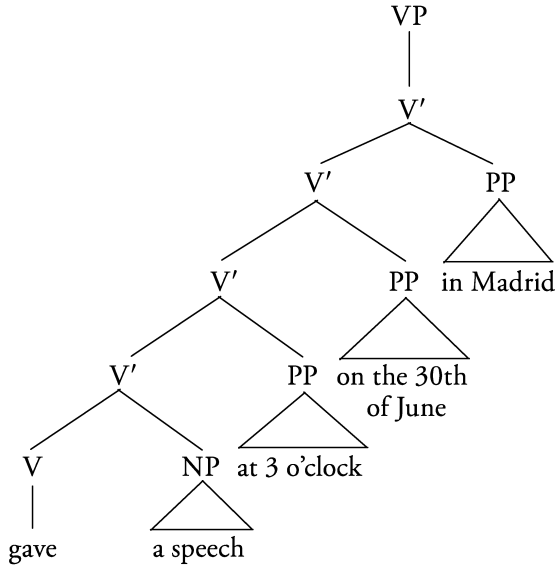
However, this doesn't in any way threaten the three-level analysis that we've proposed here. Under the simple two-level analysis, where all you have are Vs and VPs, there are no subconstituents within VP. This would mean that the only possible analysis for (62) would have a 'flat' tree, as in (70):

(70)



Remember that trees make claims about knowledge – specifically, your knowledge of English. (70) makes the claim that there's a VP *gave a speech at 3 o'clock on the 30th of June in Madrid* and no other constituents. And that seems to be a wrong claim about your knowledge of English. As we've seen, the substitution test indicates that you know that *gave a speech*, *gave a speech at 3 o'clock* and *gave a speech at 3 o'clock on the 30th of June* are all subconstituents of VP. Put another way, you know that the sentences in (63a) to (63c) are grammatical. Therefore, the structure of the VP in (65) must be (71):

(71)



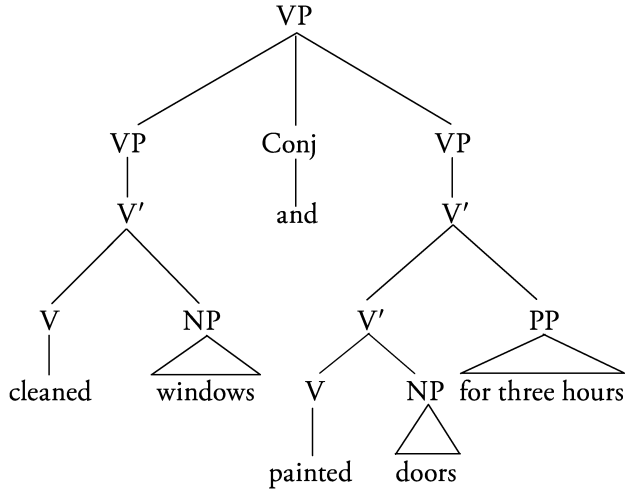
In addition to the substitution test, we can also call on the coordination test to verify the need for an intermediate level between V and VP. Consider the following sentence:

(72) John cleaned windows and painted doors for three hours.

(72) is actually ambiguous between two possible conjunctions. One is a conjunction of VPs, in which John cleaned windows for some unspecified amount

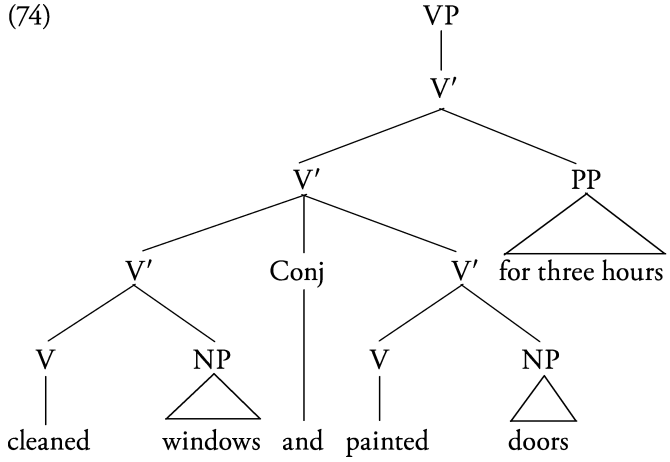
of time and then painted doors for three hours. That interpretation would be associated with the tree in (73):

(73)



However, of more interest to us is the fact that there is a second interpretation for (72) in which John was engaged during a three-hour period in window-cleaning and door-painting. That interpretation is only possible if *cleaned windows and painted doors* forms a constituent within the VP to the exclusion of the PP *for three hours*. But if the VP itself has no internal structure, as is the case under the two-level analysis, then there's no way to explain this. Under a three-level analysis, however, we can explain this second interpretation as arising from a conjunction of V′s, as illustrated in (74):

(74)



This would seem to confirm the three-level analysis for VPs.

On the basis of all this, we'll have to modify our phrase-structure rules for VP in the same way we did for NP. Instead of having two levels, as in (75), we need three levels, as in (76):

(75) VP → V (and various other things)

(76) a. VP → V′ (XP)
b. V′ → V′ (XP)
c. V′ → V (XP)

Again just as with NPs, (76b) is for when V′ iterates and (76c) is for when we finally get to the head. 'XP' in (76) is just a placeholder to indicate that any kind of phrase is possible (PPs, NPs, etc.).

Intermediate levels and other categories

We've now seen two categories, NP and VP, where we can construct arguments from substitution for a three-level system. Is there anything to be said about the other major syntactic categories, AP and PP? Well, yes. Consider the following AP and PP:

(77) very nervous about the exam

(78) completely over the line

Once again, it looks as though there's a proform which can substitute for A′ and P′: *so*. It's a word that does a lot of duty in English, and we saw it above in the V′ proform *do so*. In (77), it looks as though the string of words *nervous about the exam* can be replaced by *so*:

(79) John was very nervous about the exam, but Mary was less so.

So is not substituting in (79) just for *nervous*. It doesn't mean that Mary was less nervous in general. It means specifically that Mary was less nervous about the exam. With respect to (78), *so* can substitute for *over the line* to the exclusion of *completely*.

(80) The first shot went completely over the line, but the second one only partially so.

The conclusion that we have three levels in APs and PPs as well as NPs and VPs would seem to be confirmed by the coordination test.

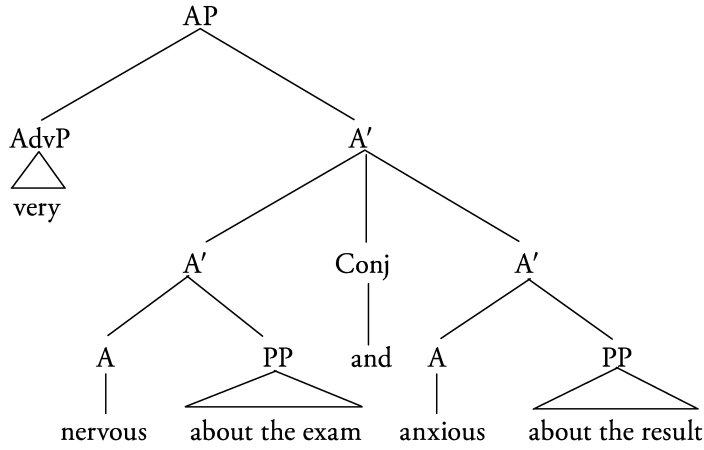
(81) very [nervous about the exam and anxious about the result]

(82) nearly [over the line and into the net]

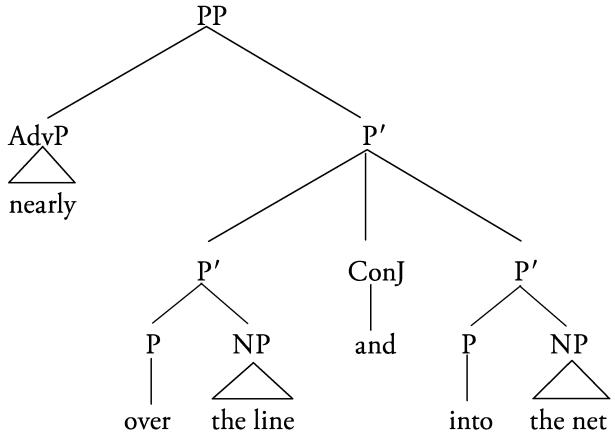
The fact that *nervous about the exam* and *anxious about the result* can be conjoined to the exclusion of the adverb *very* suggests that A's are being conjoined.

Similarly, (82) seems to show a conjunction of P′s. Representing (77) and (78) using intermediate bar-level categories, we have the following:

(83a)



(83b)



We would, of course, need to then modify our phrase-structure rules for APs and PPs in the same way as we have for NPs and VPs to add in the third level.

2.5 Constituency tests: more applications and case studies

As we mentioned above, the main purpose of phrase-structure trees is to represent the hierarchical organization of the constituents of a phrase or sentence – how the units of the sentence are built up into larger and larger structures. If you're unsure about how a particular phrase or sentence is organized (in other words, how to draw the tree), the tests for constituency that we've been taking about in this chapter can help.

Where do APs go when they modify nouns?

Let's consider, for example, how you might draw the tree for a noun phrase with an adjective in it, like (84):

(84) the blue book

We know that the whole thing is an NP, with *book* as its head. *The* is a determiner, and *blue* is both an adjective and an adjective phrase, but how is the adjective phrase hooked up to the head noun? To figure this out, we run through the various substitution tests to find out how many intermediate level categories we have, and build up the tree around those.

So, in the case at hand, we're talking about NPs. So what we want to test for is N′, and to do that we need to use the *one*-substitution test. Wherever we can substitute *one*, we have an N′.

First, *blue book* seems to be an N′, as (85) illustrates:

(85) I like this blue book more than that one.

In (85), *one* is not just substituting for *book*. If it were, (85) should mean 'I like this blue book more than that book, which could be of any colour'. But that's not what it means. (85) means 'I like this blue book more than that blue book'.

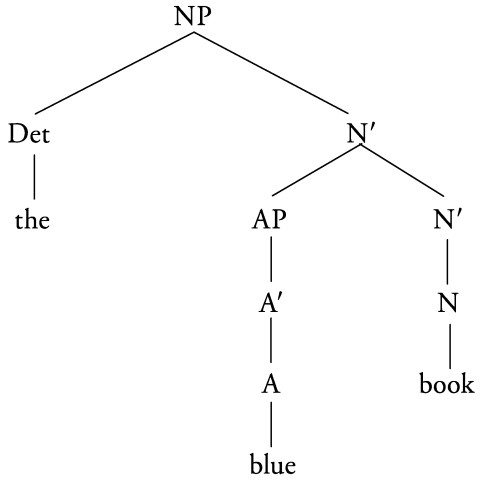
The next question is whether the adjective phrase *blue* is a sister to N or a sister to N′. Put another way, is *book* by itself both an N′ and an N (like *student* in (43) above), or is *book* just an N and not an N′ (like *student* in (45) above)? All we need to do to find out is see whether *book* by itself can be substituted by *one* to the exclusion of the AP *blue*. It looks like it can:

(86) I like the blue book more than the red one.

So, since *book* by itself can be substituted by *one*, *book* is both an N′ and an N.

Therefore, the tree for (84) is (87):

(87)

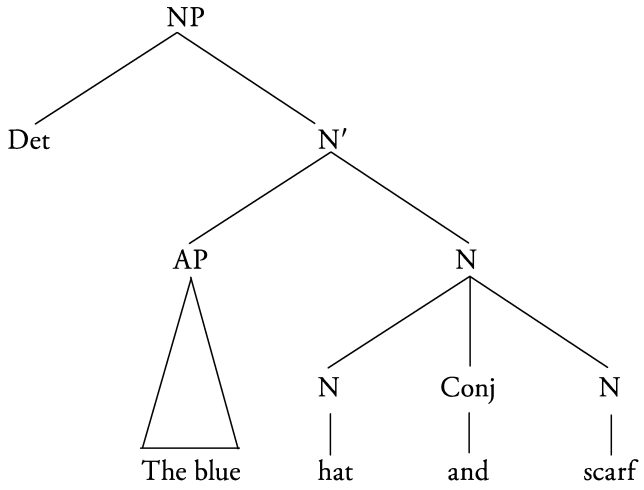


The NP immediately dominates Det and N′, since we established that *blue book* is an N′. This N′ in turn immediately dominates AP and another N′, since we established that *book* by itself is also an N′. The AP node immediately dominates A′ and A, since there are no other lexical items in the AP. The N′ then in turn immediately dominates the N, and we're all set. Notice just in passing here that the *coordination* test doesn't definitely establish one of the key questions with respect to (84), which is whether the AP *blue* is a sister to an N′ or a sister to an N. One might observe that (88) is grammatical:

(88) The blue hat and scarf are on the table.

However, all this tells us is that *hat* and *scarf* are constituents. It doesn't tell us what *kind* of constituent. The grammaticality of (88) is compatible with (87), which is the right analysis, but it's also compatible with (89), where *hat* and *scarf* are Ns, which is the wrong analysis.

(89)



The *one*-substitution test is more specific, in that it identifies *book* (or *hat*/*scarf*) not just as a constituent, but specifically as an N′.

Sentence-final PPs

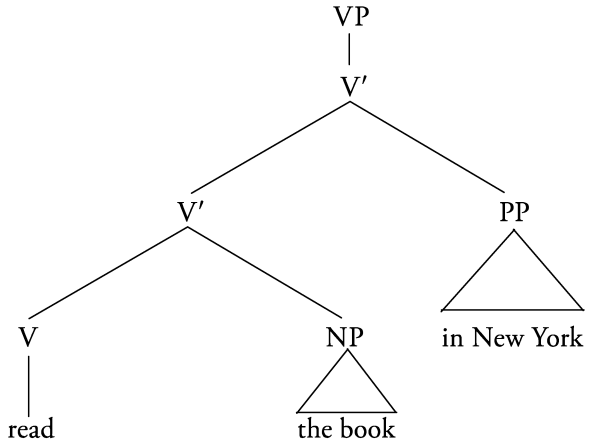
Applying the constituency tests in a systematic and methodical way can also help deal with what can be an annoying problem when it comes to drawing trees: final prepositional phrases – that is, prepositional phrases which come at the end of a phrase or, particularly, at the end of a sentence. PPs can appear in all kinds of different places modifying different constituents. The problem is that when they appear at the end of a phrase or sentence, the word order doesn't give you a clue about where they go.

Now more often than not, context and real-world knowledge will give you a good idea of whether a PP is, for example, part of an NP or a VP. Consider (90), for example:

(90) John read the book in New York.

Just about the only sensible interpretation for PP in (90) is one in which the PP gives the place in which the book-reading took place. As such it is a sister to the V′ *read the book*, as illustrated by the tree in (91):

(91)

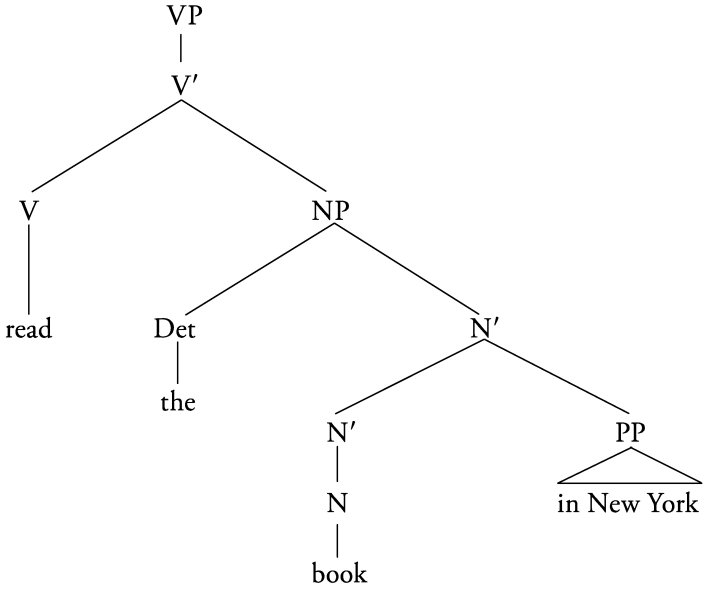


The tree in (91) can be confirmed by the *do so* test, by noting that the sentence in (92) is grammatical.

(92) John read the book in New York and Mary did so in Boston.

However, notice what happens if the PP *in New York* were modifying *book*, like this:

(93)



Constituents are sisters to constituents that they modify semantically, so (93) corresponds to an interpretation in which *in New York* modifies *book*, but not *read*. This would have to correspond to a situation where the book is, say, located in New York, but the reading didn't necessarily take place there. Without any context, this reading for (90) is not very strong, I don't think. However, the thing to notice is that the string of words in left-to-right order is exactly the same as the tree in (91). In both cases, the PP would be spelled out as the final element of the sentence. This is a case where the potential for confusion is more theoretical than actual, though, because of the pragmatics of the sentence.

Notice though that, however implausible the reading expressed by (93) is, it's *completely impossible* in (92). The *do so* test explicitly disambiguates the sentence. Since *do so* is substituting for the V′ *read the book*, that V′ (and therefore *the book*) explicitly does not include the PP *in New York*. It can therefore be interpreted only as the location where the reading of the book took place.

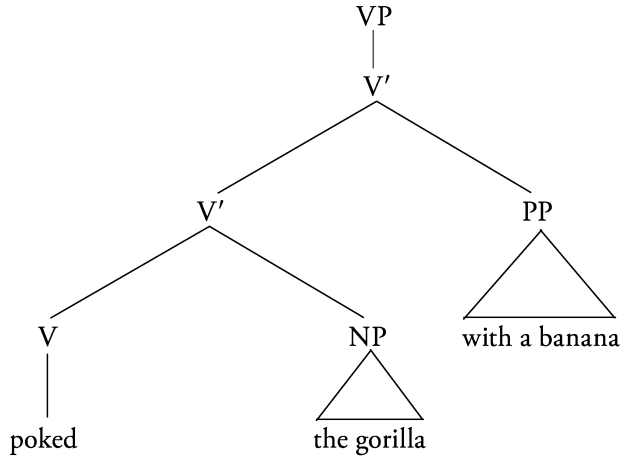
We can use this explicit disambiguation in other cases which are more difficult, or even where a sentence in isolation is fully ambiguous. Consider (94), focusing on the sentence-final PP *with a banana*.

(94) The zookeeper poked the gorilla with a banana

Without any supporting context, (94) is ambiguous between two different interpretations: one where the banana is the instrument of poking and one where the gorilla is the possessor of a banana. Those two interpretations differ precisely with respect to where the PP is in the tree.

The reading where the banana is the instrument of poking corresponds to a tree like (91) above. Semantically, the PP is modifying *poked the gorilla* and therefore is a sister to V′ within the VP.

(95)



We can use the *do so* test to bring out this reading. If indeed the PP is a sister to the V′ *poked the gorilla*, then we should be able to substitute *do so* for that phrase.

(96) The zookeeper poked the gorilla with a banana and Bill did so with an umbrella

The substitution by *do so* forces the interpretation in the second conjunct to have an instrument-of-poking reading for the *with* PP, and the parallelism of the construction then forces that interpretation back onto the first conjunct. If *with a banana/umbrella* had been modifying *gorilla*, it would have been substituted by *do so* along with the rest of the direct object.

On the reading where the gorilla is the possessor of a banana, the PP *with a banana* is modifying *gorilla*.

BEFORE CONTINUING

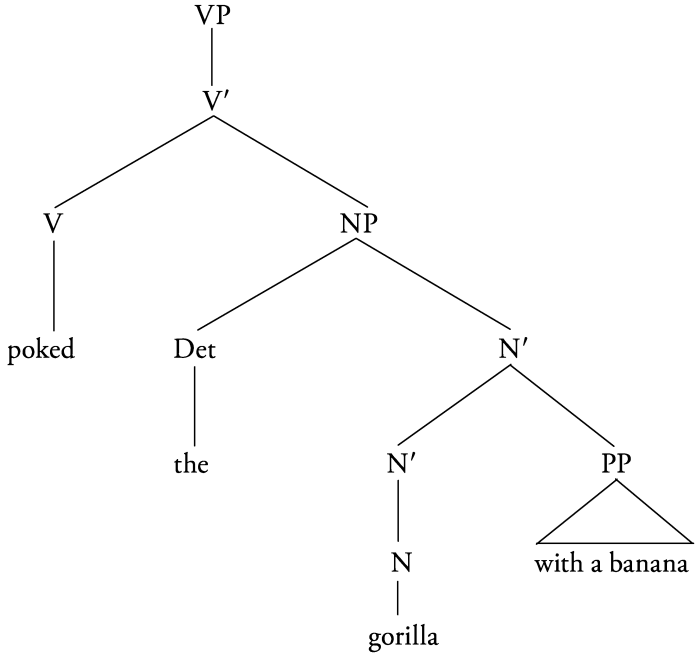
Is the PP *with a banana* a sister to N′ or a sister to N?

If we use the *one*-substitution test, we see that *gorilla* is in fact both an N and an N′:

(97) the gorilla with a banana and the one with a coconut

Therefore the VP for (94), on the reading where the gorilla is a possessor of a banana, will be (98):

(98)



And we can now use the *one*-substitution test again to force this reading:

(99) The zookeeper poked the gorilla with a banana and Bill poked the one with a coconut

So, to summarize, the various constituency tests that we've talked about in this chapter, particularly the various substitution tests, are a valuable resource for helping you figure out what the tree should be for a given sentence. If you don't have a good feeling for what a sentence final PP should be modifying, for example, or if you just want to double-check your intuitions, the various constituency tests are available.

Summary

In this chapter, we've made a start at characterizing some of the things that you know when you know a language. Sentences are not perceived as one big blur, even if that's exactly what they are from an acoustic point of view. The language faculty which you were born with breaks the acoustic blur into words, which are then organized into larger constituents. We saw that various tests, such as substitution, coordination and movement, seem to operate only on constituents, thereby providing evidence for their existence.

In particular, we argued that NPs, VPs, APs, etc. actually have three levels to their structure. In addition to the phrase level (e.g. *the book about linguistics*) and the head level (e.g. *book*), there is also an intermediate level of organization (e.g. the N′ *book about linguistics*). We found various proforms which substitute for this intermediate level (*one* for N′, *do so* for V′ and *so* for A′ and P′) and these results were confirmed by the coordination test.

Open issue: The structure of the double object construction

One area of sentence structure which has been something of a thorn in linguists' sides has to do with the phrase-structure representation of what is called the *double object* construction. Certain verbs, such as *give*, *show* and *send*, have alternations like the one seen in (100):

(100) a. I showed a book to John.
b. I showed John a book.

In both cases, *show* is taking both a direct object (*a book*) and an indirect object (*John*). In (100a), the direct object appears in a standard direct object position following the verb and the indirect object appears in a PP. Nothing particularly weird there. In (100b), on the other hand, we have the 'double object' version. The indirect object has shifted in front of the direct object, and neither NP has a preposition.

We need to figure out what the structure is in (100a) and (100b). The way to do that is run the sentences through various tests which we know depend on structure and see how they come out. As mentioned in Section 2.3, a lot of

syntactic processes seem to depend on c-command. If we can get a handle on what the c-command relationships are within the VPs in (100a) and (100b), we can figure out what the right structure is. We'll have much more to say in later chapters about these sorts of tests, but for now you'll have to just trust me that c-command is relevant for these things.

First of all, as we'll see in Chapter 5, it turns out that reflexives like *herself* and *ourselves*, as well as reciprocals like *each other*, need to be c-commanded by their antecedents. Just to give a quick example, *John likes himself* is grammatical because *John*, the subject (and the antecedent of *himself*), c-commands the direct object reflexive *himself*. On the other hand, *himself likes John* is ungrammatical because *John* fails to c-command *himself*. With that in mind, consider (101):

(101) a. I showed Mary to herself.
b. *I showed herself to Mary.

The fact that (101a) is grammatical while (101b) is not suggests that the direct object (immediately following the verb) c-commands the NP object of the preposition, but not vice versa.

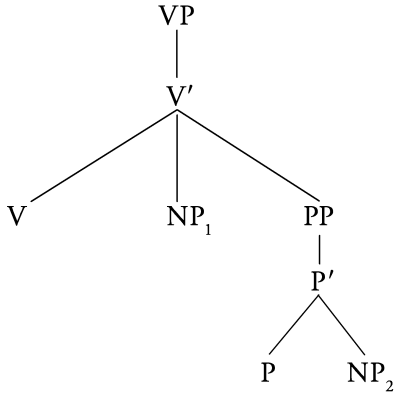
Other tests which also require a c-command relationship between two elements also seem to give the same result. In structures of the form *each ... the other*, the NP containing *each* must c-command the NP containing *the other*. Additionally, so called *negative polarity items* must always be c-commanded by a negative expression. Testing these in a double object environment, we have (102) and (103):

(102) a. I showed each partner to the other.
b. *I showed the other to each partner.

(103) a. I told nothing to anyone.
b. *I told anything to nobody.

You might think that the fact that the second object is a PP is somehow relevant. Perhaps part of the problem is that *nobody* in (103b) is an object of the preposition *to*, and so will never c-command anything other than the preposition. (The first branching node dominating *nobody* will be P′, which dominates only the preposition and the NP itself.) You could then analyse sentences like (100a) as having a 'flat' structure.

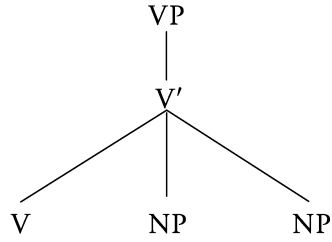
(104)



The direct object is a sister to the verb, it c-commands the second NP, but not vice versa.

However, let's see what impact that has on (100b). Neither the direct object nor the indirect object is introduced by a preposition, so, thinking about what we've seen so far, we might expect that something like (105) would be the right structure, again allowing for triple branching (this was the structure for double object constructions originally proposed by Oehrle (1976)):

(105)



Since the two objects are sisters, they c-command each other, and we predict that, in the double object versions of (101)–(103), both the (a) and (b) sentences should be grammatical.

The problem is that this isn't what we find. Oddly enough, as observed by Barss and Lasnik (1986), we *still* see an asymmetry. It looks like the first object still c-commands the second object, but not vice versa.

(106) a. I showed Mary herself.
b. *I showed herself Mary.

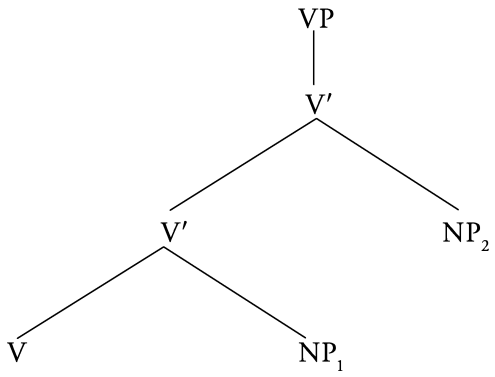
(107) a. I showed each partner the other.
b. *I showed the other each partner.

(108) a. I told nobody anything.
b. *I told anyone nothing.

Therefore, it would seem that the structure in (105) proposed by Oehrle isn't right.

(109) represents another structure which has been proposed for the double object construction. This one comes from Chomsky (1981).

(109)



(109) does have certain virtues, which will become more apparent after the discussion in the next chapter. However, the problem is that this doesn't get you the facts that we've seen in (106)–(108) either. In the tree in (109), the second object (NP_2) asymmetrically c-commands the first one (NP_1). Therefore, we'd predict the opposite pattern of grammaticality judgements. It should be the (a) sentences which are ungrammatical in (106)–(108).

Unfortunately, there's no really appealing way to resolve this problem. It looks like the only way out would be to simply deny that c-command was relevant to these phenomena, and so the sentences in (101)–(103) and (106)–(108) really tell us nothing about the structure whatsoever. That would be really unpleasant, though, as well as running counter to a ton of work that says that c-command *is* relevant.

For the moment, we'll have to leave the problem there. For illustration purposes in the next few chapters, we'll just use Chomsky's structure in (109) on those couple of occasions when we need to draw the tree for a double object structure to illustrate something else. However, Larson (1988) has proposed a really neat way of solving the problem of double objects, which we'll discuss in Chapter 9.

FURTHER READING

The book which probably deals most extensively with phrase structure and constituency is Jackendoff (1977). I think this is a terrific book, with lots of interesting data and arguments. He's assuming an even more complicated system than we are though (four levels, not three), so some of the trees and sentences will need 'translation'. However, it's worth checking out to see things applied to examples beyond the basic data that we've discussed here. As we'll see in the next chapter, this three-level system is referred to in the Chomskyan literature as X′-Theory, and so it's under that heading that you'll find at lot of the discussion about the issues we've been talking about in this chapter, like constituency and the three-level system of structure. Chapter 1 of Webelhuth (1995a) has a pretty good discussion of X′-Theory. It's not a textbook per se – more like a handbook of Government and Binding Theory, but the core of Chapter 1 is pretty accessible. Of the 'real' textbooks covering Government and Binding Theory, Haegeman (1994) has probably the most extensive discussion of X′-Theory.

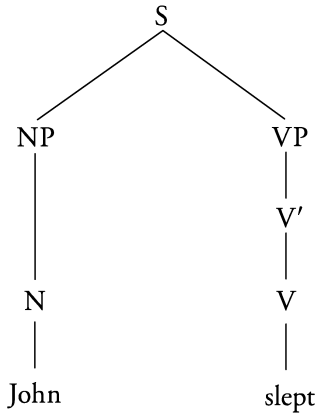
ANSWERS TO IN-TEXT EXERCISES

Exercise 1

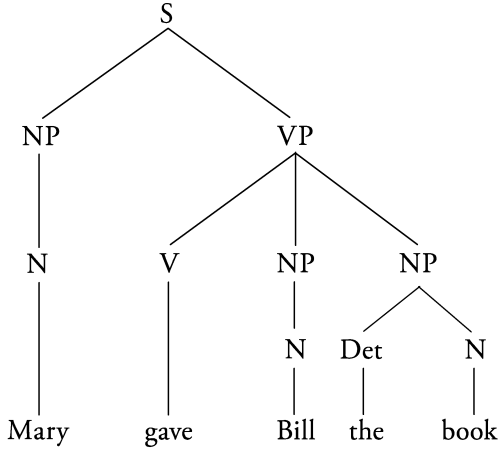
(a) the: determiner; book: noun; about: preposition; linguistics: noun. The whole phrase is an NP, with *book* as its head.
(b) at: preposition; the: determiner; end: noun; of: preposition; the: determiner; platform: noun. The whole phrase is a PP with *at* as its head.
(c) see: verb; the: determiner; answer: noun. The whole phrase is a VP with *see* as its head.
(d) aware: adjective; of: preposition; the: determiner; answer: noun. The whole phrase is an AP with *aware* as its head.

Exercise 2

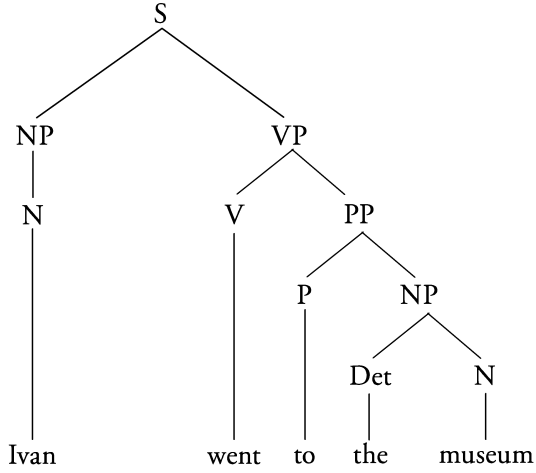
(16a)



(16b)



(16c)



Exercise 3

(a) The NPs are *the new gorilla*, *the new gorilla's attendant*, *a tasty banana* and *the purple table*.

His attendant put a tasty banana on the purple table.
He put a tasty banana on the purple table.
The new gorilla's attendant put it on the purple table.
The new gorilla's attendant put a tasty banana on it.

(b) The NPs are *the person responsible for security*, *the visiting dignitaries*, and *a minuscule antechamber.*

He left the visiting dignitaries in a miniscule antechamber.
The person responsible for security left them in a miniscule antechamber.
The person responsible for security left them in it.

Exercise 4

(a) The PPs are *in Sue's jacket* and *on Thursday.*

Mary put the book there on Thursday.
Mary put the book in Sue's jacket then.

(b) The PPs are *on the table*, *at lunchtime* and *on the very day he was hired.*

The agent who left the computer there at lunchtime was dismissed on the very day he was hired.

The agent who left the computer on the table then was dismissed on the very day he was hired.

The agent who left the computer on the table at lunchtime was dismissed then.

Exercise 5

There are several possible constituents one can identify:

John's predilection for original thinking

[John's predilection for original thinking and his reckless disregard for his safety] placed his family in dire peril.

original thinking

John's predilection for [original thinking and direct action] placed his family in dire peril.

his family

John's predilection for original thinking placed [his family and his many friends] in dire peril.

dire peril

John's predilection for original thinking placed his family in [dire peril and mortal danger].

Exercise 6

There are several possible constituents one can identify:

her friends

Her friends, Mary can see on Thursday in Birmingham, but her enemies, she can see anytime.

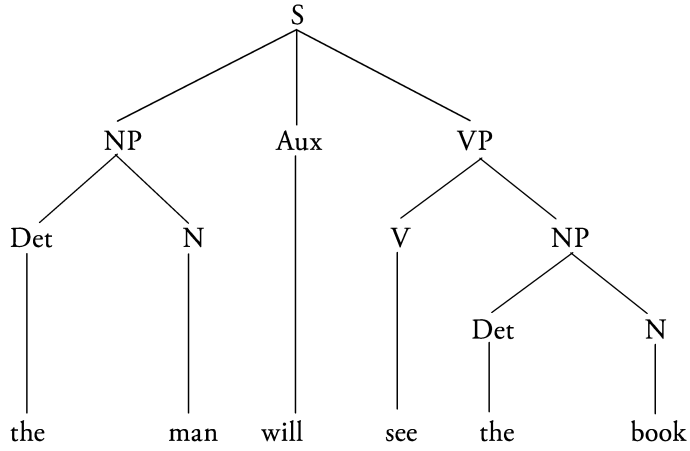
on Thursday

On Thursday, Mary can see her friends in Birmingham, but on Friday, she has to go to the conference.

in Birmingham

In Birmingham, Mary can see her friends on Thursday, but in Newcastle, she has to see her business partners.

Exercise 7



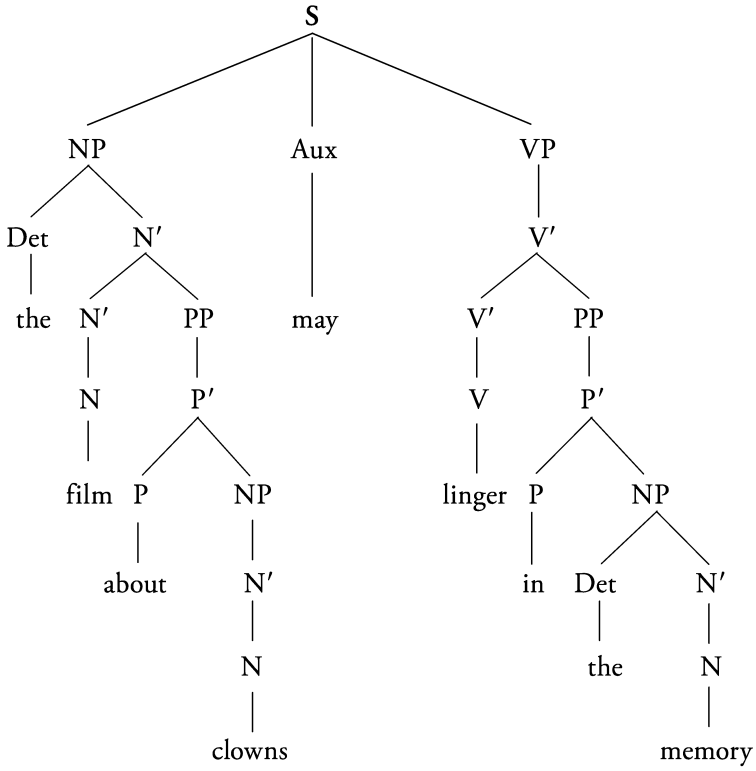
There are several pairs of nodes which are sisters: {NP, Aux}, {Aux, VP}, the two instances of {Det, N} and {V, NP}

Aux c-commands: {NP, Det, N, VP, V, NP, Det, N}

FURTHER EXERCISES

Exercise 1

Consider the following phrase-structure tree:



(a) Give two nodes which are immediately dominated by S.
(b) Give two others which are dominated, but not immediately dominated, by S.
(c) What nodes does V c-command?
(d) What is the sister of the N′ *film*?
(e) What is in the specifier of NP?
(f) What is the complement of P?

Exercise 2

For each of the following XPs, identify the head, complement, adjuncts and specifier (if any). Draw them using the phrase-structure rules of this chapter.

1. Phil's album of 40 big-band hits
2. quite cognizant of the dangers to life and limb
3. over by the temple to the frog-god
4. receive the package from Toledo
5. books on the hall table
6. John might see the picture
7. right on the ceiling
8. rather disappointed with the outcome

Exercise 3

Draw the trees for the following sentences:

1. The elderly clerk can issue the ticket.
2. The player from London will see an agent.
3. John has taken a nap.
4. I liked the book about linguistics in the market.
5. I liked the book about linguistics in the twentieth century.
6. The test of time will reveal the true winners.
7. A green shirt is a big advantage in certain situations.

Exercise 4

We've seen in Section 2.5 how the 'one-substitution' test can be used to give arguments for a three-level system. In particular, we saw that *of* PPs are sisters to N and not sisters to N'.

(1) *Jack met the student of philosophy, and I met the one of Greek.
However, consider the following sentence:
(2) Bill has two quarts of wine and one of water.

Does this represent a problem for the theory we've proposed? Does the following contrast have any relevance to this issue?

(3) a. I met the ones from France.
b. *The quarts of wine and the ones of water were left behind.

X′-Theory and Functional Categories

3

In the previous chapter, we talked about the fact that, as a native speaker of a language, you know that sentences are not merely 'beads on a string' – with one word coming right after the other with no larger organization beyond that. Instead, what you know is that the elements of a sentence are organized hierarchically into larger groups, which we called 'constituents'. In this chapter, we'll talk about a specific way of expressing this knowledge, called *X′-Theory*, and look at some of its implications. We'll also introduce some new syntactic categories, called *functional* categories (as opposed to *lexical* categories like N and V) and examine the role that they play in expressing your knowledge about your language.

3.1 Introduction to X′-Theory

So far we've done pretty well getting our phrase-structure rules in order, but you may have noticed a problem creeping into things, particularly when we introduced the X′ level. The problem is redundancy. The changes we made to the NP phrase-structure rules have a lot in common with the modifications that we later made to the VP phrase-structure rules.

(1) a. NP → (Det) N′
b. N′ → (AP) N′ (PP)
c. N′ → (AP) N (PP)

(2) a. VP → V′ (XP)
b. V′ → V′ (XP)
c. V′ → V (XP)

This is not just a problem with VP and NP. It's a general problem with phrase-structure rules of the sort that we've been looking at so far. For example, every NP that we've seen so far has a head N. Every AP has had an A head. Every PP has had a P head, and so on. We've never seen cases where a VP has an A head,

for example, nor do there seem to be any cases like that in languages of the world. This might seem like a completely trivial point. I can hear you saying, 'Well, it's obvious. You couldn't have a verb phrase if it didn't have a verb in it somewhere!' However, it's important to be surprised by 'obvious' things and to distinguish what's obvious from what the theory tells you should be possible.

If the cognitive system that you possess as a native speaker of a language can be correctly described as having rules of the form V′ → V (XP), why aren't there rules of the form V′ → A (XP)? Or perhaps NP → P′ (XP)? From a formalism point of view, there's nothing to distinguish rules like (2) and (1) from countless other possible rules which we never find. And our theory at this point doesn't explain why.

It also looks as though our phrase-structure rules are missing a generalization. All NPs immediately dominate an N′. All VPs immediately dominate a V′. All PPs immediately dominate a P′. There's a pattern here.

As first suggested by Chomsky (1970) (a paper termed 'monumentally important' by Webelhuth (1995b: 18)), we might be able to eliminate the redundancy and increase our explanatory power by generalizing the phrase-structure rules. Put another way, Chomsky proposed that what you have as part of your cognitive system is a *general* format for phrase-structure rules which are independent of any syntactic category (N, A, V, etc.). So, instead of having rules like (1) and (2), we have the general schema in (3). This approach is known as X′-Theory (where X, Y and Z are just variables standing for any syntactic category):

(3) a. XP → (Spec) X′
    b. X′ → X′ (YP)
    c. X′ → X (ZP)

Although the X can stand for N, V, A or P, you must fill in the X′s on both sides of the arrow with the same choice in any given instance of (3a), (3b) or (3c), so VP must dominate a V′, an A′ can only immediately dominate another A′ or an A, and so on. In (3a), 'Spec' is short for 'specifier', a term which we'll define properly in the next section.

If (3) is part of the cognitive system which enables you to speak and understand your language, then we can explain why many potential rules (like VP → A′ PP) aren't found. With the schema in (3), they can't be created. Every phrase, whether it be NP, VP, AP or PP, must by the rules in (3) ultimately end up with an N, V, A or P head.

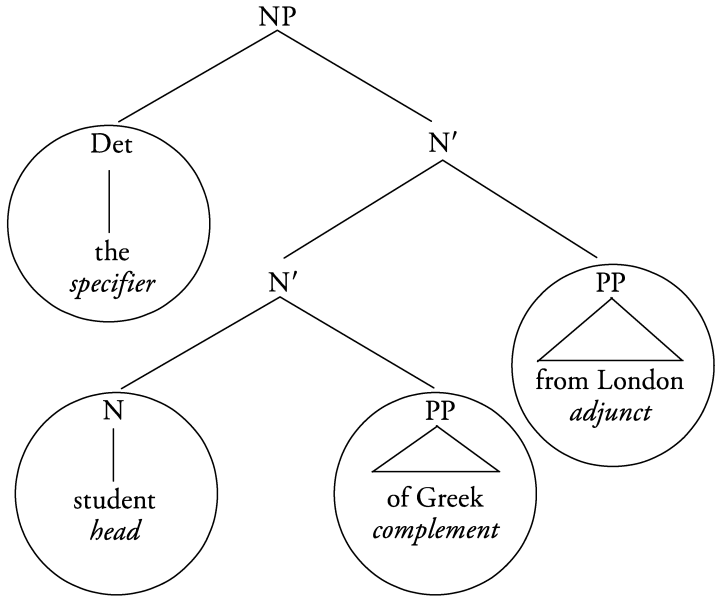
Also, using the general format in (3) eliminates the redundancy among categories. The reason why all VPs immediately dominate a V′ is because they're following the general schema in (3).

3.2 Inside the XP

Having looked at some general terms which are used to describe relations between nodes in a tree in the previous chapter, let's take a look at some configurational terms relating to the structure of phrases now that we've introduced

the X′-system. Again, these terms will become second nature in no time, so don't worry if they seem a little complicated at first.

(4)



Let's start at the bottom. The most important part of the phrase is the *head*, which we've mentioned before in discussing the two-level system. Within the X′-system, it's common to see the head X of an XP written as 'X°', which is read as 'X-zero'. (The idea is that the single-bar level is level 1, so one level below that is level zero.)

Another very common configurational term when discussing XPs is the *complement*. Looking at rule (3c) of the general X′ schema, the complement is the optional sister ZP to the head X°. A complement is always defined relative to a head. So in the NP that we drew the tree for in (4), the complement is the PP *of Greek*. That PP is the sister to the head N° of the NP. The complement often corresponds to what would traditionally be called the 'object' of the head, particularly with verbs and prepositions.

Not all heads have complements, which is why the complement is optional in the X′-schema rule in (3c). However, when they do, it's often because they are obligatory for that head. Consider the two different VPs in (5):

(5) a. John slept
b. John destroyed the town

The verb *slept* in (5a) does not have a complement, but the verb *destroyed* does in (5b). The complement is its direct object, and it is obligatory.

(6) *John destroyed

EXERCISE 1

Identify the head and the complement of the head within the following XPs:

(a) jealous of Federico's achievements
(b) into the fray
(c) buy the record

Working our way upward in the diagram in (4), we distinguish complements from *adjuncts*. Whereas complements, when they appear, are usually obligatory, adjuncts are optional. Adjuncts are immediately dominated by X′ and a sister to X′, and so are introduced by rule (3b) above. Being immediately dominated by X′ and a sister to X′ they can be easily identified using the substitution tests discussed in Chapter 2. Consider the VP in (7):

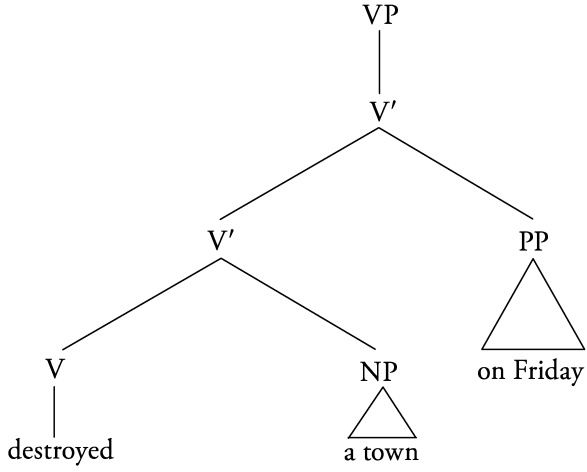
(7) destroyed a town on Friday

Using the *do so* test, we see that *destroyed a town* is a V′ and that *destroyed a town on Friday* is a V′.

(8) John destroyed a town on Friday and...
Godzilla did so on Saturday.
Godzilla did so as well/so did Godzilla.

Therefore the tree for the VP in (7) is (9).

(9)



Of the two NPs, then, *the town* is a complement, as it is the sister to the head V of the VP, and the PP *on Friday* is an adjunct. And they would seem to accord with the characterizations given above. The complement of *destroy* is obligatory. It's part of the essential meaning of *destroy*. (We'll talk about this more in the next chapter.) *On Friday*, by contrast, is optional. It simply gives additional

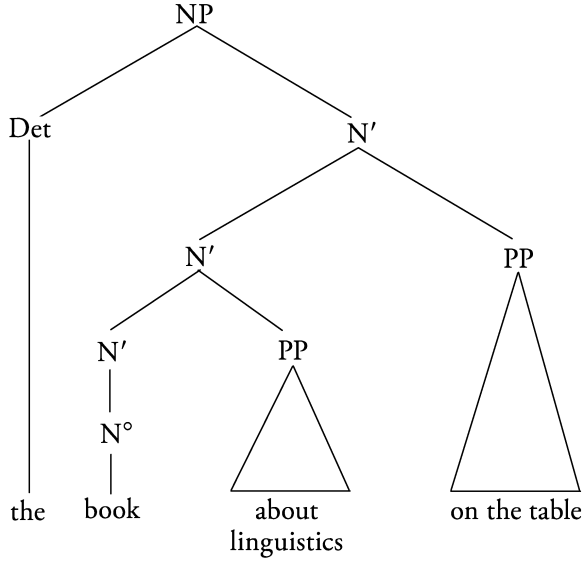
information about when the destroying the town action took place. In the tree for the NP in (4), the PP *from London* is an adjunct.

The fourth and final important configurational term within the XP is the *specifier*. Semantically, it often gives more information about the head or makes it more specific, but for our purposes it is simply the element which is immediately dominated by XP and is a sister to X′ which we saw in rule (3a) of the X′ schema. So in the tree for the NP in (4), the determiner *the* is the specifier of the NP.

It's important that you remember the 'immediately dominated by XP' part of the definition though, because that's what distinguishes specifiers from adjuncts. Both adjuncts and specifiers are sisters to an X′, but only the specifier is immediately dominated by XP. Consider (10), with the tree in (11):

(10) the book about linguistics on the table

(11)



Notice that both the determiner *the* and the PP *on the table* are sisters to an N′. However, only one of them is the sister to an N′ which is immediately dominated by NP. That's the determiner. Therefore, *the* is the specifier of the NP in (10). Both the PPs *about linguistics* and *on the table* are adjuncts.

EXERCISE 2

Identify the specifiers in the following XPs. Which of the various PPs are adjuncts?

(a) several recordings of whales on cassette
(b) three miles down the road
(c) rather appalled by the situation in New Jersey

3.3 Functional categories

In the previous sections we gave arguments for adopting the X′-system in (12) as a description of what you know about your language:

(12) a. XP → (Spec) X′
b. X′ → X′ (YP)
c. X′ → X (ZP)

However, in doing so, we focused on phrases like NPs and VPs. The majority of the categories that we've seen so far, specifically N, V, A and P, are usually referred to as *lexical* categories. These contrast with the so-called *functional* categories Inflection (I), Complementizer (C) and Deteterminer (D). In this section, we're going to argue that functional categories, just like their lexical counterparts, also have structures fully consistent with the X′ schema just outlined. We'll also argue that X′-Theory is just as valid for seemingly more exotic languages like Turkish and Japanese.

From Aux and S to IP

When we were discussing X′-Theory earlier on in the chapter, you may have been wondering what happened to the following phrase-structure rule:

(13) S → NP Aux VP

This phrase-structure rule really looks like a problem. If we're right to say that (3) is the general format that all phrase-structure rules have to follow, then we'll have to change (13). It doesn't fit the X′ format at all. Within the X′ system, there's always a one-to-one relationship between heads and phrases. Every NP has a head N° in it and every head N° is the head of some NP and so on. However, in (13), we've got an S on the left side, but it doesn't appear to dominate anything of category S. On the right side, all we've got are NP, Aux and VP. Going the other way, the Aux category would seem to be a head also, but there's no Aux′ or AuxP anywhere to be found.

We can kill two birds with one stone if we say that what's really going on is that S is essentially an 'Aux-phrase' which has Aux as its head. Also, just to make a terminological change to bring us into line with the way linguists talk, we'll also need to change the name of the Aux node to 'Inflection', which is usually abbreviated to 'I' or 'Infl'. With the addition of an intermediate I′ category, we can replace (13) with something much neater and consistent with X′-Theory:

(14) a. IP → NP I′
b. I′ → I° VP

The subject of the sentence is now the specifier of IP (Inflection Phrase) and the VP is the complement of I°.

Although there is no element which can act as a proform for I′, we can use the coordination test to show that IP needs three levels of structure.

(15) John will read the book and may agree with the conclusions.

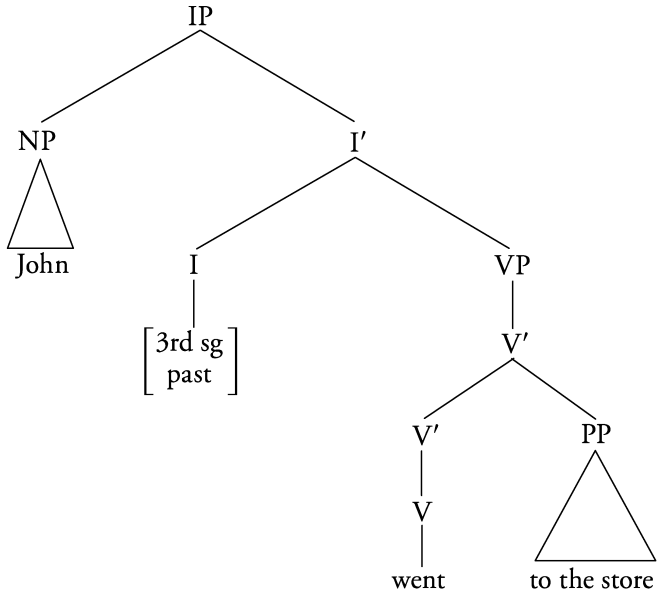
What we have in (15) is a conjunction that includes the auxiliary and the VP, but not the subject. That suggests that these elements form a constituent to the exclusion of the subject. However, the phrase-structure rule in (13) claims that these elements do *not* form a constituent. For that reason, (13) seems to be an incorrect characterization of what you know when you know English. On the other hand, the X′-consistent phrase-structure rules in (14) *do* predict that the auxiliary (in I°) and the VP *do* form a constituent, namely I′.

EXERCISE 3

Using IP, draw the tree for the following sentence:
Mary will read several newspapers.

In those instances where there is no auxiliary, we'll assume that the I° head of IP contains two pieces of information: the tense of the sentence (e.g. past or present, since future tense is indicated by the auxiliary *will*) and the subject-verb agreement information, as in (16). It is because tense and subject-verb agreement are marked by inflectional endings on the verb that the whole sentence is called 'IP'.

(16) John went to the store.



To see concretely why I° needs to have tense and agreement information, let's examine a construction called 'VP-ellipsis'. Under certain circumstances, when two sentences are conjoined, the second VP can be deleted:

(17) Mary can't go to the store but John can.

In (17), the VP in the second conjunct is identical to the VP in the first conjunct (*go to the store*), but it remains unpronounced, leaving behind the auxiliary *can*. But consider what happens when there is no overt auxiliary:

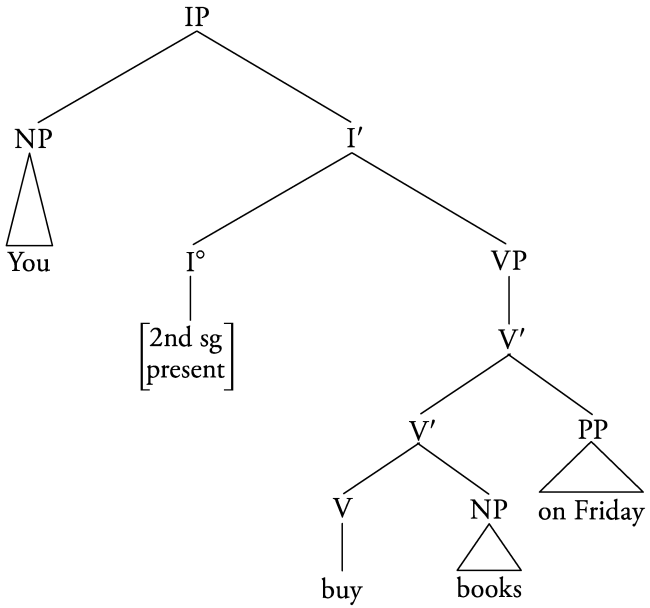
(18) I buy books every day but John doesn't.

Because there is no auxiliary, the dummy verb *do* must be inserted in the second conjunct. However, notice that the dummy verb shows up in specifically the present tense (*doesn't* vs *didn't*) and third person singular form (*doesn't* vs *don't*). Therefore, it must be the case that the tense and subject-verb agreement information are (inflectionally) present in I° even when there is no auxiliary to 'spell them out'.

In order to remember this, when drawing a phrase-structure tree with no auxiliary, be sure to indicate the tense and agreement information in I°. Thus the sentence in (19) has the tree in (20):

(19) You buy books on Friday.

(20)

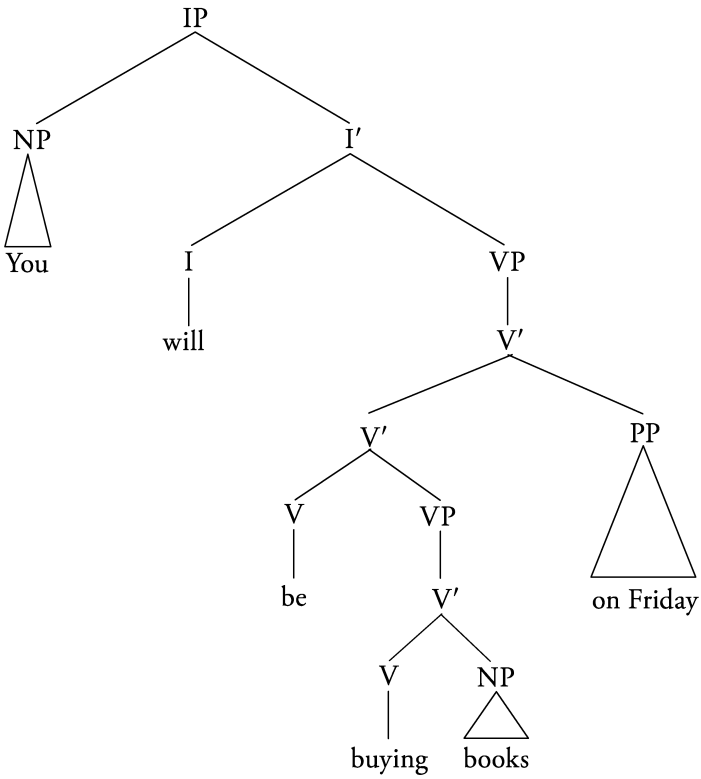


In addition to the cases where there is no auxiliary, we also need to say something about cases where we have more than one auxiliary, as in (21):

(21) You will be buying books on Friday.

We'll analyse these cases using what are called 'VP-shells'. We'll see another use for VP-shells in Chapter 9, but for the moment we'll keep the focus on how they provide a structure for sentences with multiple auxiliaries. The first auxiliary is in I°, but any other auxiliaries that you might have are within the VP, as a series of V°s, each of which takes a VP as its complement, until you get to the main verb:

(22)



The first auxiliary *will* is in I°, but the second auxiliary *be* is a verb, which takes a VP headed by *buying* as its complement. If we had had another auxiliary, say *will have been buying*, then *will* would be in I°, with *have* as a verb taking a VP headed by *been* as its complement. *Been* would then in turn have taken a VP headed by *buying* as its complement.

VP-ellipsis again suggests that this analysis is on the right track. Consider (23):

(23) You have been working hard, but John hasn't.

This suggests that *be working hard* is a VP, since it can be elided under VP-ellipsis. *Working hard* is also a VP, as (24) is also possible:

(24) You have been working hard, but John hasn't been.

Complementizers and CP

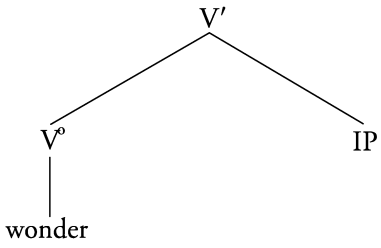
Consider the following sentences in light of the discussion in the previous section:

(25) Mary will meet her friend at the station.

(26) I was wondering whether *Mary will meet her friend at the station*.

Clearly (25) is a sentence, so it must be an IP. It seems equally clear that the italicized string in (26) is also an IP. The question is what to do about *whether*. Clearly, we can't just say that the V° *wonder* takes an IP complement, as in (27):

(27)



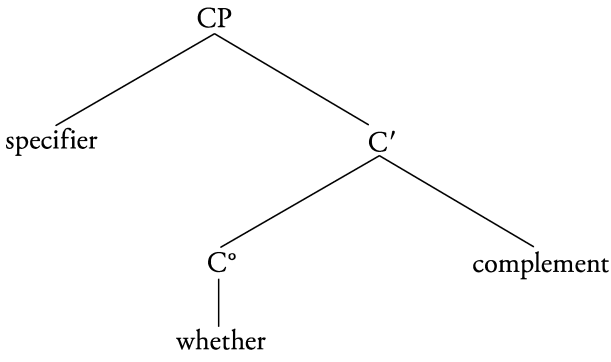
That would predict that (28) should be grammatical, but it isn't:

(28) *I wonder Mary will meet her friend at the station.

In order to account for (26), while not allowing (28), what we need to say is that the IP *Mary will meet her friend at the station* combines with some other category, whose head is *whether*.

Elements like *whether* or *that* (as in *I know that Bill is here*) are traditionally referred to as *complementizers*. They get that name because they are added to the beginning of an IP in order to help the IP fit in as the complement of a verb. If the complementizer C (*whether* in this case) is the head of CP, then X′-Theory tells us that it must have a specifier and a complement in a three-level structure:

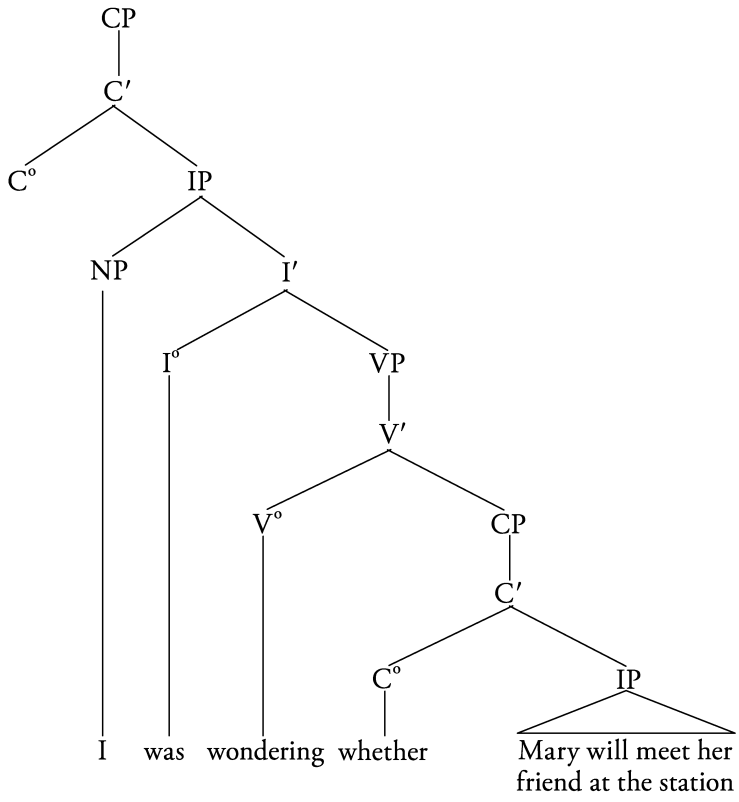
(29)



In this case, the IP *Mary will meet her friend at the station* is the complement of C. It has no specifier, although we'll come to some examples later which do.

Putting all of this together, we would draw the tree for (26) as (30):

(30)



You may be wondering why we've started off with a CP in (30). It's certainly true that sentences don't start with complementizers:

(31) *That I was wondering whether Mary will meet her friend at the station.

However, this seems to be an idiosyncratic feature of English. There are many other languages in which main clauses can start with complementizers. In Spanish, just to pick one example, main clauses can begin with a complementizer when the sentence expresses a wish (32), or when something is being repeated or emphasized (33).

(32) Que se calle Juan
that self silence-3sg Juan
'Let Juan keep silent.'
(Demonte and Fernández-Soriano 2009: (4c))

(33) Que no quiero verla
that not want-1sg to see her/it
'I don't want to see her/it.'

(Butt and Benjamin 2000: 448)

We haven't really encountered any non-English data before, so a quick word is in order regarding how to read the examples. The first line is the sentence in the original language, converted to a Roman alphabetic script if necessary (as with, for example, Japanese or Russian). The second line is a morpheme-by-morpheme 'translation'. For example in (32), the Spanish verb *calle* contains the root 'to silence' and a third person singular, present tense marker. The third line is a translation of the original example sentence into grammatical English.

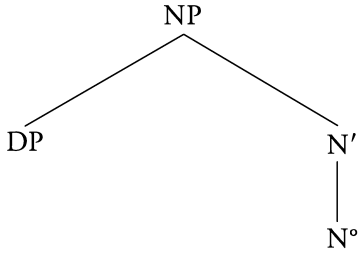
Although we're focusing on English (since this book is aimed at speakers of English), we're assuming that the basic structure of the sentence is provided by UG and so therefore should be the same in all languages. So we don't want to be misled by idiosyncracies of English. (This is going to come up again later on in this chapter, and in an even more important way in the next chapter, when we talk about the concept of 'Case'.) We'll see in later chapters anyway that there are some English sentences which do start with CP, so, really, the idiosyncratic nature of English here is in fact only apparent. From here on, we'll be assuming that all English sentence trees start with CP.

Getting into Det: The DP hypothesis

In the previous sections we've been trying to tighten up our phrase-structure assumptions a bit to bring them into line with the X′ system, eliminating heads like Aux that had no associated phrases and eliminating 'phrases' like S which didn't seem to immediately dominate an element of the same category.

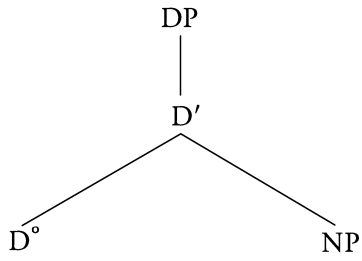
Continuing this exercise, let's examine the element Det, which is still outside the X′ system. Determiners are clearly heads, not phrases, but we've just been sticking them in the specifier of NP position without saying anything more about them. In this section, we'll examine some arguments that Dets, in fact, have associated D′s and determiner phrases (DPs). However, it will turn out that it's not as simple as sticking a DP in the specifier position, as in (34):

(34)



There have been various arguments given in the literature that what we've been calling *noun* phrases are actually *determiner* phrases. That is, the structure of *the book* actually has a determiner head which takes an NP as its complement:

(35)



Abney (1987) pulls together some observations regarding gerundive 'Poss-ing' constructions in English to argue for the existence of determiner phrases. An example of the gerundive 'Poss-ing' construction in English is (36):

(36) [John's building a spaceship] caused a huge commotion.

There's a simple question about phrases like *John's building a spaceship* which turns out not to have an obvious answer: what syntactic category is it? Is it an NP? Or is it something sentential, like an IP or a CP?

On the one hand, (36) as a whole looks like it must be an NP as it can turn up in NP positions. In (37), we see that a 'Poss-ing' phrase can be a subject, and in (38) the object of a preposition, which is impossible for a CP:

(37) a. I wondered if [John's building a spaceship] had upset you.
b. I wondered if [$_{NP}$ the book] had upset you.

(38) a. I told you about [John's building a spaceship].
b. I told you about [$_{NP}$ the book].
c. *I told you about [$_{CP}$ that John built a spaceship].

However, it also seems clear that *building a spaceship* represents a VP. In particular, the '-ing' element of the gerundive 'Poss-ing' is not a noun which has been derived from a verb, like *destruction* or *referral* (from *destroy* and *refer* respectively). Instead it has all the characteristics of a verb. It takes direct objects, for example, without requiring any kind of preposition, such as *of*.

(39) a. John destroyed the spaceship.
b. John's destroying the spaceship (caused a panic).
c. *John's destruction the spaceship (caused a panic).
d. John's destruction of the spaceship (caused a panic).

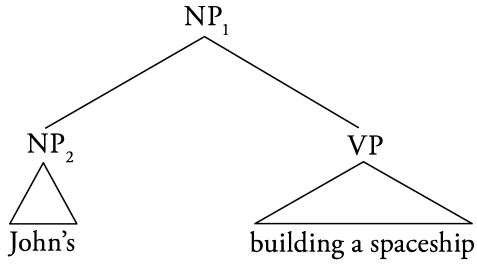
The '-ing' element also takes so-called 'verbal particles', which can shift position just as they would with a 'real' verb:

(40) a. John explained away the problem/John explained the problem away.
b. John's explaining away the problem/John's explaining the problem away (didn't satisfy his critics).
c. *John's explanation away of the problem/*John's explanation of the problem away...

What we seem to be stuck with are two pieces of a puzzle that don't fit together.

If we just bolt together what we've got so far in a Frankenstein's monster-like fashion, it's this:

(41)



Unfortunately, this is not a structure which is allowed by X′-Theory.

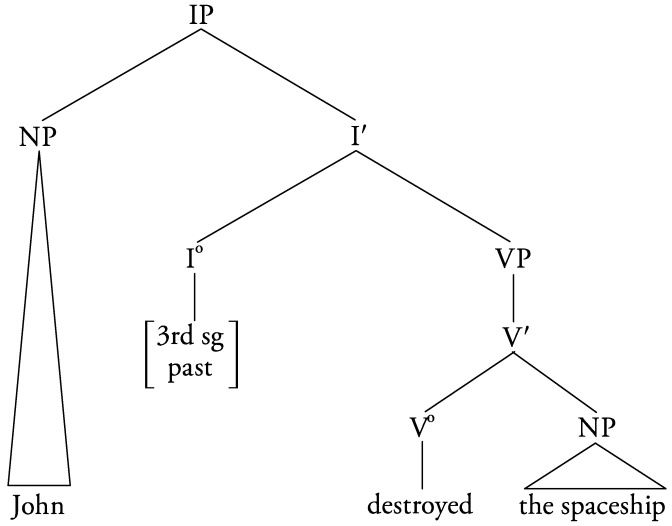
BEFORE CONTINUING

Explain why (41) is not allowed by X′-Theory.

The problem is that NP_1 is a noun phrase, but it fails to immediately dominate an N′. NP_1 doesn't seem to have an N° head either.

However, perhaps the sentence-like nature of 'Poss-ing' gerundives can offer a clue as to what to do. Consider the tree for (39a):

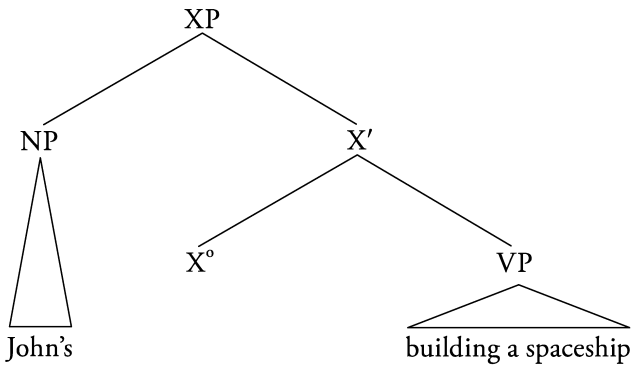
(42)



Here, we have a subject *John* and a VP *destroyed the spaceship*. A mediating role is played by the functional category I. It 'hooks up' the VP with its subject. It

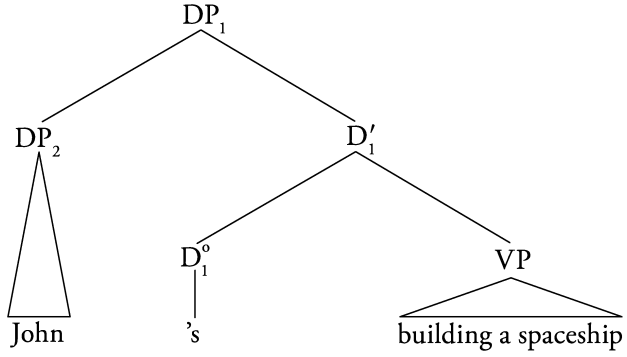
allows the two to get together. If we assume that there's some mediating category in the gerundive 'Poss-ing' construction, that would provide us a way out of the X′-Theoretic problem that we had in (41):

(43)



The only question now is 'What is "X", our noun phrase correlate of I?'. Well, there aren't really too many options. It's at least possible that X is some category which never has any overt words, but that would be an unpleasant thing to have to say. Even the other functional categories like I and C have at least *some* actual words which belong to them (modals like *can* and *will* for I and complementizers like *that*, *whether* and *if* for C), and it's not completely clear how a child learning English would even figure out that this category existed if there were no overt words of that category. However, it really looks like everything falls together nicely if we assume that the mystery category is Det, the determiner, and what we've previously thought of as NPs are actually DPs.

(44)



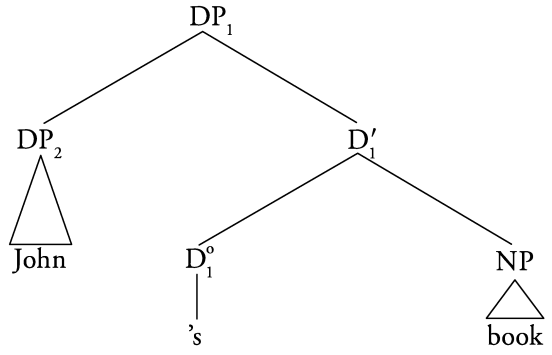
More on DPs: possessives, proper names and pronouns

In this section, I want to tie up a few potentially loose ends concerning this re-analysis of NPs as DPs. We'll start with a slightly counter-intuitive analysis of possessive DPs like *our book*. We'll then talk about what the implications of the DP hypothesis are for the analysis of bare plurals, proper names and pronouns.

Possessive DPs

Given the discussion in the previous section concerning the 'Poss-ing' construction, you can probably guess what our analysis of possessive DPs like *John's book* is going to be. We want to emphasize the parallelism between *John's book* and *John's building a spaceship*, so the structure for the two of them will be essentially identical. In *John's building a spaceship* (44), *John* is the specifier of the DP, the head D of the DP contains the possessive marker *'s*, and the complement of D is the VP *building a spaceship*. The analysis of *John's book* is identical, only with NP, instead of VP, as the complement of the determiner *'s*.

(45)



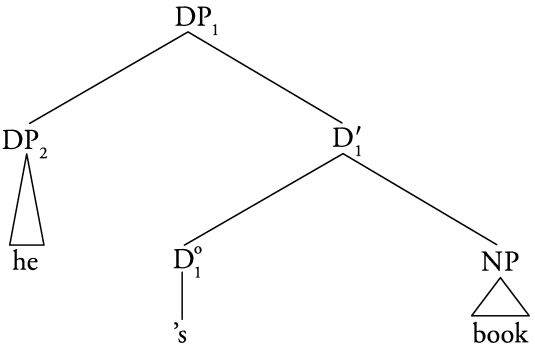
This leads, though, to an interesting, and maybe slightly counter-intuitive, analysis of DPs in which the possessor is not itself a DP, like *John* or *the student*, but instead is a pronoun, as in (46):

(46) his book

The only real difference between *John's book* and *his book* seems to be that the former has a more specific third person singular masculine entity as its possessor (*John* vs *he*). Yet, at least on the surface, there's a possessive *'s* in the former but not the latter.

What we'll suggest is that *John's book* and *his book* are actually more parallel than the English morphology might suggest, and that the tree for (46) is actually (47), fully parallel to the tree in (45) above:

(47)



Just as in (45), the DP *his book* has a specifier. But rather than *John* it is the third person singular pronoun *he*. The head D of the DP contains the possessive marker *'s*, and it takes the NP *book* as its complement. In other words, as far as the syntax is concerned, *his book* is 'really' *he's book*. *He's book* is realized as *his book* because of a low-level morpho-phonological rule of English which says that *he's* is 'spelled out' as *his*. In a similar way, *our book* is 'really' *we's book*, *their book* is 'really' *they's book*, etc.

For the moment, we'll have to leave only the conceptual argument based on simplicity, parallelism and theory-neatness considerations as a reason for assuming (47). To be fair, that's not nothing. The alternative is one in which *John's book* and *his book* have rather different structures. If, for example, we were to make *his* a determiner, then one DP would have a specifier and the other wouldn't. If, on the other hand, *his* is a specifier, it's not a DP, since *his* doesn't occur in places where other DPs do (like subject, object of a preposition, etc.).

Furthermore, the 'low-level morpho-phonological rule of English' that we have to assume which changes *he's* into *his* doesn't seem too unreasonable either. The possessive forms *his*, *our* and *their*, like many other pronoun forms, are clearly the reflection of an earlier stage of English, and are something speakers of Modern English simply have to memorize. We will, however, come back to the question in Chapter 5 and see that Binding Theory provides an *empirical* argument for the analysis that we've proposed. So if you're unhappy about (47), hang on and you might be a little happier later.

Proper names and bare plurals

A sort of theme is beginning to emerge about ways in which what we see on the surface can be slightly misleading when it comes to the analysis of certain phenomena. In this section, that theme will continue when we turn to proper names and bare plurals. The question is what to do with elements like *Alice* or *tigers*. They clearly have to be DPs because they appear in typical DP positions:

(48) a. I like Alice/tigers.
b. Bill left all his money to Alice/tigers.

Nonetheless, they do not have any overt determiner, and, in the case of proper names in English, *cannot* have an overt determiner:

(49) *I like the Alice.

However, the interpretation of bare plurals does seem to include something determiner-like. (50a), for example, is paraphrasable as (50b):

(50) a. Cows eat grass.
b. All cows eat grass.

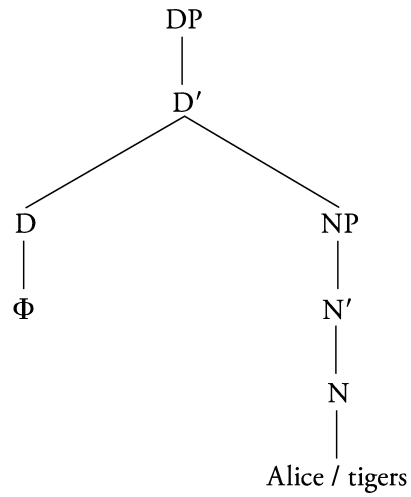
Neither is the paraphrase limited to *all*: other contexts appear to allow other paraphrases with other determiners, like *most* or *some*...

(51) Cats are smaller than dogs/Most cats are smaller than most dogs.

(52) People approached the UFO cautiously/Some people approached the UFO cautiously(, but a few stayed back).

Furthermore, in the case of proper names, other languages *do* allow them to be used with (definite) determiners. Southern varieties of German, Ancient and Modern Greek, and certain non-prestige dialects of Castilian Spanish are all examples. We're assuming that the determiner system, and functional categories in general, form part of UG. So, bearing in mind the discussion about description and explanation back in Chapter 1, we want to check a wide range of data to make sure that we're not being misled. Therefore, we will assume that bare plurals and proper names are DPs which contain a null determiner, as illustrated in (53).

(53)



Pronouns

Moving on from proper names, you might assume that we would analyse pronouns in the same way. (We haven't been explicit about their structure so far, indicating them merely as a DP with a triangle.) However, it looks as though pronouns have some important differences. First, pronouns can occur together with Ns in a way that proper names can't:

(54) We linguists have got to stick together.

These cannot themselves occur with a determiner, so can't, for example, be analysed as compound nouns:

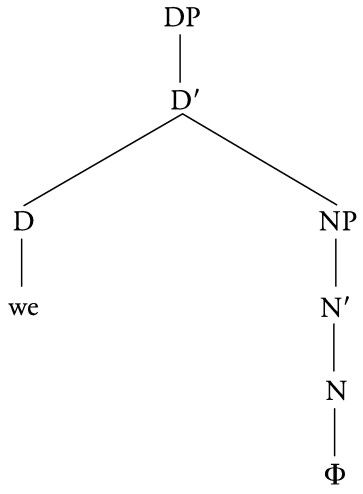
(55) a. the rose garden
b. *the we linguists

Also, pronouns can occur together with adjectives and nouns:

(56) You blue aliens need to fill in these forms.

This suggests that pronouns are actually the reverse of proper names. Rather than being DPs which contain a null determiner, the pronouns themselves are Ds (which take a null NP sister when they appear alone).

(57)



3.4 Coordination and X′-Theory

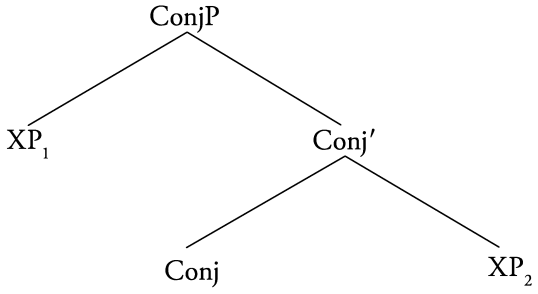
The final functional category which we'll consider in this chapter is conjunctions, as illustrated by (58):

(58) John and Bill went to the store.

In Chapter 2, when we were illustrating the coordination test, we weren't explicit about what phrase-structure trees we were assuming for coordinated structures. Our hypothesis is that X′-Theory is provided by UG and creates a general framework which all phrase structure conforms to. If that's correct, then we would hope to find an X′-consistent way of incorporating sentences like (58).

Various researchers have suggested (59) as a way of meeting our requirements:

(59)



The conjunction *and*, in this case, is the head of an X′-consistent functional category ConjP. It takes the first conjunct as its specifier and the second conjunct as its complement.

Now, it's all very well to say that we can *create* a structure for conjunctions which is X′-consistent, but the real question is whether there's any *evidence* that (59) is the right analysis. In other words, we have a *conceptual* reason for proposing the structure in (59), but we now need to find some *empirical* support for it.

One argument which has been advanced in support of the structure in (59) has to do with subject-verb agreement. There are various examples from languages of the world in which a verb shows subject-verb agreement only with the *first* of the two conjoined subjects. Consider the following example from Czech:

(60) Půjdu tam [já a ty].
go-1sg.future there I and you
'I and you will go there.'

In (60), the verb is marked with first person singular agreement marking. It seems as though it agrees with only the first conjunct *I* and not the second conjunct *you*. We see a similar situation in the following Palestinian Arabic example:

(61) Gatalen [ʔel-banat we-l-wlad] ʔel-bisse.
killed3pl.fem the-girls and-the-boy the cat
'The girls and the boy killed the cat.'

Again, the verb seems to agree with only the first conjunct, which is third person plural feminine.

The structure in (59) provides a natural way to accommodate these facts, through the mechanism of *specifier-head agreement*. There seems to be a special relationship that heads have with elements in their specifiers. Specifier-head agreement is a process by which information can be passed between the specifier of a functional category and the head.

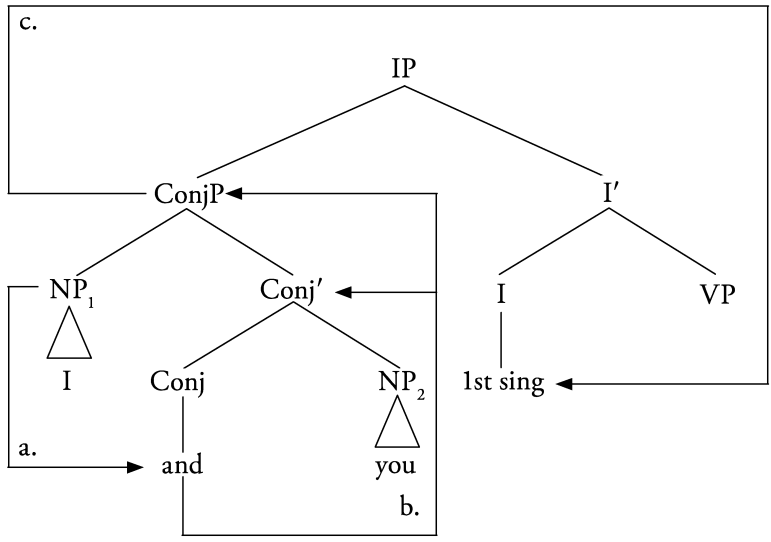
The most common example of specifier-head agreement is subject-verb agreement within IP. As we discussed in Section 3.1, we can see from VP-ellipsis data that the I° head of IP contains the verbal agreement information. When the subject is third person singular in English, for example, the verb takes its characteristic '-s' ending. It is because the subject is in the specifier of IP that I° 'knows' what person and number the subject is.

The structure in (59) explains why we should see the facts in (60) and (61), where the entire conjoined phrase seems to have only the person, number and gender features of the first conjunct. What we have is simply another case of specifier-head agreement. The first conjunct, under the analysis in (59), is the specifier of ConjP. Via specifier-head agreement, it and the Conj head will agree person, number and gender features.

Now, because Conj is the head of ConjP, the features are said to 'percolate' up to ConjP, with the result that the entirety of the conjoined phrase has

only the person, number and gender features of the first conjunct. By specifier-head agreement again, ConjP and I agree in first person singular features. This process is illustrated in (62):

(62)



a. specifier-head agreement between Conj and its specifier
b. feature 'percolation' from Conj to ConjP
c. specifier-head agreement between ConjP and I

Thus, we see at least one piece of support for an X′-consistent approach to conjunction, and we seem to have a pleasing convergence between conceptual and empirical considerations.

3.5 Head-initial and head-final languages

So far, it looks as though X′-Theory does at least a decent job of accounting for the data that we've looked at so far. However, what are we to make of the following data?

(63) Ben Newcastleda otoruyorum. (Turkish)
I Newcastle-in live-lst sg.pres.
'I live in Newcastle.'

(64) Watashi-wa Newcastle-ni sunde-iru. (Japanese)
I-subj Newcastle-in live-present
'I live in Newcastle.'

Even just looking at the PPs in Turkish and Japanese, it looks as though we're going to have to do some serious modification to the X′-system. These languages have what are called *postpositions* as opposed to the *prepositions* which English has. That is to say, the head P° of the PP comes *after* its complement instead of before. And the VP works the same way. The PP *in Newcastle* is itself the complement of the verb *live*, and once again we see the complements of the VP coming *before* the head V°. If you look at the X′-schema as illustrated in (3) at the beginning of this chapter, you'll see that they only allow for prepositions. The complement must *follow* the head. The schema in (3) does not allow for postpositions. But that's what we seem to have here.

How should we respond to this problem? We could simply say that Turkish and Japanese are different and leave it at that. We could then write another set of rules to account for Turkish and Japanese. The X′-system already does a lot of work for us in eliminating redundancies in English. Perhaps there's nothing more to be said about it than that. However, remember that as *linguists* we're interested not just in a theory about English, but also in a general theory of all human languages. In other words, we're interested in investigating UG – in what all languages have in common. Ideally, then, we want a *single* set of very general rules covering all languages.

But how do we go about generalizing over the English rule for prepositions and the Japanese/Turkish rule for postpositions:

(65) a. English: PP → P NP
b. Japanese/Turkish: PP → NP P

What the two rules in (65) have in common is that they both say that a PP consists of a P and an NP. The only difference seems to be the order which is imposed on the two subconstituents. Therefore, all we need to do to generalize over the two is to say that linear order is *not* imposed. This can be indicated by the use of the 'comma' notation:

(66) English/Turkish/Japanese: PP → P, NP

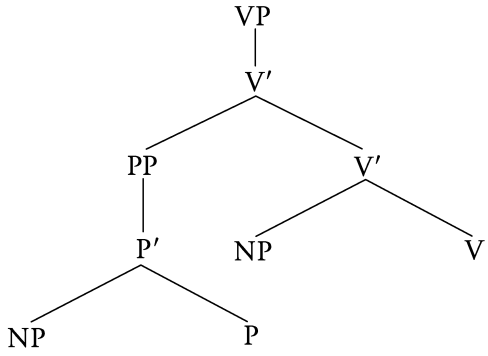
The phrase-structure rule in (66) is read as indicating that a PP is made up of a P and an NP, which can come in either order.

Therefore, if we remove the linear ordering from the X′-schema in (3), we have (67):

(67) a. XP → (Spec), X′
b. X′ → X′, (YP)
c. X′ → X, (ZP)

By doing this, and claiming that (67) is provided by UG, we can account for the Turkish and Japanese examples in (63) and (64) by assuming that their structure represents the 'mirror-image' of English. So if we were to draw the trees for (63) and (64), they'd look like (68):

(68)



Languages like English, where the order of (39c) comes out as (65a), are referred to as *head-initial* languages because the head precedes its complements. Languages like Turkish and Japanese are referred to as *head-final* languages because the head follows its complements.

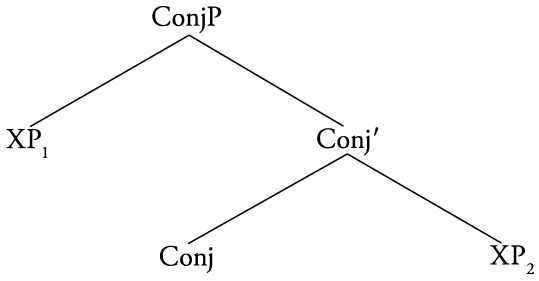
Summary

In this chapter, our goal has been to generalize the phrase-structure rules we discussed in Chapter 2. We introduced X′-Theory as a way of capturing generalizations that we seemed to be missing about the phrase-structure rules we were discovering. We then extended X′ consistency throughout the phrase-structure system. We saw that some categories, like S, Aux and Det, didn't fit easily into the X′-system. If our hypothesis is correct that the X′-system is provided by UG as a general format for phrase-structure rules, then we want to recast any non-X′-consistent categories. We showed how an X′-consistent recasting of S and Aux into IP, Det into DP, and coordination into ConjP might be accomplished. We also saw some empirical evidence that an X′-consistent approach to these categories was the right one. In the final section, we saw that not only is X′-Theory independent of particular syntactic categories (i.e. it describes everything from DPs to PPs to VPs), but it is also independent of linear order. That is to say, the rules in (3) are intended to characterize not only *head-initial* languages, where the head X precedes the complement ZP, but also *head-final* languages, where the head X follows the complement ZP.

Open issue: problems with coordinate structures

In Section 3.4, we proposed the structure for conjunction in (69):

(69)



However, this view, while X′-consistent, and therefore a convenient one to adopt for our purposes, is somewhat controversial, and various authors have noted that certain facts about conjunction are hard to explain under the analysis in Section 3.4.

We noted in Section 3.4, for example, that (69) allows us to account for the fact that agreement is typically with the first conjunct, rather than with the second conjunct, or with some union of their features. Under specifier-head agreement, the grammatical features of XP_1 are shared with the Conj head, which then pass them up to ConjP. However, Borsley (1994) notes that there are problems with this approach, casting doubt on whether we can account for the agreement facts of conjunction using the mechanism of specifier-head agreement. Looking at subject-verb agreement, it is clear that only person, number and possibly gender features are shared via specifier-head agreement. Crucially, however, categorial features are *not* shared. In other words, the DP in the specifier of IP does not share its 'nouniness' with the head I of IP, turning the IP into a DP. If it did, we would have to allow complementizers to take DP complements, which is clearly not possible in general.

(70) a. *I said [$_{CP}$, that [$_{NP}$ the book]].
b. *I asked [$_{CP}$ whether [$_{NP}$ the file]].

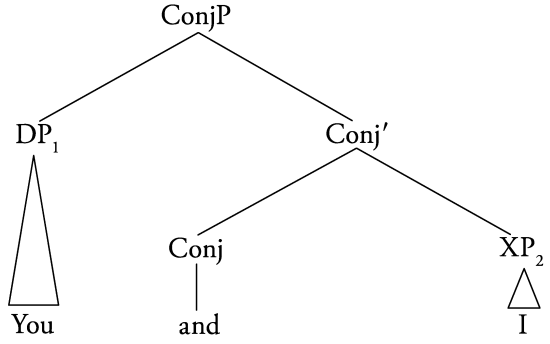
Therefore, it seems as though we must not allow specifier-head agreement to extend to copying the syntactic category of the specifier on to the functional category. However, this sharing of syntactic categorial features seems to be precisely what is needed in the case of conjunction. The conjoined entity has the same syntactic category as the two individual conjuncts.

Furthermore, there are also cases where we would want to say that feature sharing between the first conjunct and the head does not happen. This is exemplified by sentences such as (71) (from Borsley 1994):

(71) You and I understand ourselves/*yourself well.

In (71) it would seem that the subject must be first person plural, in order to account for the appearance of *ourselves* rather than *yourself*. However, under the specifier-head agreement account in Section 3.4, we would expect the subject to be second person singular.

(72)



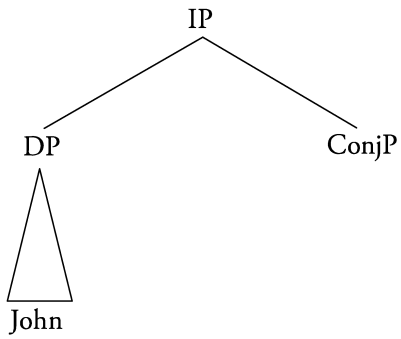
Since *you* and *and* agree via specifier-head agreement, ConjP must also therefore be second person singular, and we predict the opposite pattern of grammaticality to the one we see in (71).

Another problem noted has to do with the bar-level of the conjoined categories. The traditional structure in (69) claims that the conjoined elements together form a maximal projection, not an intermediate X′ projection or a head. However, there are cases where the conjoined constituent appears to be just such a thing. Consider (73) and (75):

(73) John [may like Mary] and [must like Sue].

In Section 3.3, we gave sentences like (73) as evidence for the existence of I′. Both conjuncts contain an auxiliary verb, which means we must have two Infls. However, we have only one subject, *John*, which suggests that we have only one Spec IP. Unfortunately, though, the traditional structure in (69) doesn't give us a way to express (73). By X′-Theory, the daughters of IP must be an XP (the specifier) and I′. IP cannot have as its daughters a DP specifier and ConjP, another maximal projection.

(74)



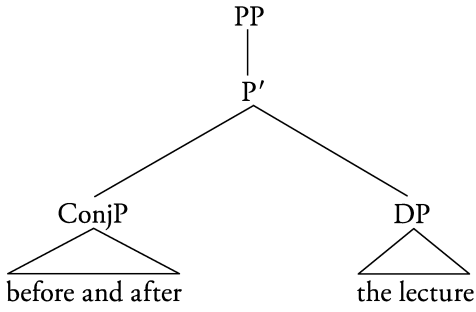
We run into a similar situation in (75):

(75) a. We [handed] or [sent] a copy of the letter to every student.
b. He was there [before] and [after] the lecture.

In this case, however, what we appear to have is conjunction of heads, rather than intermediate projections. In (75a), we have two verbs *handed* and *sent*, but only one set of objects. In (75b), we have two prepositions sharing a single object.

Just as in the previous cases, however, the traditional structure for conjunction in (69) doesn't give us an X′-consistent way to express this. Consider the tree for (75b) for example:

(76)



The structure in (76) seems to be a clear violation of X′-Theory. P′ has two maximal projections as its daughters, ConjP and DP, rather than a head P and a DP complement.

In summary, it looks as though we may need to temper the initial enthusiasm we might have had back in Section 3.4 for the X′-consistent approach to conjunction. However, we will continue to assume it for the rest of the book, keeping in mind the problems that we've seen.

FURTHER READING

Chomsky (1986) suggested the idea that the X′-system could be extended from lexical categories to functional categories, and introduced CP and IP. In the chapter we discussed Abney's arguments for DP in English, which were taken from the introductory chapter of his PhD dissertation (Abney 1987) and some of that introductory material is reasonably accessible. Other early work on DPs was also done by Fukui (1986) and Fukui and Speas (1986). Finally, an excellent overview of the status of ConjP is Borsley (1994). Some of his criticisms and observations about ConjP were discussed in the Open Issue.

ANSWERS TO IN-TEXT EXERCISES

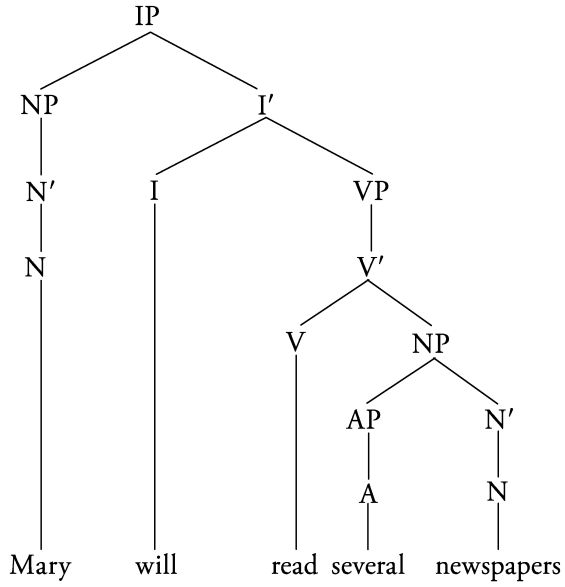
Exercise 1

(a) head: book; complement: about linguistics
(b) head: jealous; complement: of Federico's achievements
(c) head: into; complement: the fray
(d) head: buy; complement: the record

Exercise 2

(a) Brian's
(b) three
(c) rather

Exercise 3



FURTHER EXERCISE

Draw trees for the following sentences using IP, CP, DP and ConjP. (Warning: some may be head-final!)

1. The report by the police will affect the community and its members.
2. John's friend and Bill's mother can send the letter.
3. I know that the cheque will arrive on time.
4. Mary performed the experiment and completed the report.
5. Lions in my garden are a source of irritation.
6. I was wondering whether Bob said that Mitch was leaving.
7. Maria's growing five hundred plants caused an overflow on the patio.
8. Jun-ka chinkwu eykey chayk-ul ilkessta. (Korean)
   Jun friend to book read
   'Jun read the book to a friend.'
9. Juan dijo que María va a Madrid. (Spanish)
   Juan said that Maria is going to Madrid
   'Juan said that Maria is going to Madrid.'
10. Rajakah kanyaya saha vadati. (Sanskrit)
    The laundryman the girl with speaks'
    The laundryman speaks with the girl.'

θ-Theory and Case Theory

4

In the last chapter we looked at how words and phrases are organized into larger groups. In this chapter, we will focus in on words and phrases themselves in more detail. In particular, we'll be looking at the relationship between *predicates* and *arguments*. These are terms from formal logic which have been borrowed into linguistics, but what we're basically talking about are verbs, adjectives and prepositions in the case of predicates, and DPs (and sometimes PPs) in the case of arguments.

By way of illustration, let's take a look at an example of a predicate that you'll be familiar with from mathematics: the 'equals' sign. The predicate has two arguments – that is, two required elements. You have to say what the two things are that are equal to each other. If you leave either one out, the (mathematical) expression is 'ungrammatical':

(1) a. *x =
b. * = y
c. x = y

Verbs, adjectives and prepositions can be thought of as predicates which take arguments in exactly the same way. For example, *destroy* can be thought of as a predicate which also takes two arguments (a destroyer and a destroyee). Leave either one out and the (linguistic) expression is ungrammatical:

(2) a. *John destroyed.
b. *destroyed the building.
c. John destroyed the building.

We'll see that predicates in language have *argument structure*, which indicates what kinds of XP arguments the verbs require in order to complete their meaning. They assign *θ-roles* to these required XPs, which indicate the semantic role that the arguments play in the syntax. (The symbol *θ*, by the way, is the Greek letter theta, so 'θ-roles' is read 'theta-roles'.) The subpart of the theory which

concerns θ-roles and θ-role assignment is called *θ-Theory* ('Theta-Theory'). For those arguments of predicates which are DPs, we'll see that the predicate must assign *case* to the argument. Subjects get nominative case, direct objects get accusative case, etc. It's a traditional idea, and the idea of cases may be familiar to you if you've studied a language like Latin or German. This subpart of the theory, not too surprisingly, is called *Case Theory*.

4.1 An introduction to argument structure

Let's begin the discussion by looking at a woefully inadequate grammar of English (i.e. a model of what you know when you know English). First, we have X′-Theory from Chapter 3. This gives us the general format for our phrase-structure trees:

(3) a. XP → (Spec) X′
b. X′ → X′ (YP)
c. X′ → X (ZP)

However, in addition to what our phrase-structure trees can look like, we also need to have words to put in them. The place in the grammar where words and associated information are stored is called the *lexicon*, and the words themselves are usually referred to as *lexical items*. The lexicon is just the mental equivalent of a dictionary.

What we're interested in here is the information in the lexicon which is relevant for the syntax. Although we didn't talk about it in these terms, we've already seen one piece of information that comes from the lexicon: the syntactic category of a given lexical item. So, for example, part of the lexical entry for *sincerity* will indicate that it's a noun, as opposed to a verb or an adjective.

To account for what you know about English, we need the lexicon to specify more than just the syntactic category though. For example, *see* and *put* are both verbs. However, there is more to specify than that. Consider the contrast in (4):

(4) a. John saw the book.
b. *John put the book.

If all you knew about *see* and *put* was that they were both verbs, we wouldn't be able to explain why you know that (4a) is grammatical and (4b) is ungrammatical.

What we need is some indication about the kinds of complements that a verb can have. *See*, for example, is able to take either a DP complement or a CP complement:

(5) a. John saw the book.
b. John saw that the government was corrupt.

Other verbs, like *sleep*, must not have a complement of any kind:

(6) a. John slept.
b. *John slept the bed.
c. *John slept that Bill left.

However, one thing that *sleep* and *see* do have in common is that they need a subject:

(7) a. John saw the book.
b. John slept.

(8) a. *Saw the book.
b. *Slept.

So a verb like *sleep* is said to be a *one-argument* verb. It requires a subject, but may not have a direct object of any kind. A verb like *see* on the other hand is a *two-argument* verb, with the second argument being either a DP or a CP. There are even *three-argument* verbs, like *put*:

(9) a. John put the book on the table.
b. *John put the book.
c. *John put on the table.
d. *put on the table.

EXERCISE 1

Identify the following verbs as one-, two-, or three-argument verbs: confirm, show, fall.

It's more than just verbs that have arguments, though. For example, consider the preposition *at*. This preposition also has an argument:

(10) a. John left the book at the library.
b. *John left the book at.

This information about the number and category of the arguments that a predicate takes is called the *argument structure* of the predicate.

4.2 θ-roles

Clearly, the argument structure of a predicate is related to its semantics. Whatever it is that makes *sleep* a one-argument predicate (i.e. intransitive) is related to what *sleep* means. Another addition to the lexical entry which

we need to make, and which also seems connected to the semantics of the verb, concerns the *θ-roles* that a given element assigns. θ-roles encode the semantic roles possessed by the various elements which are required by the predicate.

By way of illustration, consider (11):

(11) Bob destroyed the statue.

Part of what you know about (11) is that *Bob*, the subject, is the voluntary agent of the 'destroying' action and *the statue* is the object affected by that action. This is expressed in the theory by saying that *destroy* is a predicate which assigns the *Agent* θ-role to its subject and the *Patient* θ-role to its object.

Now compare (11) with (12), where *destroy* has been replaced by *like*:

(12) Bob liked the statue.

In (12) the subject of *like* isn't really an agent in the same way as (11). Bob's not really *doing* anything as the subject of (12) in the same way that he is in (11). Neither is Bob's action necessarily voluntary. Bob could, for example, like the statue in spite of himself. What's going on in (12) is that Bob is experiencing an emotion rather than performing an action. In order to encode the differences between sentences like (11) and (12), we say that *like* assigns the *Experiencer* θ-role to its subject rather than the Agent θ-role.

In the same way, the direct object *statue* in (12) isn't really an object affected by the action. In order to distinguish direct objects of verbs like *like* (where nothing is really happening to the statue) from direct objects of verbs like *destroy* (where something *is* happening to the statue), we say that the direct object of *like* is assigned the *Theme* θ-role.

Our main focus in this book is not lexical semantics or argument structure in themselves, so I won't go into tremendous detail about the nature of θ-roles. However, there are some aspects of θ-roles which do impinge on syntax. For example, part of your syntactic knowledge of English includes the fact that *destroy* and *like* behave differently with respect to the syntactic construction called 'wh-clefts':

(13) a. What Bob did was destroy the statue.
b. *What Bob did was like the statue.

It seems that the requirement for participating in this 'What X did was...' frame is that X must be assigned the Agent θ-role by the verb.

We also can see contrasts like the one in (14):

(14) a. What happened to the statue was Bob destroyed it.
b. *What happened to the statue was Bob liked it.

Since the direct object of *like* is assigned the Theme θ-role as opposed to the Patient θ-role, and is therefore not affected by the action of the verb, it can't participate in alternations like the one in (14). Other θ-roles, primarily

assigned by prepositions as opposed to verbs, include Goal, Source and Location:

(15) a. Mary gave the book to John.
GOAL
b. Alice checked the book out from the library.
SOURCE
c. I saw the match in Bilbao.
LOCATION

Expletives and the θ-criterion

In this section, we're going to look at so-called *expletive* elements. Now I know you'll be disappointed, but, in this context, we don't mean by 'expletive' things like '****!' and '%?^&!'. When we speak of expletives in the context of syntax, we mean certain kinds of DPs which look like they must *not* be assigned a θ-role. Expletives tell us a lot of interesting things about what you know about your language.

Consider first (16):

(16) It seems that John doesn't clean his house very often.

What would the θ-role be that *seems* assigns to its subject *it*? Well, it doesn't look like there is one. Certainly *it* is not an agent, or an experiencer, or any of the other semantic roles that subjects normally play. In fact, it appears as though *it* is contributing nothing to the semantic interpretation of the sentence whatsoever. For example, the following is completely nonsensical:[1]

(17) What seems that John doesn't clean his house very often? (Answer: it.)

More importantly, notice what happens if we try to put a determiner phrase like *the room* into the subject position of *seems*:

(18) *The room seems that John doesn't clean his house very often.

It's not that there's no coherent interpretation for (18). Quite the contrary, (18) could perfectly well mean there's something about the room which gives you the impression that John doesn't clean his house very often. The room could be a complete mess, for example. However, despite (18) representing a perfectly meaningful thought, it is simply ungrammatical.

We can explain the contrast between (16) and (18) in terms of θ-roles. Let's suppose that there's a general requirement that says that all non-expletive DPs must be assigned a θ-role. This requirement is called the θ-criterion:

(19) *The θ-criterion (preliminary)*
All non-expletive DPs must be assigned a θ-role.

Now let's assume that the subject position of *seems* is not a position where a θ-role is assigned. That will account for the ungrammaticality of (18). The DP *the room* requires a θ-role, but is not in a position where it can get one. But what about (16), then? Well, if we assume that an element like expletive *it* cannot bear a θ-role, we account for the fact that (16) is grammatical. *Seems* is a one-place predicate. Its one argument is the CP. It therefore has no θ-role to assign to its subject, but that's OK because the expletive *it* doesn't require one.

It isn't the only element which appears able to function as an expletive. *There* can also do so as well when it appears as a subject:

(20) There seems to be a problem.

We can use expletives as a general test for whether a position is assigned a θ-role or not. If expletive *it* or expletive *there* can appear in the position and a non-expletive DP can't, then the position isn't assigned a θ-role. For example, the subject of the predicate *be likely* is not a position which is assigned a θ-role when it takes a CP complement:

(21) a. *John is likely that he's a spy.
b. It is likely that John is a spy.
c. There is likely to be a spy in the room.

The fact that we can get expletive *it* and *there*, but not any other kind of DP, shows that the subject position of *be likely* is not a position where a θ-role is assigned.[2]

A question arises, though, as to why English has expletives. They don't contribute anything to the semantics of the sentence. If *seems* and *be likely* don't assign θ-roles to their subjects, why can't the subjects just be left out? The fact of the matter is that they just can't be:

(22) a. *Is likely that John is a spy.
b. *Is likely to be a spy in the room.

It's not clear what information is present in (21b) and (21c) which is not present in (22), but the sentences are still ungrammatical.

On the basis of facts like this, Chomsky (1981) proposed the *Extended Projection Principle* (EPP):

(23) *The Extended Projection Principle*
All clauses must have a subject.

(The EPP is an extension of an older principle, the Projection Principle, which said that all information in the argument structure of a lexical item had to be present in the syntax.) The EPP just appears to be one of those 'brute force' things. (23) doesn't explain why clauses have to have a subject, but it seems we need it on the basis of the ungrammaticality of sentences like (22a) and (22b).

All clauses must have a syntactically represented subject even if that means they have a semantically empty (expletive) subject.

To finish this section, let's go back to the θ-criterion for just a second. As we've got it so far, the θ-criterion says that all non-expletive DPs need to be assigned a θ-role, but it looks like we need to say something more. Consider (24):

(24) *John showed Bill.

BEFORE CONTINUING

What does the θ-criterion predict about (24)?

Notice that (24) does *not* violate the θ-criterion as we formulated it back in (19). We've got two DPs in (24), *John* and *Bill*. The θ-criterion in (19) says that, if you're a non-expletive DP, then you must have a θ-role. *John* and *Bill* are both non-expletive DPs. *John* has the Agent θ-role and *Bill* has the Goal θ-role. So, as far as (19) is concerned, (24) is fine. The EPP isn't violated either, as (24) has a subject, *John*.

Given the discussion at the beginning of the chapter, though, it's clear what's gone wrong. *show* is a three-argument verb, not a *two*-argument verb. It requires an Agent, Goal and a Theme. So we need (25) instead of (24):

(25) John showed Bill the book.

All three of *show's* arguments have to be realized. You can't just decide not to use one of them. Therefore, it looks like what we need the θ-criterion to say is that there's a one-to-one relationship between θ-roles that a predicate has and elements which those roles are assigned to.

(26) *The θ-criterion (revised)*
All syntactic arguments of a predicate must be assigned a θ-role, and all θ-roles of a predicate must be assigned.

I've changed from talking just about DPs in the preliminary version (19) to 'all syntactic arguments' in (26) just to cover the fact that, as we've alluded to above, other categories like CPs and PPs are sometimes arguments of a predicate.

θ-Theory and PRO

Another impact that θ-Theory has on the syntax has to do with the analysis of infinitival constructions. Infinitives differ from the kinds of verbs that

we've seen so far because they don't seem to have any independent tense interpretation. Instead, the way they're interpreted depends on the tense of the verb in the main clause. The infinitive is the same in (27a) and (27b), but the 'going' action is interpreted as taking place in the present in (a) and in the past in (b):

(27) a. We are trying to go.
b. We tried to go.

Consider now (28), from the point of view of θ-Theory:

(28) We hope to hit the target.

We've got two verbs here, *hope* and *hit*. *Hope* is pretty straightforward. It assigns the Experiencer θ-role to its subject and takes a clausal complement, either a *non-finite* (tenseless) clause, as in (28), or a *finite* (tensed) clause as in (29).

(29) We hope that John will arrive tomorrow.

However, *hit* in (28) is more interesting. Under normal circumstances, *hit* takes two arguments – a subject, to which it assigns the Agent θ-role, and a direct object, to which it assigns the Patient θ-role.

(30) John hit the target.

Hit is a two-argument verb, and, as we discussed in the last section, the θ-criterion says that both of those θ-roles must be assigned. That looks right, given (31):

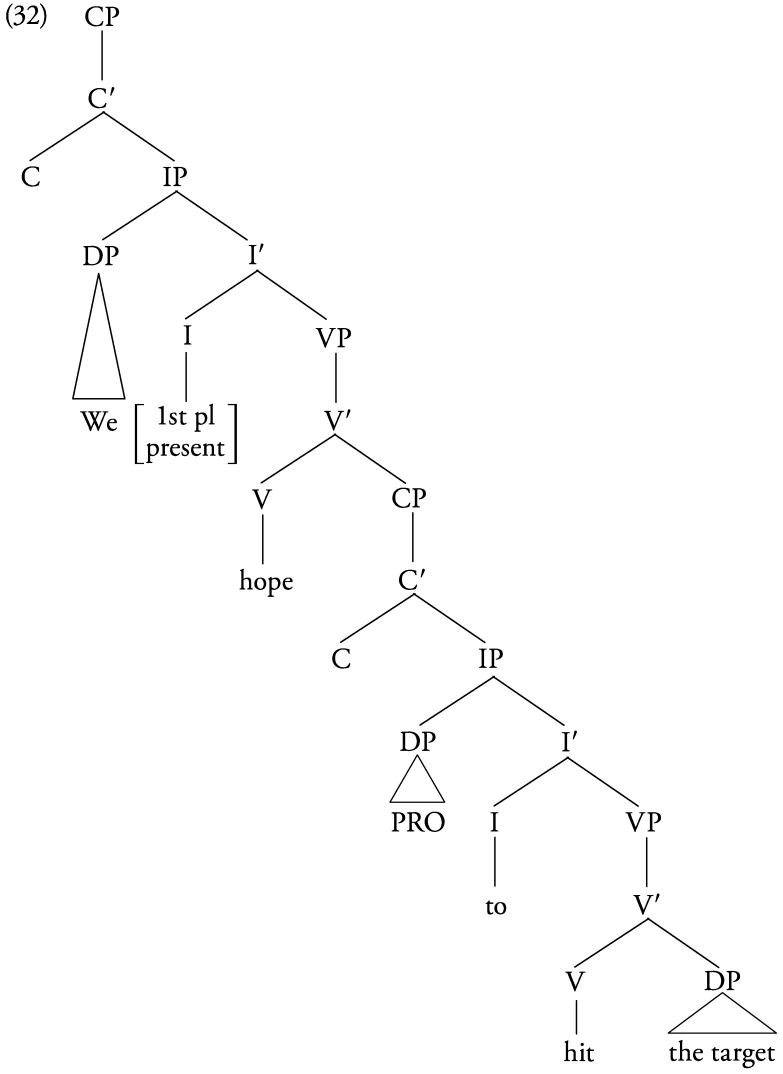
(31) a. *John hit.
b. *hit the ball.

So what's going on in (28)? In (28), *hit* has a direct object, *the target*, but it doesn't appear to have a subject, in violation not only of the θ-criterion (26) but also the EPP (23). Not only does *hit* seem to have a θ-role that it's not assigning (even though the θ-criterion says it has to), but the infinitive clause *to hit the target* also violates the EPP by having no subject.

These two problems have a single solution. If there were an *unpronounced* DP of some kind in the specifier position of the infinitival IP (i.e. in its subject position), we'd be all set. Rather than being different, *hit* would be exactly the same both in finite clauses and in infinitives. It would have two obligatory θ-roles to assign. The only difference would be that in infinitives the Agent θ-role is assigned to an unpronounced element. Also, the EPP would be obeyed rather than violated, since the clause would have an unpronounced DP as its

subject. Let's refer to this unpronounced subject of infinitivals as *PRO*. The structure of (28) would therefore be (32):

(32)



You might wonder why in (32) the complement of *hope* is a CP rather than an IP. There's no overt complementizer, and you can't even put one in:

(33) *We hope that PRO to hit the target.

It really is a CP, but actually showing it requires a bit more theory than we've got right now. However, we'll come back to the question in Chapter 8 and see how it follows from something called the *PRO Theorem*. For the

moment, it's something that you can just memorize: infinitives that have PRO are CPs.

By the way, don't be confused by the fact that the infinitive marker *to* is homophonous with the preposition *to*. They're different. In *we hope to hit the target*, *to* can't be a preposition. This is because the sister of *to*, namely *hit the target*, is a VP, and prepositions take DPs as their complement, not VPs:

(34) a. I told John about [$_{DP}$ the delivery].
b. *I told John about [$_{VP}$ hit the target].

It's I°s that take VPs as their complements. So given that *to* takes a VP complement in *we hope to hit the target*, it must be an I°.

Assuming that there is an unpronounced PRO as the subject of infinitivals also explains a fact about their interpretation. (28) is not interpreted as something like 'we hope that somebody or other hits the target' or 'we hope that people in general hit the target'. What (28) means is 'we hope that *we* hit the target'. We can account for this fact by saying that the subject of *hope*, which is *we*, 'controls' PRO, and that *hope* is a *subject-control* verb. That is, PRO obligatorily refers back to the subject of *hope*.

Many languages of the world have subject-control verbs, but English is unusual in that it also has *object-control* verbs like *persuade*:

(35) John persuaded Bill [$_{CP}$ PRO to go to college].

Here, it's not the subject *John* which is the controller of PRO. It's not John who's going to college in (35). Not is it some unspecified person or people in general. Rather the controller of PRO is the object *Bill*. John persuaded Bill that *Bill* should go to college.

In addition to 'controlled' PRO, there are also instances of what is called 'arbitrary' PRO:

(36) a. To return to the scene of the crime is never a good idea.
b. PRO to return to the scene of the crime is never a good idea.

Just as in (28), *return* seems to be missing a subject in (36a). Again, though, we will analyse (36a) as having PRO as the subject of *return*, as illustrated in (36b). In this case, though, PRO does not have a controller and instead has a generic interpretation. Unlike the cases of controlled PRO, what (36) means is 'for anyone to return...' or 'for people in general to return...'.

4.3 Morphological case

Having introduced θ-Theory, and seen the implications that it has for our model of what speakers of English know, let's now turn to another area in

which aspects of a word's lexical entry have an impact on the syntax: Case Theory. To begin our discussion of Case, let's take a look at the pronoun system of English. One of the interesting things about pronouns in English is that, as you know, they change form depending on their function in the sentence. So, for instance, when the first person singular pronoun is used as a subject, it is 'I'. However, when the same pronoun is the object of a verb or the object of a preposition, it takes the form 'me'. So we have contrasts like (37) and (38):

(37) a. I like The Beatles.
b. *Me like The Beatles.

(38) a. Mary gave the book to me.
b. *Mary gave the book to I.

This change of form in the pronoun is just about all that's left of a system which existed in older stages of English where not just pronouns, but also determiners and nouns in general, changed form depending on their grammatical function in the sentence. Also, the number of different functions which demanded a change was greater. So, for instance, in Old English we had contrasts like the following (from Freeborn 1992):

(39) a. þæt wif andwyrde.
The-NOM woman-NOM answered
'The woman answered.'

b. God cwæþ to þam wif-e
God-NOM said to the-DAT woman-DAT
'God said to the woman ...'

In (39), the DP *the woman* changes form depending on its function in the sentence. When it is the subject, it appears as *þæt wif*. When it is an indirect object, it appears as *þam wife*. And there are languages of the world which are more complicated still. Nouns in Latin, for example, take on different forms depending on whether they are the subject or the object of the verb, or even depending on exactly which preposition they are the object of.

Taking over some traditional terminology from the study of Latin grammar, these changes in the form that the words take fall under the general heading of 'Case'. When a noun is in a certain form because it is the subject of a sentence, it is said to be in the *nominative case*. When it is in direct object position, it is said to be in the *accusative case*. When it is the object of a preposition, it is usually said to be in an *oblique case* (for example, the ablative case in Latin). Finally, as we mentioned in the previous chapter, when a noun is a possessor (e.g. 'John' in *John's book*), it is in the *genitive case*. The important point to note is that the case a DP is in depends on other elements in the structure, usually a verb, a preposition or a determiner. So part of the lexical entry for Vs and Ps, for example, will

need to include what case they assign to the various elements that they're in a relationship with.

4.4 Morphological case vs abstract case

I want now to clarify something that we've so far been assuming regarding case assignment. We started the discussion by talking about pronouns in English, which actually change form depending on whether they are nominative or accusative case. However, from now on we are going to assume that *all* DPs, in English as well as in Latin, are actually assigned a case, even when there's no change in form:

(40) a. Brutus hit the ball.
b. The ball hit Brutus.

Even though the two *Brutuses* look identical in (40), what we want to say is that the 'Brutus' in (40a) is really *Brutus-nominative* and the 'Brutus' in (40b) is really *Brutus-accusative*. It's just an accident of the morphology of English that the case marking in English is unpronounced. In Latin, the case marking would be pronounced (*Brutus* in (40a) but *Brutum* in (40b)). It's the same situation with *the ball*. It's really *the ball-accusative* in (40a) and *the ball-nominative* in (40b).

To formalize this distinction, when we're talking about *morphological* case, as in 'The ending for the instrumental case in Russian is "-om"', we'll use a lower-case 'c'. When we're talking about *abstract* case, we'll use a capital 'C'. The two (case and Case) are clearly related. Morphological case is the realization of abstract Case, but this realization can happen in different ways in different languages.

So, staying with English, in the sentence *John hit him*, *him* is assigned (abstract) accusative Case by the verb *hit*. Because it's a pronoun, that abstract accusative Case is realized as morphological accusative case (*him*). *John* on the other hand is assigned (abstract) nominative Case, but there's no realization of that with a visible (morphological) case-ending. When pronouns occur as the objects of prepositions, they are assigned abstract oblique Case, but this is realized morphologically as accusative case in English (*from him*). (It is not **from he*, or something which is neither nominative nor accusative (like **from hime*).) For the most part, we'll be concerned in the rest of the book with abstract Case rather than morphological case.

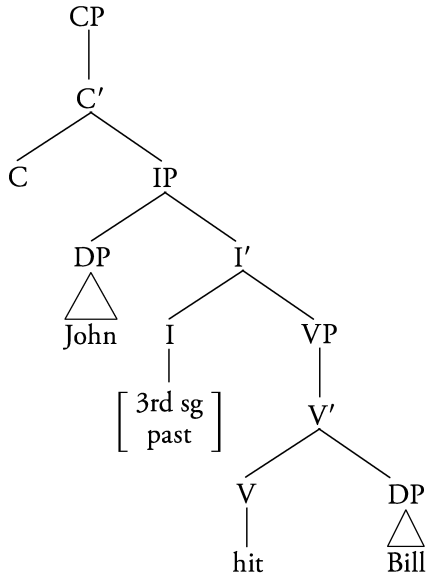
Case assignment and M-command

Let's move on now to examine in a little more detail where Case assignment comes from. Under what conditions is a given DP assigned Case? We hinted just a moment ago that Case is assigned to DPs in certain structural configurations by certain other elements. Objects of verbs and prepositions are assigned

accusative and oblique Case respectively, subjects are assigned nominative Case and the DP specifiers of possessive DPs are assigned genitive Case by *'s*. In this section, we're going to examine the configurations under which Case assignment can take place. Put another way, we'll see that a particular structural/configurational relationship needs to obtain between the Case assignor and the assignee.

Let's start by looking at nominative and accusative Case assignment.

(41)



The functional category I° seems clearly to be responsible for nominative Case assignment because different kinds of I° correlate with different Case-assignment possibilities. When I° is finite (42a), it assigns nominative Case to its subject. When I° is non-finite/infinitival (42b), it doesn't.

(42) a. He hit the ball.
b. *He/Him to hit the ball would be amazing.

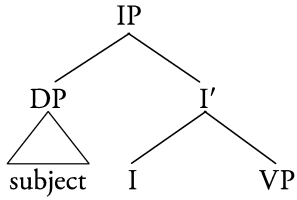
With respect to accusative Case assignment, it's clearly the V that's relevant. Certain V°s, which in traditional grammar literature are called 'intransitive' verbs, don't assign accusative Case, but others ('transitive' verbs) do:

(43) a. John bought the bed.
b. *John slept the bed.

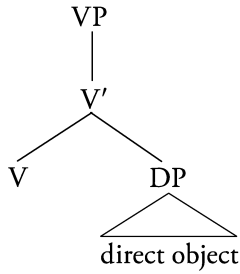
The question is, though, what the precise relationship has to be between the Case assignor and the Case assignee. Let's start by considering the two standard cases that we have: nominative Case assignment by finite I and accusative Case

assignment by V:

(44a)



(44b)

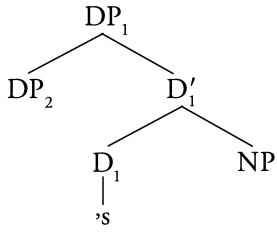


If we were just looking at the VP case in (44b), you might think that c-command is the relevant relationship between Case assignor and Case assignee. The V does c-command its direct object. However, if we take a look at (44a), it's clear that c-command won't work. I doesn't c-command the subject in (44a). The first branching node dominating I° is I′, and I′ doesn't dominate the DP subject.

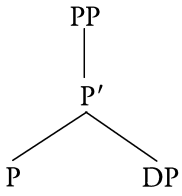
There is one thing that (44a) and (44b) do have in common, though: the Case assignment is taking place *within* the maximal projection of the Case assignor. That is, the subject is in the specifier of IP position. It is within the IP which has the Case assignor I as its head. Similarly, in (44b), the direct object is also within the VP headed by V, which is the Case assignor.

This also appears to be true of the other two instances of Case assignment that we've seen. The DP assigned genitive Case by *'s* is in the specifier of the DP headed by the *'s* determiner, just like assignment of nominative Case by I (45a), while assignment of oblique Case by prepositions, like accusative Case assignment by verbs, takes place in the PP headed by P (45b).

(45a)



(45b)



We can formalize this notion of 'within the XP' by introducing the notion of m-command (where 'm' is after 'maximal', since the XP is the 'maximal' projection of the head X^0):

(46) *m-command*
A node α m-commands a node β if and only if the first maximal projection dominating α dominates β.

We can now define Case assignment in terms of m-command:

(47) α assigns Case to β if and only if
(a) α is a Case assignor (i.e. V, finite I, P or D), and
(b) α m-commands β.

If the definition of m-command sounds a little familiar, it's because it's almost exactly the same as the definition of c-command from Chapter 2. The only difference is that for c-command you find 'the first branching node dominating α' and for m-command you find 'the first maximal projection dominating α'. They're two different relations, though. While V c-commands the direct object DP in (44b), I^0 does *not* c-command the subject DP in (44a), so Case-assignment can't be defined in terms of c-command. So we definitely seem to need m-command in addition to c-command.

The need for barriers

Although m-command seems to be the right relationship between Case assignor and assignee, we need to rein things in a bit. It's not that you can just assign Case to anything that you m-command. To see why consider (48):

(48) I read the book in Moscow.

We want things to work out so that *read* (a verb) assigns accusative Case to *the book*, but, crucially, not to *Moscow*. *Moscow* must get oblique Case from the preposition *in*. We can see this even more clearly if we look at the equivalent of (48) in a language like Russian:

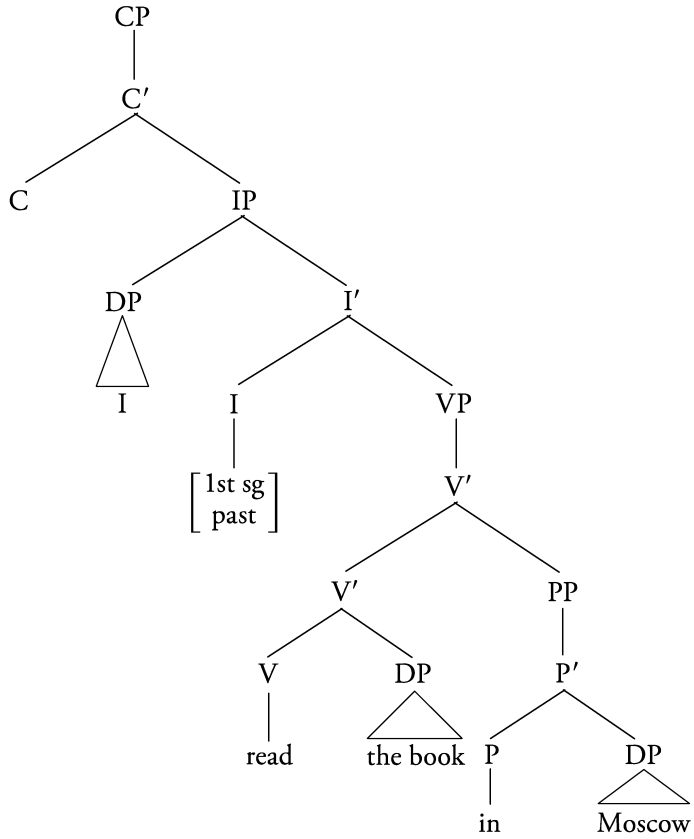
(49) a. Ja pročital knigu v Moskvje.
I read book-ACC in Moscow-LOC
'I read the book in Moscow.'

b. *Ja pročital v Moskvu
I read in Moscow-ACC

Read's attempt to assign accusative Case to *Moscow* seems to be blocked by the preposition *in*.

However, consider the tree for (48):

(50)



Read does m-command *Moscow*. (The first maximal projection dominating *read* is VP, and VP does dominate the NP *Moscow*.) Therefore, if Case assignment depends on m-command and nothing else, we have a problem.

What we need is some way of 'protecting' objects of prepositions from being assigned Case by verbs. To do this, let's assume that heads exert a 'zone of control' within their maximal projections, preventing Case assignment by anything from outside. This idea is worked into the definition by adding a clause which says that maximal projections (i.e. XPs) form barriers to Case assignment from outside the maximal projection.

(51) α assigns Case to β if and only if
(a) α is a Case assignor, and
(b) α m-commands β, and
(c) no barrier intervenes (where any maximal projection is a barrier).

With this definition of Case assignment, although the verb still m-commands the object of the preposition, it can't assign Case to it because of the intervening barrier PP. The PP barrier blocks Case-assignment from outside.

EXERCISE 2

Which elements assign Case to the DPs in the following sentence?
John showed the book to Bill on Friday.
Draw the tree.

4.5 Exceptional case marking

Now that we have the standard Case assignment configurations under our belts, let's take a look at one configuration which has received a fair bit of attention over the years:

(52) I believe Mary to be the best candidate.

(52) is most commonly referred to as an Exceptional Case Marking (ECM) configuration. However, because *Mary* doesn't differ between *Mary-nominative* and *Mary-accusative*, (52) doesn't tell us as much as it could. By replacing *Mary* with a pronoun, we can use morphological case to tell us more about the abstract Case-assignment situation:

(53) a. I believe her to be the best candidate.
b. *I believe she to be the best candidate.

On the evidence of *her* in (53a), Mary is receiving accusative Case in (52), not nominative Case. The lack of nominative Case is not surprising. We've seen that infinitival I° (*to*) does not assign nominative Case. (Only finite I° assigns nominative Case.) The accusative Case is surprising, though. The only potential source for accusative Case is the verb *believe* in the higher clause.

However, unlike the cases of accusative Case marking that we've seen so far, *Mary* is not the direct object of the verb in (52). It's the subject of the lower infinitival clause. This is why it's called 'Exceptional' Case Marking. If you think about what (52) means, you can see pretty clearly that *Mary* is the subject of the lower clause and not the direct object of *believe*. The object of my belief isn't actually Mary. Rather what I believe is that some situation or state of affairs is true, and that situation or state of affairs is represented by the IP *Mary to be the best candidate*.

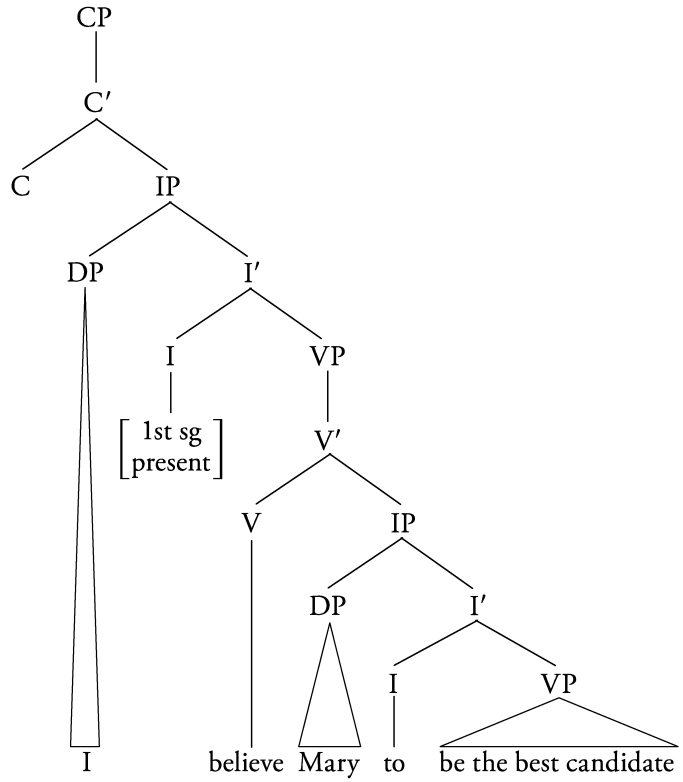
You can see this even more clearly by comparing two cases with *believe*: one where it takes a DP direct object and one where it takes an infinitival complement:

(54) a. #I believe the politician, but I think he's lying.
b. I believe the politician to be lying.

The '#' symbol is used to indicate a sentence which is not syntactically ill-formed, but which is deviant for some other reason. In this case, it's because (54a) represents a contradiction. When *believe* takes a DP complement, the DP is a true direct object. Therefore, in (54a) the object of belief is the politician, and it is contradictory to say that you both believe and don't believe the politician. In (54b) on the other hand, where *believe* is followed by an infinitival phrase, the complement of *believe* is the IP and it is the *situation* or state of affairs which represents the object of belief. Therefore, there's no implied contradiction in (54b) because you don't believe the politician. You believe a situation to be true, namely that the politician is lying.

Therefore, the structure for ECM configurations will be (55), with the IP as the sister to the V:

(55)



Other ECM verbs include *want* and *expect*.

(56) a. I want Mary to be the best candidate.
b. I expect Mary to be the best candidate.

Notice how an ECM verb like *believe* contrasts with a superficially similar verb like *persuade,* which we mentioned earlier:

(57) I persuaded Mary to leave the room.

In this case, *Mary* is clearly the *object* of *persuade.* Unlike *believe,* you're not persuading a situation or a state of affairs. You're persuading *Mary. Persuade* therefore has two objects, the persuadee, and what you persuaded him or her to do. *Believe* has one object, the clause, where the DP is the subject of that lower clause.

(58) a. I believe [Mary to be the best candidate].
b. I persuaded [Mary] [PRO to leave the room].
(cf. I persuaded Mary that Bill left the room.)

IP: it's not a barrier

However, it turns out that ECM configurations present a problem for what we've just said about Case assignment, and this is why they are called 'Exceptional'. We argued just a minute ago that *Mary* is not the object of the verb *believe,* but rather is the subject of the IP *Mary to be the best candidate.* But if that's the case, how is it that *believe* is able to assign accusative Case to *Mary*? IP is a maximal projection just like any other, and therefore should present a barrier to government by *believe.* It doesn't seem to though. If it did then (52) should be ungrammatical. This, unfortunately, is just one of those situations like the EPP, where brute force is the order of the day. The response to this problem which is taken in the syntax literature (e.g. Chomsky 1986) is to say that *infinitival* IP is 'defective'. Unlike other maximal projections, like PP, it is not a barrier to Case assignment.

At least there is a certain amount of initial plausibility to this idea. Infinitival I°, which is realized as *to,* is quite different from other kinds of inflection. *To* has no person and number or tense information, for example. It doesn't assign nominative Case either; as we saw in (42) above. We therefore need the subject of the IP to be exceptionally Case marked with accusative Case from the matrix verb. We'll also see with respect to other phenomena in other chapters that infinitival IP doesn't act as a barrier the way you'd expect it to, so at least the description seems right, even if it doesn't give any particular insight into what the ultimate explanation might be.

If, on the other hand, a CP is present, then ECM is blocked by the barrier, just as you'd expect. This can be seen in the following contrast, adapted from Haegeman (1994):

(59) a. I know [$_{IP}$ Mary/her to be an honest person].
b. *I know [$_{CP}$ whether [$_{IP}$ Mary/her to be an honest person].

In (59a) we see that *know* is an ECM verb. However, when we insert the complementizer *whether,* we know that we must have a CP. The ungrammaticality of (59b) suggests that Case is genuinely coming from the higher verb *know,* rather than infinitival I assigning oblique Case. Furthermore, as our definition of Case assignment in (51) would predict, the CP barrier blocks Case

assignment from *know*. Notice that it's not as though there's something incoherent about the thought expressed by (59b). If we replace the infinitival I with a finite I, the sentence becomes grammatical:

(60) I know [$_{CP}$ whether [$_{IP}$ Mary/she is an honest person]].

Here *Mary* is assigned nominative Case by the finite I. This suggests that, after we make the stipulation that infinitival IP is 'defective', the general theory of barriers accounts for Case assignment in ECM contexts.[3]

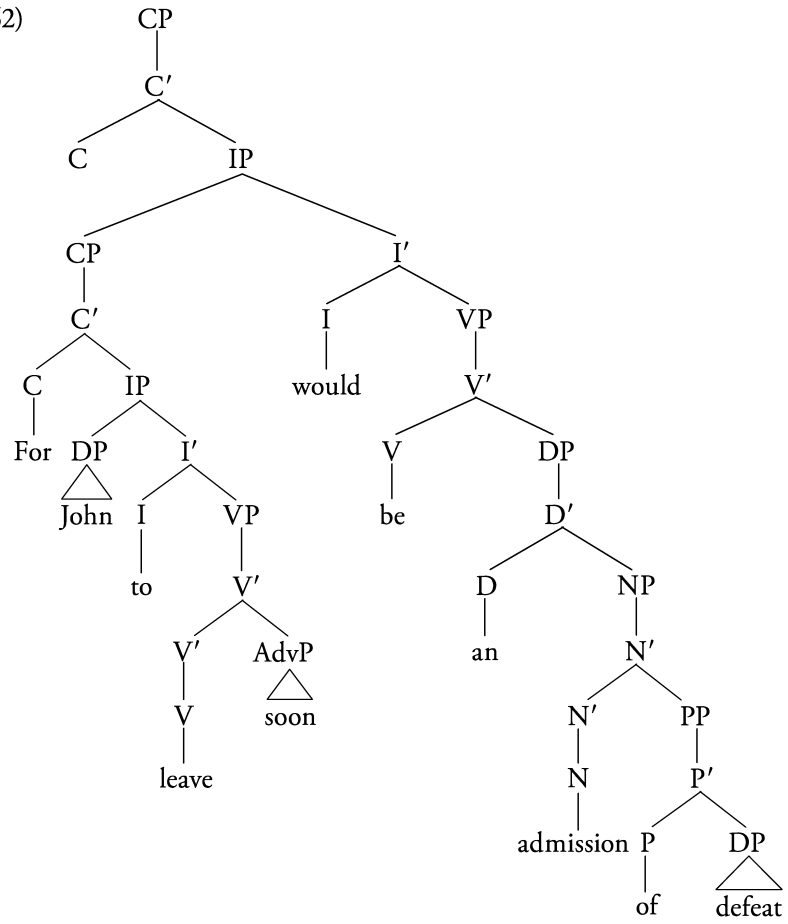
Case marking in other non-finite clauses

Let's now turn to a couple of other constructions in which the subject of an infinitival clause can be 'saved' by being Case-marked by something outside IP, like the accusative Case marking in ECM constructions which we just saw.

'For-to' infinitives are one instance where the subject of an infinitival clause is 'saved' by being case-marked by something outside the IP:

(61) For John to leave soon would be an admission of defeat.

(62)



Unlike the infinitives that we've seen so far, (61) doesn't have a PRO subject, but rather has an overt subject, *John*, introduced by the complementizer *for*. (Just as before, when we were talking about infinitival *to*, *for* can't be a preposition in (61) because *John to leave tomorrow* must be an IP, and it's *complementizers*, not prepositions, that take IP as their complement.) What happens in (61) is that the complementizer *for* assigns accusative Case to the subject of the IP *John to leave tomorrow*. Without *for*, (61) is ungrammatical:

(63) *John to leave soon would be an admission of defeat.

Here again, we see more evidence that IP does not constitute a barrier to Case assignment. Otherwise we would predict that both (61) and (63) should be ungrammatical.

Similarly, in the 'poss-ing' construction that we discussed in the previous chapter, the fact that the *'s* determiner assigns genitive Case to the DP is what allows it to appear as the 'subject' of the gerund.

(64) a. John's building a spaceship surprised his neighbours.
b. *John building a spaceship surprised his neighbours.
c. *John the building a spaceship surprised his neighbours.

Without a source for genitive Case (i.e. the determiner *'s*), the 'poss-ing' construction is ungrammatical.

4.6 The Case Filter

In discussing ECM and other non-finite constructions the way we did in the previous section, we hinted at one of the central arguments for the existence of abstract Case assignment, which I now want to make explicit. This argument was first made by the linguist Jean-Roger Vergnaud in a famous letter to Noam Chomsky in the late 1970s. Vergnaud noted that it appears we need to say that *abstract* Case assignment is taking place in order explain certain contrasts in grammaticality. Consider (65):

(65) a. They believe [$_{IP}$ John to be the best candidate].
b. *They hope [$_{CP}$ John to be the best candidate].
(cf. They hope that John is the best candidate.)

In the previous section, we discovered that *believe* is an ECM verb, able to assign accusative Case to the specifier of its infinitival IP complement. *Hope* by contrast is not an ECM verb because *hope* takes a CP complement. If we have a rule which says that DPs must always be assigned Case, we can account for the contrast in (65). *Believe* can provide a source of Case, but *hope* cannot because CP (as a maximal projection) is a barrier to Case assignment. This requirement that DPs be assigned Case is usually referred to as the Case Filter:

(66) *The Case Filter*
A DP must be assigned Case.

However, having just reminded ourselves that infinitival *to* does not assign Case to the subject position, we do need to say something about PRO, as in (67):

(67) John hoped PRO to win.

PRO must be a DP. It's DPs that turn up in subject position and it's DPs that verbs assign their θ-roles to. But if PRO is a DP, we incorrectly predict that (67) is ungrammatical, as things stand. The Case Filter in (66) says that DPs need to be assigned Case, but there isn't anything around. There's no nominative Case from I^{o} (since *to* is non-finite) and any potential Case-assigner outside the CP is blocked by the CP barrier.

The usual way to account for (67) is to say that the Case Filter is restricted to *overt*, that is to say *pronounced*, DPs. Unpronounced DPs like PRO don't require Case. Viewed in this way, the Case Filter is essentially a morphological requirement.

One reason why this pronounced/unpronounced distinction might look like the right one is that the Case Filter applies to expletives just as it applies to other DPs. If I've just killed John by dropping an undetectable poison in his coffee, I can say (68a), but not (68b):

(68) a. I want it to seem that John died of natural causes.
b. *I hope it to seem that John died of natural causes.

The DP *it* in (68) is an expletive, as *seem* does not assign a θ-role to its subject. However, the contrast suggests that expletive *it* does require Case. The only difference between (68a) and (68b) is that *want* can function as an ECM verb but *hope* cannot. Therefore *it* can be assigned Case in (68a) but not (68b), thereby flouting the Case Filter.

The same contrast is observed with expletive *there*:

(69) a. I want there to be six reports on my desk by noon.
b. *I hope there to be six reports on my desk by noon.

If the Case Filter applies to all *overt* DPs, whether expletive or not, the contrasts in (67) to (69) are explained.

EXERCISE 3

Do the following sentences violate the θ-criterion, the Case Filter or both? Why?

(a) *John seems that his wife left him.
(b) *I promised it to seem that the arena was full.
(c) *I decided the room to seem that John had left.
(d) *I tried there to be three books on the table.

4.7 The visibility condition: an improvement on the Case Filter?

So far the Case Filter looks like a reasonably plausible requirement. Pronounced DPs need Case and unpronounced DPs don't need Case. However, the Case Filter is not without its shortcomings. So let's take a look at a popular replacement for the Case Filter, *the Visibility Condition*, which proposes that there's an important relationship between the two topics we've been looking at in this Chapter: θ-role assignment and Case.

Chomsky (1981) proposed a new way of looking at sentences involving Case violations. He suggested that there wasn't really a separate Case Filter at all. Instead, what it really all boiled down to was θ-role assignment. In order for V, I or P to 'see' a DP and assign it a θ-role, it had to become 'visible' to the θ-role assignor. The way to become 'visible' is to be assigned Case. This requirement is referred to in the literature as the *Visibility Condition*.

Under this approach, Case assignment is now not a direct requirement. There's no principle that says 'you must have Case'. The grammar couldn't care less whether you have Case or not. However, if you don't get Case, you're going to run into trouble if you need a θ-role down the line. You won't be 'visible' for θ-role assignment, and you'll end up violating the θ-criterion:

(26) *The θ-criterion*
All syntactic arguments of a predicate must be assigned a θ-role, and all θ-roles of a predicate must be assigned.

So, unlike with the Case Filter, the division isn't between overt and non-overt DPs, but rather between things that need θ-roles and things that don't.

Let's look at a couple of examples to see how the Visibility analysis differs from the Case Filter analysis:

(70) a. I believe [John to be the best candidate].
b. *I hope [John to be the best candidate].

The ECM verb *believe* assigns Case to *John* and it is therefore visible for θ-role assignment from *to be*. On the other hand, *John* in (70b) is not Case marked, and therefore not visible for θ-role assignment. As a result (70b) violates the θ-criterion.

It's been argued that the Visibility analysis is conceptually superior to the Case Filter analysis. There's no question that we need the θ-criterion in our theory. If we can get rid of the Case Filter, and reduce all Case-related violations to the θ-criterion, then we have a simpler theory – only one principle and not two.

However, the Visibility approach to Case is in turn not without its problems, some of them rather serious. Since Case and θ-role assignment are explicitly tied together, it makes the prediction that wherever you have the need for a θ-role you should have Case, and wherever you don't have a θ-role there should

be no need for Case. There are potential counter-examples to this prediction in both directions.

BEFORE CONTINUING

For each of these two predictions, we've seen something in this very chapter which would present a problem. What are they and why are they problematic?

On the one hand, there are elements which do receive a θ-role but which seem not to need Case. For example, PRO. Consider the contrast in (71):

(71) a. *I hope John to be here.
b. I hope PRO to be here.

The Case Filter analysis makes a correct prediction here, but the Visibility approach predicts that both should be ungrammatical. Both *John* and PRO need a θ-role as the subject of *to be*. There's no difference between them in this respect. This has led some researchers (e.g. Martin 1999) to claim that PRO is assigned a form of Case called 'null Case', which overt DPs like *John* can't bear. However, this is still a reasonably controversial suggestion.

On the other hand, the Visibility approach also predicts that expletives shouldn't need Case. Since they are not assigned a θ-role, and in fact must never be assigned one, they should be free to be Case-less. This is not the case, however, as we can see in (72) and (73):

(72) a. I want it to seem that John died of natural causes.
b. *I hope it to seem that John died of natural causes.

(73) a. I want there to be six reports on my desk by noon.
b. *I hope there to be six reports on my desk by noon.

(72) and (73) seem to show clearly that expletive *it* and expletive *there* can only appear in Case marked positions. This is predicted by the Case Filter account, but not by the Visibility account. Visibility supporters do have a potential reply for (73). It's been claimed that expletive *there* transmits Case to its 'associate' (*six reports* in (73)). Thus it might *appear* as though expletive *there* needs Case when it's actually the associate that needs it. However, the explanation for the behaviour of expletive *it* is still unclear under Visibility.

For the rest of the book, what I'll do is retain the Case Filter and the θ-criterion as separate principles, and not adopt the Visibility analysis. However, I've discussed it because it is an idea that you will run into in the literature. And there are lines of research that you might run into in your

future reading, like the null Case for PRO idea, which seems like a weird suggestion unless you know the Visibility analysis and the problems that people have been trying to solve.

Summary

In this chapter, we focused on predicate/argument relations – that is, the relationship that DP subjects, direct objects and so on have to their respective verb, adjective or prepositional heads. The first area we looked at is called θ-Theory. This has to do with the way in which predicates assign semantic roles, called θ-roles, to their arguments. We discovered that every θ-role that a predicate has must be assigned to an argument, and that every argument must bear a θ-role. This is formalized in the principle called the θ-criterion.

(26) *The θ-criterion*
All syntactic arguments of a predicate must be assigned a θ-role, and all θ-roles of a predicate must be assigned.

We also discussed elements called *expletives*, which don't seem able to bear a θ-role, and we saw how this led to the postulation of the *Extended Projection Principle*, the requirement that all IPs must have a subject. We then concluded that infinitival clauses must in fact contain an unpronounced subject, which we referred to as PRO.

The second area that we looked at was Case Theory. We argued that abstract Case assignment takes place even though most DPs in English don't show any reflection of it. This led to the postulation of the Case Filter:

(66) *The Case Filter*
A DP must be assigned Case.

There must also be a particular phrase-structure (configurational) relationship between the Case assignor and the Case assignee.

(51) α assigns Case to β if and only if
(a) α is a Case assignor (i.e. V, finite I or P), and
(b) α m-commands β, and
(c) no barrier intervenes (where any maximal projection is a barrier, except infinitival IP).

We then looked at a popular alternative to the Case Filter, the Visibility Condition, in which there is no direct requirement that DPs be assigned Case. Instead, Case plays an indirect role. Case assignment can take place or not. The grammar isn't bothered about it one way or the other, metaphorically speaking. However, if a DP doesn't have Case, then it is not 'visible' for θ-role

assignment. If it later transpires that that DP needed a θ-role, then it will trigger a θ-criterion violation.

Open issue: The case filter, visibility and degrees of ungrammaticality

In this chapter, we've discussed Chomsky's (1981) suggestion that the Case Filter be eliminated and replaced by the Visibility Condition. However, Epstein (1990) argues for retaining the Case Filter as an independent principle of grammar. The core of Epstein's argument is that a theory in which the Case Filter is replaced by the θ-criterion is incapable of correctly accounting for intuitions about different *degrees* and different *kinds* of ungrammaticality in certain cases. On the other hand, he argues, a theory which keeps the Case Filter and the θ-criterion as separate principles explains those contrasts. I'll adapt and simplify what Epstein says a bit, but it's worth taking a look at as an example of how evidence that you wouldn't normally think of looking at can be relevant for the theory that we're developing.

Citing observations going back as early as Chomsky (1965), Epstein notes that differing *degrees* of ungrammaticality can be used as a fairly direct test for the existence of grammatical principles. We've been relying so far on the fact that you have intuitions about which sentences are grammatical and ungrammatical. However, it's not the case that all ungrammatical sentences feel the same. You also have intuitions about whether sentences are *more* ungrammatical than others, or are ungrammatical in different ways.

We can explain these kinds of judgements by saying that a sentence which feels *more* ungrammatical does so because it violates more principles, and sentences which feel ungrammatical *in a different way* do so because they violate different principles. By way of example, Epstein cites the following (from Chomsky 1965):

(74) a. *John compelled.
b. *Golf plays John.
c. John plays golf.

Here, Chomsky argues that there is a three-way contrast in grammaticality. (74c) is well-formed, but (74a) and (74b) are not. However, the ungrammaticality of (74a) and (74b) is not of the same kind. (74a) is a violation of the θ-criterion. *Compel* is transitive, not intransitive, and a required argument is missing. On the other hand, in (74b) all of the arguments are present, but the DPs which are present in the subject and object positions are of the wrong semantic type. It's people that play games and not vice versa. (74b) satisfies the θ-criterion, but assigning the Agent θ-role to a game makes the sentence deviant.

With this in mind, let's consider the contrast in (75):

(75) a. *I hope John to think that Bill left.
b. *I hope John to be likely that Bill left.

In both (75a) and (75b), the DP *John* is in a position which does not receive Case. However, in (75a) but not (75b), *John* is a position where, at least in principle, a θ-role could be assigned to it. The verb *think* has the Experiencer θ-role available to assign to its subject. The predicate *be likely* on the other hand, does not assign a θ-role to its subject.

What would the Case Filter and Visibility analyses have to say about (75)? Let's take the Case Filter analysis first. Under this analysis, where the Case Filter and the θ-criterion are separate, (75a) would violate only the Case Filter and not the θ-criterion. Nothing prevents *John* from getting the Experiencer θ-role which *think* has to assign. It's just that *John* doesn't get Case because *hope* isn't an ECM verb. In (75b), on the other hand, *John* violates both the Case Filter and the θ-criterion.

Under the Visibility analysis, in which Case assignment is merely a prerequisite for θ-role assignment, (75a) violates only the θ-criterion. There is no Case Filter, but if *John* doesn't get Case then it can't get the θ-role it needs. Since *John* doesn't get Case, it's not visible to receive the Experiencer θ-role which *think* has to assign. The Visibility analysis of (75b) is exactly the same. *John* doesn't receive Case, and so is invisible for θ-role assignment. The fact that in (75b) there's no θ-role for *John* to get even if it were visible is irrelevant. We're still left with just a θ-criterion violation and nothing else.

So, what do the Case Filter and the Visibility analyses predict with respect to *degrees* of ungrammaticality if, as Epstein and Chomsky suggest, judgements about degrees of ungrammaticality are relevant for our theory (in addition to judgements about straight grammaticality vs ungrammaticality)? Well, under the Case Filter analysis, (75b) is predicted to be more ungrammatical than (75a). Both (75a) and (75b) violate the Case Filter, but (75b) violates the θ-criterion in addition. On the other hand, the Visibility analysis says that there's no difference between (75a) and (75b). They both violate the same single principle: the θ-criterion.

So that's the prediction about (75a) and (75b), now what's the judgement? This is admittedly a little tricky. Epstein's conclusion, after checking with both linguists and non-linguists, is that (75b) is indeed worse than (75a). As he puts it (1990: 317): 'in both [(75a)] and [(75b)], there is "something wrong with" the lexical subject of the infinitive: the two cases are identical in this regard. However, [(75b)] differs from [(75a)] in that, in this example, "something else is wrong", namely something interpretive'.

However, notice that there's a slightly easier way we can ask the question if we turn things around. The Visibility analysis predicts that (75a) and (75b) should feel exactly the same. Both violate the θ-criterion and nothing else. Does that seem right? I don't think so. Whatever's going on in (75a) and (75b), they certainly don't feel the same. The important point seems to be that you can kind of figure out what (75a) is trying to say. If a foreigner produced it, you could correct him or her with (76):

(76) I hope that John *thinks* that Bill left.

On the other hand, (75b) just doesn't make any sense. Replacing the infinitival with a finite verb doesn't create a grammatical sentence:

(77) *I hope that John is likely that Bill left.

Therefore, whatever the subtleties of judgement, (75a) and (75b) don't feel like they're ungrammatical in the same way. But that's what the Visibility analysis predicts. And that might be an argument that the Visibility analysis is on the wrong track.

What this Open Issue illustrates is the complex relationship between our intuitions and our theory. Because we've chosen to investigate our linguistic capacities using the scientific method, it is perfectly legitimate to try to reduce the complexity of the overall theory by eliminating principles and 'neatening things up'. However, we must make sure that at the end of the day our theory still makes the right predictions about the data.

FURTHER READING

Ray Jackendoff is one author who has looked extensively at the relationship between the lexicon and the rest of the grammar. His (1984) book outlines a theory of 'lexical-conceptual structure' which is further elaborated in Jackendoff (1990). Grimshaw (1990) is also a good general book on argument structure. On θ-Theory and Case, the standard reference is again Chomsky (1981), but again that book can be difficult. Textbooks such as Lasnik and Uriagereka (1988) and Haegeman (1994) contain more user-friendly discussion. There is discussion as well in Webelhuth (1995a). Chapter 1 contains some relatively accessible discussion of Case Theory and Chapter 2 discusses θ-Theory.

ANSWERS TO IN-TEXT EXERCISES

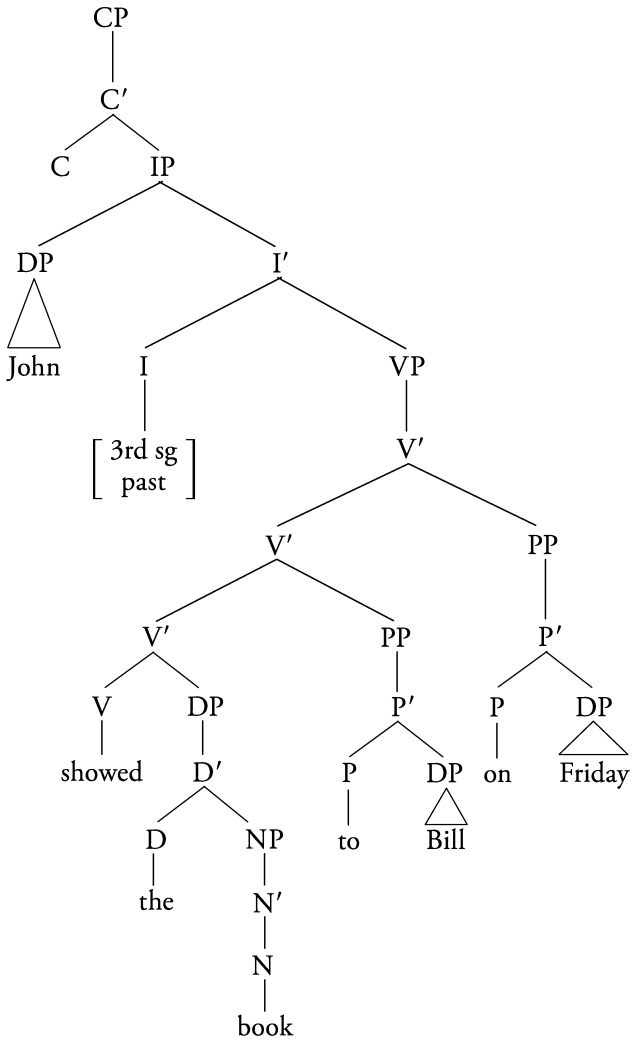
Exercise 1

Confirm is a two-argument verb. It requires a subject (*'confirmed the result') and also a direct object DP or CP. ('John confirmed the result' or 'John confirmed that the match had taken place', but not *'John confirmed'.)

Show is a three-argument verb under normal usage. It again requires a subject (*'showed the book to John'), and both the direct object and the indirect object are required as well. ('Mary showed the book to John' and not *'Mary showed the book' or *'Mary showed to John'.) You can say 'Mary shows pure-bred dogs', but *show* has a special meaning here, something like 'to enter (animals/plants) in competitions'.

Fall is a one-argument verb. It requires only a subject ('John fell') and may not take any objects. (*'John falls the stairs every day'.)

Exercise 2



John is assigned Case by the finite I, which contains the tense and agreement information. The finite I m-commands *John* (the first maximal projection dominating *John* is IP and IP dominates *John*) and no barrier intervenes between I and the DP subject. Finite I is also a Case assignor.

The book is assigned Case by the verb. The verb m-commands *the book* (the first maximal projection dominating *gave* is VP, which dominates *the book*) and no barrier intervenes between V and DP. V is also a Case assignor.

Both *Bill* and *Friday* are assigned Case by their respective prepositions *to* and *on*. They both m-command their objects (with PP_1 and PP_2 being the first maximal projections

dominating *to* and *on* respectively) and no barrier intervenes between P and DP. P is also a Case assignor.

Exercise 3

(a) violates the θ-criterion only. The subject position of *seems* is a position to which Case, but not a θ-role, is assigned. Therefore, *John* never receives a θ-role and the sentence violates the θ-criterion, but not the Case Filter.
(b) violates only the Case Filter. *It* is not assigned Case by non-finite (infinitival) I (since it is not a Case assignor) and the CP barrier blocks any potential Case from *promise* (which is not an ECM verb in any case). The subject position of *seem* is not assigned a θ-role, but expletive *it* does not require one. Therefore only the Case Filter is violated.
(c) violates both the Case Filter and the θ-criterion. *The room* requires both Case and a θ-role, but neither is available to the subject of a non-finite CP. Therefore both requirements are violated.
(d) violates the Case Filter. *Try* is not an ECM verb, and the CP barrier blocks Case in any event. However, since expletive *there* does not require a θ-role, there is no θ-criterion violation.

FURTHER EXERCISES

Exercise 1

(a) Draw the trees for the following sentences, being sure to indicate PROs where needed.
(b) Identify the Case assignors for all DPs.

1. I hope that John will go to the store.
2. I want to read the book.
3. Mary persuaded John to run for political office.
4. I need to return John's video.
5. John thinks that it is likely that Mary has left.
6. To run from danger can be a sensible decision.

Exercise 2

In the following sentences, what θ-roles are assigned to the subject and object? Use the tests in the chapter ('What X did was ...'/'What happened to X was...') to back up your intuitions.

1. Sue struck the ball cleanly.
2. John opened the door.
3. The book fell off the table.
4. Mary surprised Bill.
5. The court summons surprised Bill.

Exercise 3

Which of the following sentences involve ECM and which involve PRO?

1. The police believe John to be in London.
2. I forced John to go to the audition.
3. I expect John to go to the audition.
4. I know Mary to be an honest person.

Exercise 4

Which of the following predicates assign a θ-role to their subject position? Give actual linguistic data (i.e. grammatical and ungrammatical sentences) to support your conclusions:

(a) appear (when it takes a finite clause CP complement)
(b) consider
(c) strike (in the metaphorical sense of 'consider' and when it takes a CP complement)

Exercise 5

Modern Icelandic seems to allow objects to be dropped in certain conjoined contexts:

(1) Ég tók bókina og las.
I took the book and read
'I took the book and read (it).'

(a) Assuming that 'read' has the same argument structure in Icelandic and English, would you be suspicious if (1) were the only example of a 'dropped object'? Why? What other interpretation for (1) would there be?
(b) Would the grammaticality of (2) lead you to reconsider your answer in (a)? Why or why not?

(2) Ég elska þig og dái.
I love you and admire
'I love you and admire (you).'

Introduction to Binding Theory

5

In this chapter, we're going to look at one of the most investigated areas of syntax: Binding Theory. Binding Theory is concerned with the conditions under which two DPs in the sentence can be interpreted as referring to the same person or thing. Consider the sentences in (1), under the assumption that we're talking about our mutual friends John and Bill:

(1) a. John likes himself.
b. John likes him.
c. John thinks that Bill likes him.

In (1a), *himself* can only refer back to *John*, and not to *Bill* or some third party. In other words, (1a) can only mean *John likes John*, and not *John likes Bill*. In (1b), on the other hand, we have the reverse situation. *Him* cannot refer to *John*. It must refer to *Bill* (or someone else). However, in (1c), the DP *him* can refer back to John. (1c) means *John thinks that Bill likes* John (or, again, some third party). It can't mean *John thinks that Bill likes* Bill.

We'll see that we need to differentiate between three different kinds of DPs, which have three different syntactic requirements. The first kind of DPs are called *anaphors*. These are reflexives, such as *herself*, *ourselves,* and *themselves*, and reciprocals, such as *each other*. The second kind are *pronouns* – elements such as *them* and *you*. The third kind are called *R-expressions*, the most common of which are proper names like *Mary* and *Pat* or DPs like *the man*. The syntactic requirements of these three types of DPs are encoded in the three principles of Chomsky's (1981) Binding Theory: Principles A, B and C. We'll spend most of the time looking at anaphors, actually, but only because we'll need to introduce all of the relevant concepts along the way. Once we have Principle A under our belts, it'll take about two seconds to explain Principles B and C. Also, we'll really emphasize the theory-building part of things in this chapter. Instead of just telling you 'here's what Binding Theory is, now do it', we'll be building it (in a basic way) from the ground up.

5.1 Anaphors and Principle A

The first kind of DP that we'll be looking at is *anaphors*. There are two kinds of anaphors: reflexives, like *herself, themselves* and so on, and reciprocals, like *each other*. What is special about these elements is that they require that an *antecedent* be present; that is, another DP somewhere in the sentence that tells you who or what the anaphor is supposed to be referring to. It turns out that the antecedent can't be just anywhere (even if 'antecedent' makes it sound as though it has to precede), and in this section we'll try to discover what the constraints are.

The minimal IP requirement

Compare the sentences in (2), where subscripts are used to identify which person or entity the DPs in the sentence are intended to refer to:

(2) a. *Himself$_i$ left.
b. John$_i$ likes himself$_i$.
c. *John$_i$ likes themselves$_j$.

In (2) we have reflexives, but reciprocals pattern exactly the same way:

(3) a. *Each other$_i$ left.
b. They$_i$ like each other$_i$.
c. *John$_i$ likes each other$_j$.

(In fact, from now on we'll just use reflexives to illustrate the discussion. Everything that we say about reflexives is true of reciprocals, though.)

Let's see if we can start to draw up a theory to explain the distribution of anaphors based on (2) and (3). That is to say, under what circumstances do we expect that a sentence containing an anaphor will or will not be well formed? As always, we should start with the simplest hypothesis and work from there. On the basis of just (2) and (3), it looks like the simplest thing we can say is (4).

(4) An anaphor must have an antecedent which agrees with the anaphor in person, gender and number.

BEFORE CONTINUING

How does (4) account for (2a)–(2c) and (3a)–(3c)?

That would account for (2) and (3). In the (a) sentences, the anaphors *himself* and *each other* don't have any antecedent at all, and therefore (4) would predict

that they are ungrammatical. By contrast, in the (b) sentences, the anaphor has an antecedent which agrees with it in person, number and gender, and so our hypothesis in (4) would again make the correct prediction. The (c) sentences are like the (a) sentences actually. The anaphor requires an antecedent which agrees with it, but there isn't one in the sentence. The only possible antecedent is *John*, which is singular, and so can't agree with an anaphor with plural features.

However, it's clear straight away that we need to modify our hypothesis in (4). It's not just that there has to be an antecedent floating around. The antecedent has to be in the same sentence as the anaphor. To see why, take a look at the sentences in (5):

(5) I saw John$_i$ yesterday. *Himself$_i$ finally quit his job.

Here, we would say that *himself* has a discourse antecedent; that is, an antecedent which is not in the same sentence, but which exists elsewhere in the discourse. Discourse antecedents are perfectly OK for things like pronouns:

(6) I saw John$_i$ yesterday. He$_i$ finally quit his job.

However, it seems that they're not good enough for anaphors.

EXERCISE 1

Show that reciprocals do not allow discourse antecedents.

So it looks like we'll have to be more specific about the kinds of antecedents that anaphors can have:

(7) An anaphor must have an antecedent in the same sentence.

(From here on, by the way, we'll just use 'antecedent' to mean 'antecedent which agrees with the reflexive in person, number and gender'. You can't *be* an antecedent for a reflexive if you don't agree with it in person, number and gender.)
Anyway, so far so good, but not for long. Consider the sentence in (8):

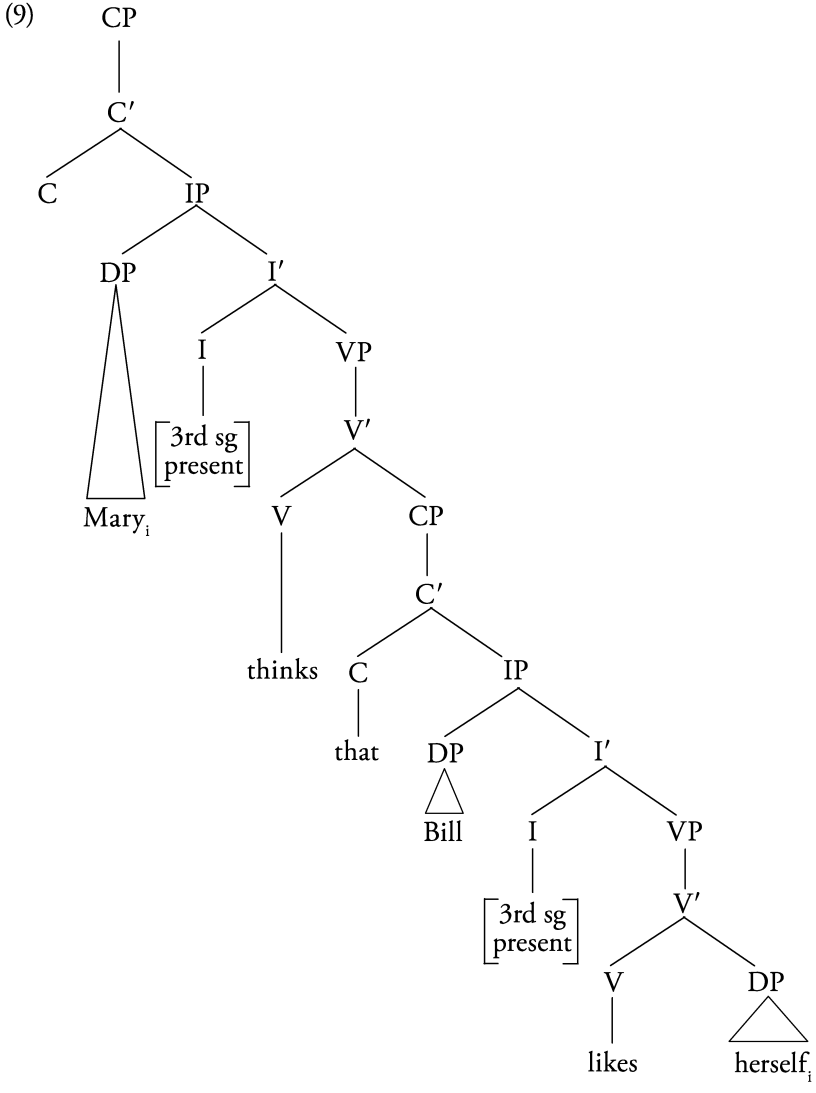
(8) *Mary$_i$ thinks that Bill$_j$ likes herself$_i$.

In this case, the anaphor *herself* is in the same sentence as the only possible antecedent, *Mary* (the whole thing is one big sentence), but (8) is still ungrammatical. (*Bill* is not a possible antecedent because it does not agree with the anaphor in gender.) It looks as though the antecedent must not only be in

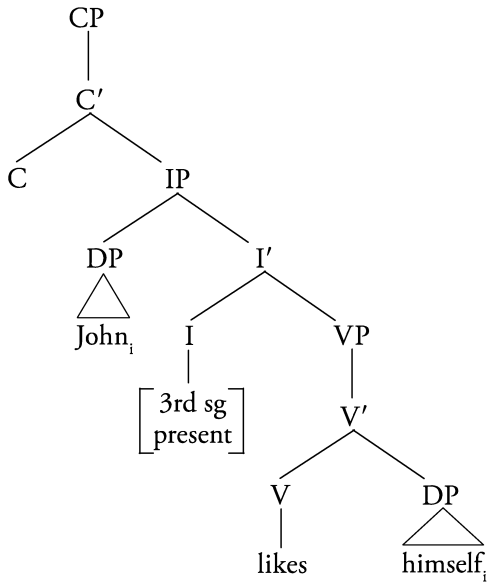
the same sentence as the anaphor, but also must be sufficiently 'close' to the anaphor.

But what does 'close' mean here? Well, notice that in (8) we have a case of embedding. The sentence in (8) *Mary thinks that Bill likes herself* has within it a second sentence, namely *Bill likes herself*. On the other hand, in (2b) we have only the one sentence *John likes himself*. If we strengthened the requirement on antecedents, so that the antecedent had to be in the same *minimal* sentence as the anaphor, then that would account for the difference. Only in (2b) is the antecedent in the same *minimal* IP as the anaphor (i.e. in the same clause). To see that, let's draw the tree for (8) and (2b).

(9)



(10)



As you can see from the trees, the minimal or closest IP which dominates *herself* in (9) is the embedded IP *Bill likes herself*. Although the anaphor is within that IP, the antecedent is not. On the other hand, in (10) there's only one IP, so the minimal IP dominating the anaphor *himself* is the IP *John likes himself*, and this time the antecedent *is* in the same minimal IP as the anaphor. So it looks like we need to make (7) even more specific:

(11) An anaphor must have an antecedent in the same minimal IP.

I want to take this opportunity to stress how important it is that we be as careful and explicit as possible when we're formulating our principles. There are ways of saying things which might look the same, but which have very different implications. For example, (12) sounds very much like the principle in (11):

(12) The minimal IP which contains the antecedent must also contain the anaphor.

BEFORE CONTINUING

Use (9) to show that (11) and (12) make different predictions and are therefore different statements.

(11) is the right principle. (12) is not. Consider again the tree representing (8), which we saw just above in (9). We want to make sure that our rule correctly predicts that (8) is ungrammatical, but (12) predicts the opposite.

The minimal IP containing the antecedent *Mary* is the IP *Mary thinks that Bill likes herself*. Is the anaphor within that IP? It sure is. So (12) incorrectly predicts that (8) should be grammatical. (11), on the other hand, makes the correct prediction. What's gone wrong is that we need to make sure that our algorithm looks for the *anaphor's* minimal IP, not the antecedent's. Otherwise, we might get something which looks right, but has things backwards.

The c-command requirement

Even though (11) and not (12) is the right principle to account for things so far, we're going to have to modify (11) still further. Consider (13):

(13) *[John$_i$'s mother]$_j$ likes himself.

BEFORE CONTINUING

Explain why (13) presents a problem for (11).

Nothing that we've said so far explains why (13) is ungrammatical. There's no embedding of clauses here. The minimal IP which dominates the anaphor is the matrix IP *John's mother likes himself*. The intended antecedent *John* is within that IP, and therefore (11) predicts that (13) should be grammatical.

(13) isn't really all that different from the grammatical (lb) above. However, there's a sharp difference in grammaticality. The crucial difference would seem to be that the antecedent of *himself* is not the subject, as it was in the grammatical (1b), but is instead buried inside the subject. We can confirm that we're focusing in on the right difference if we change the gender of the anaphor and make *John's mother* the antecedent:

(14) [John$_i$'s mother]$_j$ likes herself$_j$.

The antecedent of *herself* is now the whole subject *John's mother*, and not a subpart of the subject. As expected (14) is perfect.

How can we account for the contrast between (14) and (13)? We could simply say that anaphors must always have a subject as their antecedent. However, this would clearly be the wrong move to make, as (15) shows:

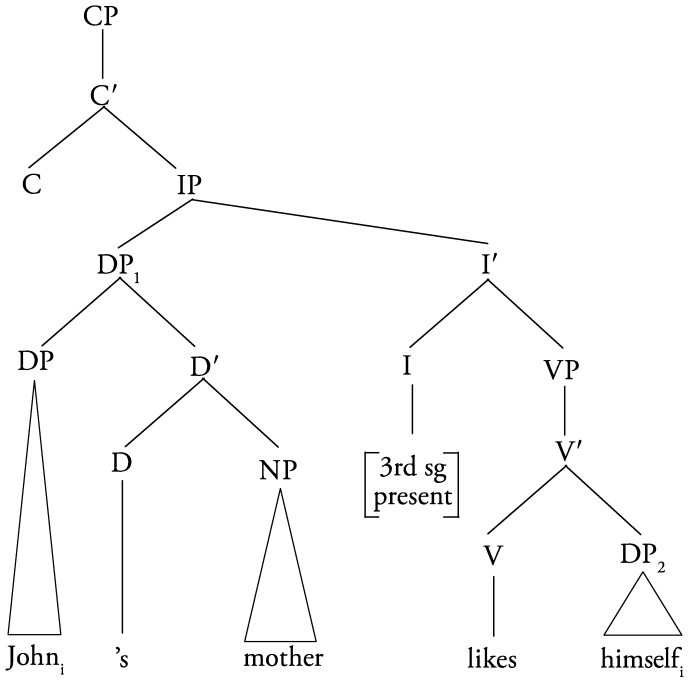
(15) John$_i$ gave Mary$_j$ a picture of herself$_j$.

If it was the case that anaphors just always had to refer back to the subject, we'd expect that (15) should be ungrammatical. However, since the anaphor *herself*

appears capable of referring to *Mary*, an indirect object, that doesn't seem like the right move to make.

As always, when we're in a situation like this, the first thing we should try to do is draw the tree. Perhaps there's a structural difference that we can look to for a solution.

(16)



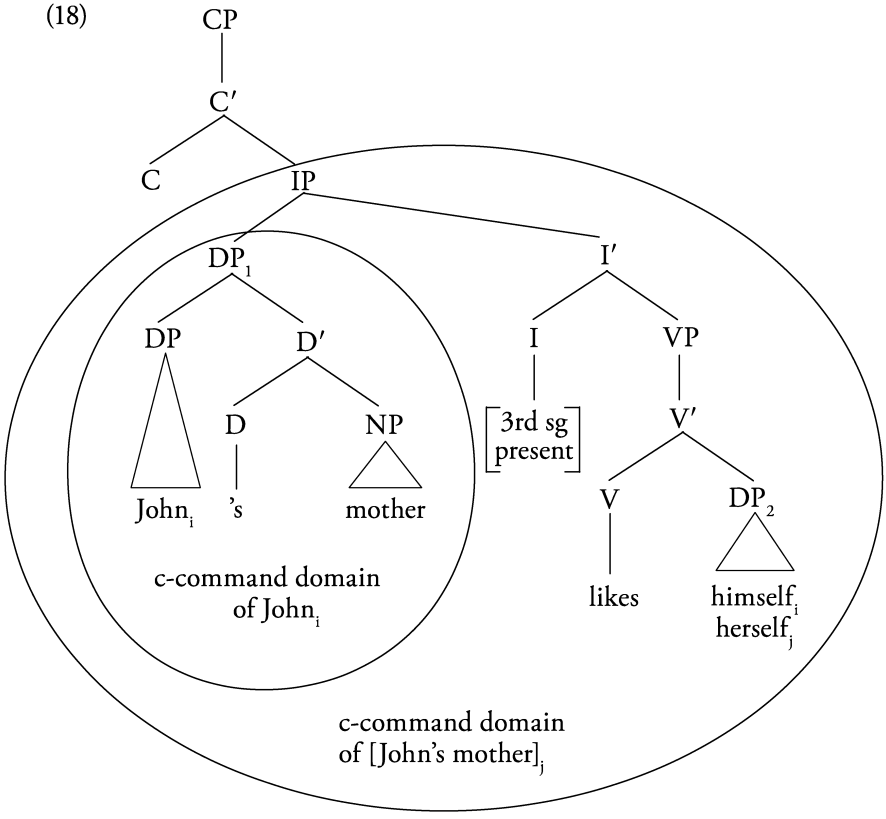
To put the problem in a nutshell, the relationship between *John's mother* and the anaphor in direct object position must be OK, but the relationship between *John* and a direct object anaphor seems not to be. What could be the difference?

Well, it looks like c-command might be relevant here. Recall the definition from Chapter 2:

(17) A node α c-commands another node β if and only if the first branching node which dominates α dominates β.

If we apply this definition to (13) and (14) above, it looks like c-command captures the difference between the two. In the ungrammatical (13), the antecedent *John* fails to c-command the anaphor *himself*. The first branching node which dominates *John* is DP_1, and DP_1 does not dominate *himself*. On the other hand, in the grammatical (14), the antecedent is the whole subject DP_1 *John's mother*.The first branching node which dominates DP_1 is the IP node. And *that* IP node does dominate the anaphor *herself*.

(18)



So c-command looks like it's definitely relevant, but we still need to hang on to our 'minimal IP' requirement. Otherwise, we'd be back to incorrectly predicting that *Mary thinks that John likes herself* is grammatical. Just like *John's mother*, the DP *Mary* will c-command the anaphor *herself*. The first branching node dominating *Mary* is the matrix IP, and the matrix IP dominates *herself*. So it looks like what we need to say is (19):

(19) An anaphor must be c-commanded by an antecedent which is in the same minimal IP.

Another – and, in fact, the usual – way of saying that an element β is c-commanded by another element α which is its antecedent is to say that α *binds* β. (That'll also make it clearer why this topic is called Binding Theory!) We can then replace (19) with (20), which just says exactly the same thing using the new terminology, including a definition for binding in (21):

(20) *Principle A (preliminary)*
An anaphor must be bound in the minimal IP which contains it.

(21) α binds β if α c-commands β and is coindexed with β.

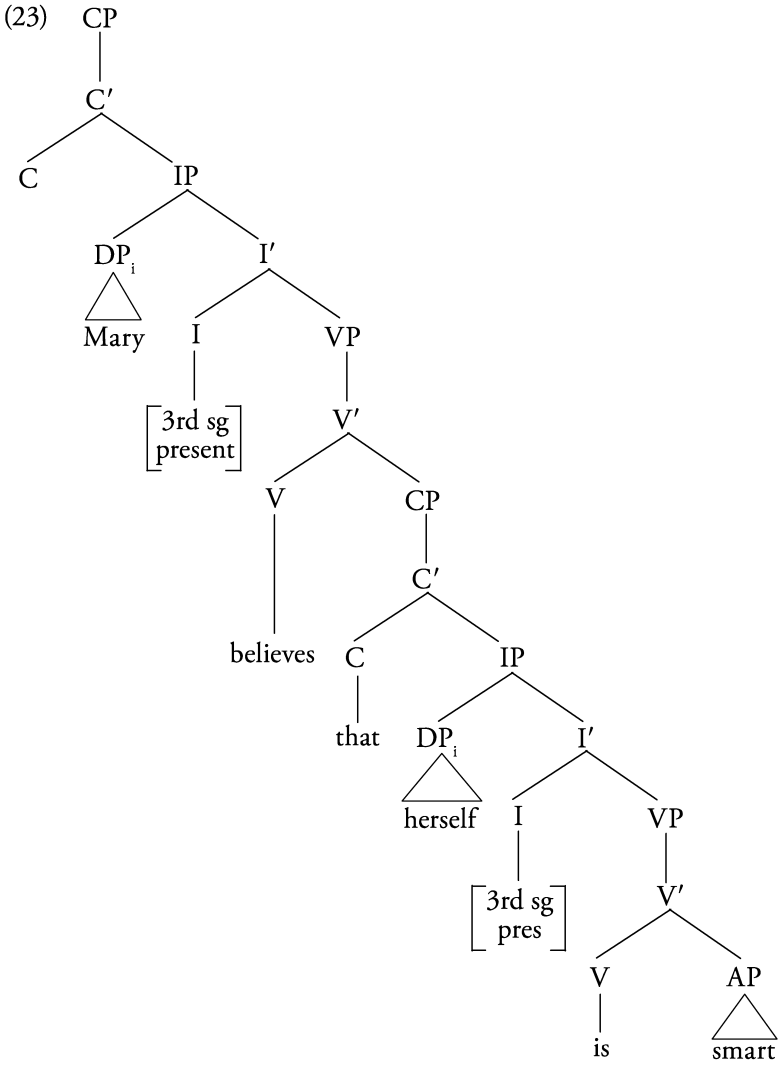
The problem of ECM contexts

ECM configurations pose a problem for the anaphor rule in (20) as we've developed it so far. Consider the contrast in (22):

(22) a. *Mary$_i$ believes that herself$_i$ is smart.
b. Mary$_i$ believes herself$_i$ to be smart.

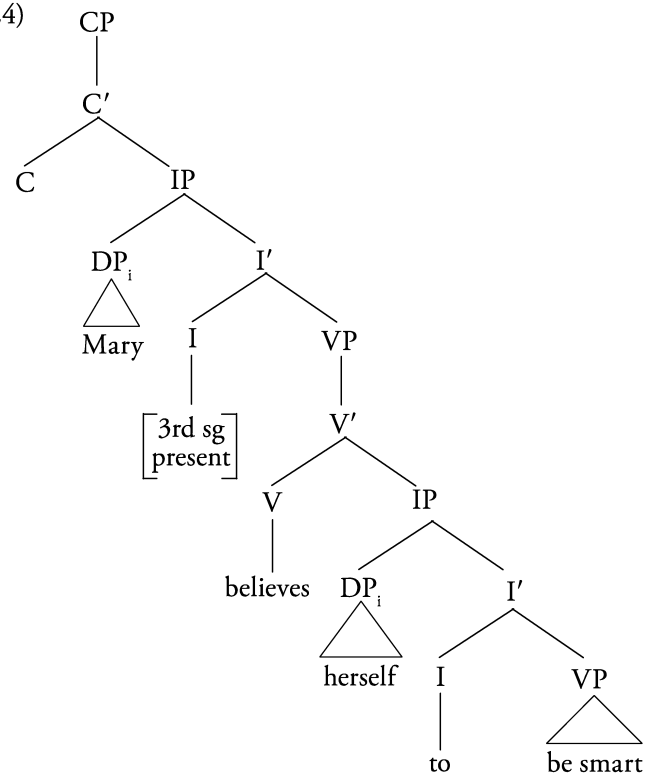
(22a) behaves as predicted.

(23)



Although *Mary* c-commands *herself*, the minimal IP containing the anaphor is the embedded IP *herself is smart*. Since the antecedent *Mary* is not within this IP, the sentence is correctly predicted to be ungrammatical. However, in (22b), the minimal IP containing the anaphor is *herself to be smart*.

(24)



The antecedent *Mary* is not within this IP, and (22b) is also predicted to be ungrammatical. But in this case that's the wrong prediction.

So what's different about ECM complements, and how can we use that difference to explain the contrast in (22)? The only structural difference between (22a) and (22b) is that in the ungrammatical (22a) *believe* takes a CP complement (even if the complementizer is not expressed, don't forget) whereas the complement is an IP in the grammatical (22b). So whatever's going on is surely related to that difference.

As we saw in Chapter 4, the fact that ECM infinitives are IPs and not CPs means that the verb *believe* is able to assign accusative Case to the subject of the infinitive, so you get contrasts like (25):

(25) a. I believe her to be smart.
b. *I believe she to be smart.
c. *I believe that her is smart.
d. I believe that she is smart.

Let's suppose that it's the verb's ability to reach down into the subject position of the infinitive clause to assign Case to an element there which is what allows an anaphor in that position to look up to the higher clause to find its antecedent. It's as if the verb's Case assigning creates a bridge between the subject position and the higher clause.

We need to work this idea into our theory, but, as it turns out, it's not quite Case assignment per se that's relevant. It's a slightly more general notion called *government*, of which Case assignment is a special case. Government is defined as in (26):

(26) α governs β if and only if
(a) α is a head, and
(b) α m-commands β, and
(c) no barrier intervenes (where any maximal projection except infinitival IP is a barrier).

Does this look familiar? It's the definition of Case assignment from the previous chapter, with one minor change. In (a), instead of saying 'α is a Case assignor', it says 'α is a head'. So heads always govern everything that they m-command (provided no barrier intervenes), and if the head assigns Case, then the head can also assign Case to something that they govern.

So, in the ECM case, because the verb governs the subject of the infinitive, an anaphor in that position can look up to the matrix clause for an antecedent. Notice that in other cases, where we don't want the anaphor to be able to reach up to the matrix clause, government from above isn't possible:

(27) a. *Mary$_i$ believes [$_{IP}$ John to [$_{VP}$ like herself$_i$]].
b. *Mary$_i$ thinks [$_{CP}$ that herself$_i$ is smart].

In (27a), the anaphor is in the direct object position of an ECM infinitive, and government by *believe* is blocked. The infinitival IP may be defective as a barrier to government, but we've still got the barrier constituted by VP. In (27b), government from outside the embedded CP is straightforwardly blocked by the CP, which acts as a barrier.

So it looks as though when the anaphor in question is governed from outside the IP that it's in, it's allowed to look up to that higher IP for an antecedent. Let's incorporate this into our rule:

(28) *Principle A of the Binding Theory (version 2)*
An anaphor must be bound in the minimal IP which contains:
(a) the anaphor, and
(b) a governor for the anaphor.

The domain in which an anaphor has to find its antecedent is usually called the 'governing category' of the anaphor, so we can rewrite (28) completely equivalently as (29), with the appropriate definitions (with 'bound' repeated from above):

(29) *Principle A of the Binding Theory (final version)*
An anaphor must be bound in its governing category.

(30) α *binds* β if α c-commands and is coindexed with β.

(31) The *governing category* for an element α is the minimal IP containing:
(a) the element α, and
(b) a governor for α.

Although we now have the final statement of Principle A of the Binding Theory in (29), it's going to turn out that we need to modify the definition of governing category just a little bit more.

'Subjects', accessible and otherwise

The last notion that we'll need to incorporate in order to give a complete account of anaphors is that of 'subject'. To see why that's going to be necessary, compare the sentences in (32):

(32) a. John$_i$ liked the drawing of himself$_i$.
b. *John$_i$ liked Mary's$_j$ drawing of himself$_i$.
c. John liked Mary's$_j$ drawing of herself$_j$.

Whatever's going on, the contrast is pretty clear. It looks like the anaphor in (32b) and (32c) wants to have *Mary* as its antecedent, not *John*, in contrast with (32a), which is happy with *John* as its antecedent. The only difference between the two is that the definite article *the* in (32a) is replaced by *Mary's* in (32b) and (32c). What could be the reason for the contrast we're seeing?

Well, notice that DPs which have a DP specifier have some similarities to sentences. For example, the DPs in (33) contain roughly the same information as the full sentences in (34), a point we made when we introduced DPs in Chapter 3:

(33) a. Mary's drawing of him.
b. Bill's destruction of Carthage.

(34) a. Mary drew him.
b. Bill destroyed Carthage.

In (34), *Mary* and *Bill* each receive an Agent θ-role and are interpreted as the subjects of their respective sentences. *Him* and *Carthage*, the complements of the verbs in question, receive θ-roles appropriate for their direct object status. If *Mary* receives the Agent θ-role in (34a), there seems to be no reason not to assign *Mary* that same role in (33a). After all, *Mary* is understood as being in exactly the same relation to *drawing* as she is to *drew* in (34a). So even though (33a) is not a sentence, but a DP, it seems reasonable to say that *Mary* is functioning like a subject in both cases.

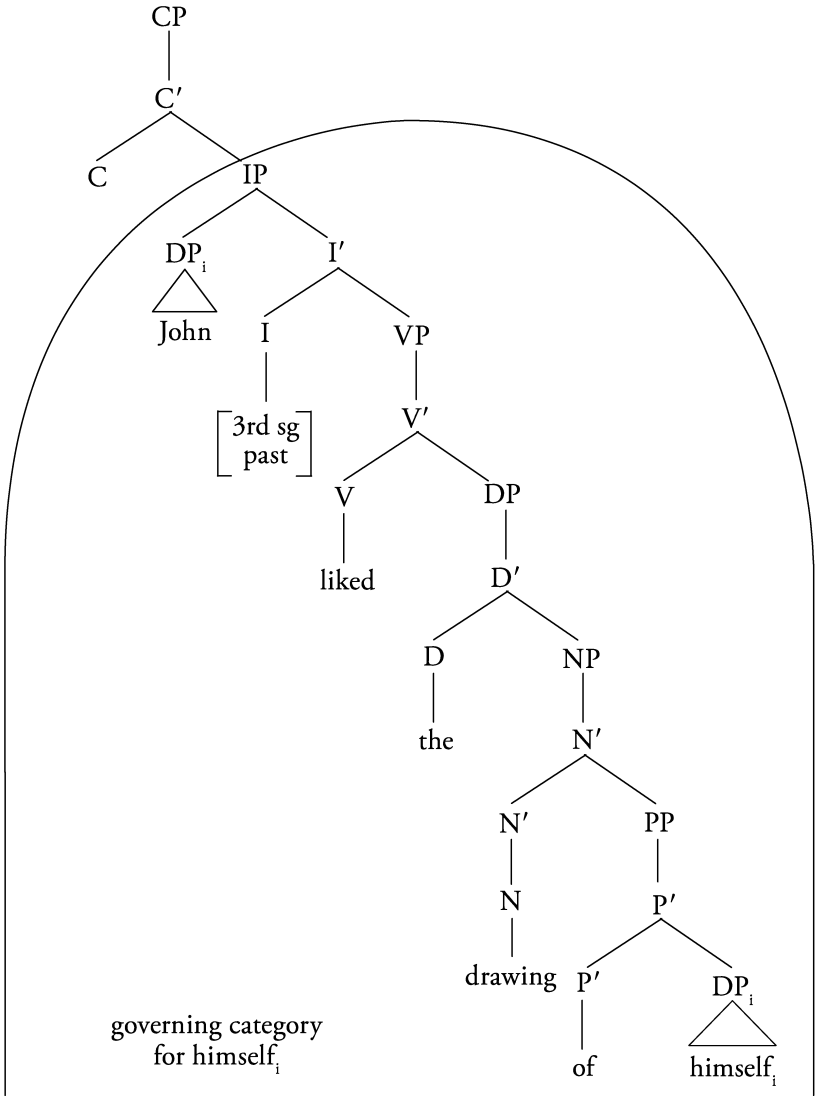
Let's try to develop this idea. Maybe, all along, when we were saying that the antecedent has to look for its antecedent in the appropriate minimal IP, it wasn't really IP per se that was relevant. Maybe the anaphor needs to find its antecedent in the minimal XP that has a *subject*. Under normal circumstances, it's usually only IPs that have subjects, so most of the time we'll be looking for IP. However, in those cases where you have a DP which has a 'subject', the

domain in which the anaphor must have its antecedent is that DP. This means we'll need to revise (31) a little bit:

(35) The *governing category* for an element a is the minimal XP containing
(a) the element α,
(b) a governor for α, and
(c) a subject.

We can now explain the pattern of grammaticality that we saw in (32). Consider first (36), which is the tree for (32a):

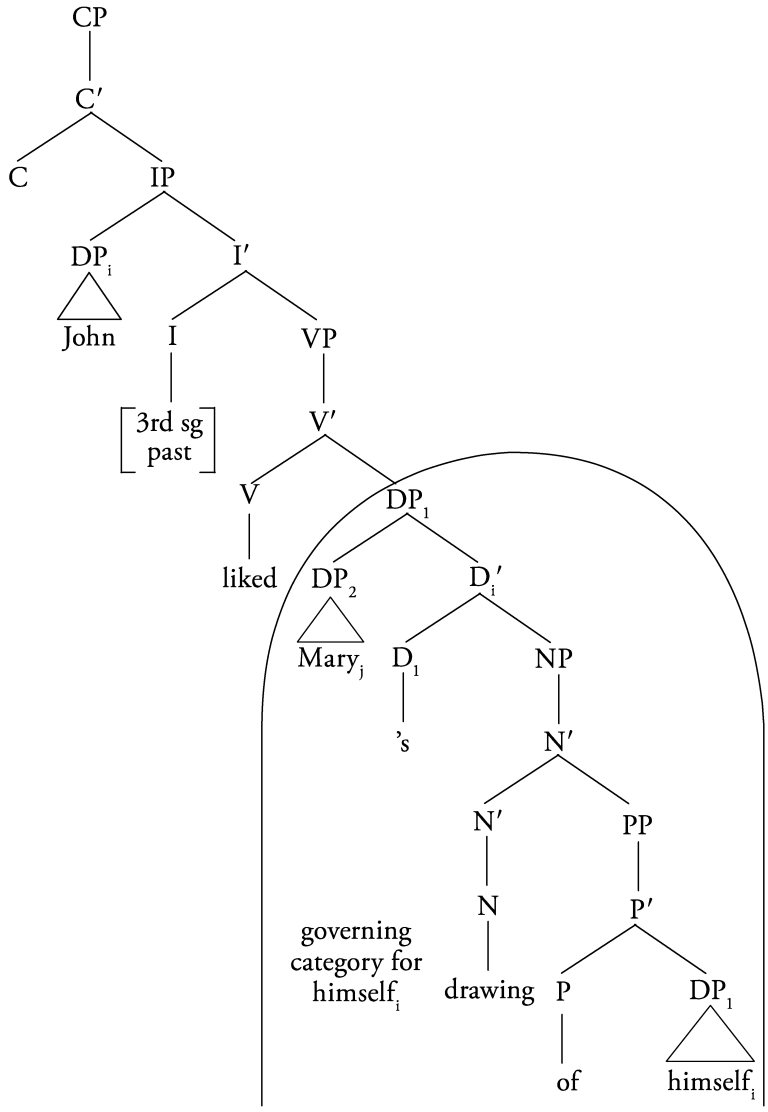
(36)



In (32a), the DP *the drawing of himself* does not have a subject. That DP doesn't have a DP specifier or indeed a specifier of any kind. Therefore, to find the governing category, we need to keep looking higher. The next higher XP is VP, but that has no specifier of any kind either. The next XP is IP, and we finally have a category with a subject: *John*. *John* is immediately dominated by IP and a sister to I′. Since the anaphor *himself* has an antecedent (*John*) in its governing category, Principle A predicts that (32a) is grammatical.

But notice how things change when we move from *the drawing of himself* to *Mary's drawing of himself*.

(37)



In (32b) we have two DPs, *Mary's drawing of himself* and *Mary*, which I've marked as DP_1 and DP_2 respectively. (There's no significance to this; it's

just for reference.) DP_1 *does* have a subject, namely DP_2. *Mary* is immediately dominated by DP_1 and is a sister to D'_1. That means that the minimal XP containing *himself*, a governor for *himself* and a subject is DP_1 and that therefore DP_1 is the governing category for the anaphor *himself*. And that means that *himself* is no longer able to satisfy Principle A. *John*, the antecedent, is not in the governing category for *himself*, and we predict that (32b) is ungrammatical.

So far, so good. All we need to do now is make one tiny change to the notion 'subject' to take care of the following situation, and we'll be all set. Consider (38):

(38) $John_i$ thinks that a picture of $himself_i$ is on sale.

BEFORE CONTINUING

Show that (35) wrongly predicts that (38) is ungrammatical.

(38) contrasts sharply with other cases where the anaphor is in an embedded CP, like (39), which we've seen before, and (40):

(39) *$John_i$ thinks that Mary bought a picture of $himself_i$.

(40) *$John_i$ thinks that $himself_i$ should win the election.

However, (35) as it currently stands predicts that (38) to (40) should all be ungrammatical. (38) is the surprising one, so let's take a closer look at how (35) wrongly predicts it to be ungrammatical. Once we clearly understand where we're going wrong, we can think about how to fix things.

We're looking for the minimal XP which contains the anaphor, the anaphor's governor and a subject. The preposition *of* is the governor of the anaphor, so the PP *of himself* is the minimal XP which contains the anaphor and the anaphor's governor. However, the PP does not contain a subject. Continuing up the tree, we eventually get to the embedded IP *a picture of himself is on sale*. Does this IP have a subject? Yes indeed. The subject is the DP *a picture of himself*. Therefore, the minimal XP which contains the anaphor, the anaphor's governor and a subject is the embedded IP. However, the antecedent for the anaphor, *John*, is not within the embedded IP. We therefore predict that (38) should be ungrammatical, and that's not the answer we want.

Clearly, we need to make some change to our rule in (35) in order to account for the difference between (38) on the one hand, and (39) and (40) on the other. First of all, let's think about where in (35) we have room to make changes. It's hard to see how we could make a change to (a) or (b) in order to account for (38) that wouldn't incorrectly predict that (39) should be grammatical too. (c) on the other hand might be more promising. At least the subject is

different in (38) vs (39). Perhaps that's the hook that we need to hang the distinction on.

Notice that in (38) the anaphor is inside the subject of the embedded clause *a picture of himself*, whereas in (39) the anaphor is inside the direct object. Perhaps the subject position is special in some way. They can't be too special, though, because as (40) shows if the anaphor is itself a subject, then the sentence is ungrammatical, just as we would expect. So the key thing must be that the anaphor in (38) is a subpart *inside* the subject. How can we implement something that will correctly predict that (38) is grammatical while ruling out (39) and (40)? Basically, what we need is something that will make the embedded IP the relevant domain in (39) and (40), but allow the anaphor to look up to the matrix clause in the grammatical (38).

Chomsky (1981) introduced the notion of the 'accessible' subject as a way of solving the problems posed by (38)–(40). In order for a subject to count as an 'accessible' subject for an anaphor, it has to meet the following condition:

(41) α is an accessible subject for an anaphor β if and only if (hypothetical) coindexation between the anaphor and the subject violates no grammatical principle.

That sounds complicated, but it's actually straightforward once you see it in action. On a personal level, I'm not sure how persuasive I find Chomsky's solution. It seems very *ad hoc*. However, it does look as though *something* needs to be said, and since I don't have anything better to suggest, I can't complain too loudly.

Let's start with the grammatical (38). What we want to make sure happens is that the embedded IP is *not* the relevant domain for the anaphor. Instead, we want the matrix IP to be the governing category. Let's see how (41) helps us. The minimal XP containing the anaphor and the anaphor's governor is still the NP *picture of himself*, so now we're looking for an accessible subject. The embedded IP has a subject, namely *a picture of himself*, but we need to see whether that subject is accessible or not. In order to do that, we need to see whether hypothetical coindexing of the proposed subject and the anaphor violates any grammatical principle. We say 'hypothetical' here because we're not proposing that the two things are supposed to be actually coindexed. A picture of something is inanimate, but a person is animate, so it wouldn't make any sense to literally coindex the two. This is just a hypothetical test of a situation – if we were to do it, would there be a problem?

Let's see what happens when we do that. If we hypothetically coindexed the anaphor *himself* with the proposed subject *a picture of himself*, we get the following configuration:

(42) John_i thinks that [a picture of himself_i]$_i$ is on sale.

Does this violate any grammatical principle? Well, Chomsky says that it does actually. The principle is called the i-within-i filter, and basically rules out any configuration in which a smaller subpart of something is coindexed with the larger whole thing, as illustrated by (43):

(43) *I-Within-I Filter*
$*[\ldots X_i \ldots]_i$

You need something like (43) to avoid referential circularity. Basically, as a general rule, if you're trying to figure out what some expression refers to, you need to know what all of the component parts refer to. So in (43), in order to figure out what the whole thing refers to, part of what you need to figure out is what X refers to. However, if X and the whole expression are coindexed, you can't figure out what X refers to until you know what the whole thing refers to. But to do that you need to know what X refers to, and so on. You go round in circles. (43) rules out these configurations.

To give a slightly less abstract illustration of the same point, consider the sentence in (44), which also contains an i-within-i violation:

(44) *[The picture of $it_i]_i$ is on the table.

In order to determine what *the picture of it* refers to, we need to know what *it* refers to. But in order to know what *it* refers to, we need to know what *the picture of it* refers to, and we're back where we started. If you try to imagine what (44) could mean, you end up with an infinite regress of pictures going into the distance, as if you had two mirrors reflecting each other. If, on the other hand, we have an indexation which does not create an i-within-i configuration, there's no problem:

(45) [The picture of $it_j]_i$ is on the table.

In (45), *it* has index 'j', and so refers to some previously mentioned entity, like your new computer or something. There's no referential circularity here, and thus no problem with interpretation.

So, let's return to (38), our binding theory problem:

(38) $John_i$ thinks that a picture of $himself_i$ is on sale.

The key intuition behind the notion 'accessible' subject is that *a picture of himself* can't serve as a subject for the anaphor *himself*, because that would require coindexation between *himself* and *picture of himself*, which is an i-within-i violation. Therefore, although the minimal XP which contains the anaphor, its governor and a subject is the embedded IP, the subject of that IP isn't accessible to the anaphor. Therefore, we are allowed (and required) to look higher in order to find an antecedent for *himself*.

Once we've gone past the embedded IP, the next category which has any kind of subject at all is the matrix IP *John thinks that a picture of himself is on sale*. The subject is of course *John*. Would hypothetical coindexation of *John* and *himself* violate any grammatical principle? In particular, would it create an i-within-i configuration? No. Neither *John* nor *himself* is contained within the other, so there's no problem. *John* is a subject which is accessible to *himself*. Therefore, the minimal XP which contains the anaphor *himself*, a governor of *himself* and a subject which is accessible to *himself* is the matrix IP. Do we find

an antecedent for *himself* within this category? Yes, it's *John*. Therefore, the sentence in (38) is predicted to be grammatical.

Now that we've found a way to make (38) grammatical, let's make sure that we haven't lost anything along the way by taking a look at (39) and (40) again, both of which were ungrammatical. We need to make sure that our changes don't result in our incorrectly predicting that those two sentences are grammatical.

(39) *John$_i$ thinks that Mary bought a picture of himself$_i$.

(40) *John$_i$ thinks that himself$_i$ should win the election.

Just as with (38), the minimal XP containing the anaphor and a governor for the anaphor is the NP *picture of himself*. Moving upward in search of a subject, the next category which we need to take a look at is the embedded IP *Mary bought a picture of himself*. Does hypothetical coindexing between *Mary* and *himself* create an i-within-i violation? No. Neither *Mary* nor *himself* is a subpart of the other. Therefore, *Mary* is an accessible subject for *himself*. So that means that the minimal XP containing the anaphor *himself*, a governor for *himself* and a subject which is accessible to *himself* is the embedded IP *Mary bought a picture of himself*. Since *himself* has no *actual* antecedent in the embedded IP, (39) is still correctly predicted to be ungrammatical.

Now let's turn to (40). Here the situation might be a little bit more confusing. The difference between (38) and (40) is that in (38) the anaphor is *inside* the subject DP, whereas in (40) it *is* the subject. The governor for himself is now the embedded I, since it's assigning *himself* nominative Case. So the minimal XP containing the anaphor and its governor is the embedded IP *himself should win the election*. Does this IP have a subject? Yes, it's *himself*.

Now we need to answer the following question: can *himself* serve as an accessible subject for itself? We know the answer that we want. We want the answer to be 'yes'. That way, the anaphor will need to find a c-commanding antecedent in the embedded clause. Since there isn't one, we would then correctly predict that (40) is ungrammatical. What would be the justification for saying that an anaphor can serve as an accessible subject for itself? Well, if we were to consider the issue of hypothetical coindexation, what we'd be doing would be just giving something the same index over again. It's hard to see why that would be a problem. I suppose it could be taken to be redundant or something, but it certainly wouldn't create an i-within-i configuration. You're not coindexing something with a *subpart* of it, which is what the i-within-i filter prohibits.

Therefore, the minimal IP which contains *himself*, a governor for *himself* and a subject which is accessible to *himself* is the embedded IP. *Himself* does not have an antecedent within this IP, and (40) is correctly predicted to be ungrammatical.

So it looks like we're all set; adding the 'accessibility' requirement to the subject would seem to correctly allow the grammatical (38), while ruling out the ungrammatical (39) and (40). So let's incorporate that into our complete version of Principle A:

(46) *Principle A of the Binding Theory (final version)*
An anaphor must be bound in its governing category.

(47) α *binds* β if α c-commands and is coindexed with β.

(48) The *governing category* for an element α is the minimal XP containing
(a) the element α,
(b) a governor for α, and
(c) a subject which is accessible to α.

(49) α is a subject which is accessible to β if and only if (hypothetical) coindexation between the anaphor and the subject violates no grammatical principle.

(50) *I-Within-I Filter*
$*[\ldots X_i \ldots]_i$

5.2 Pronouns and Principle B

Fortunately, as promised, now that we have Principle A under our belts, Principles B and C are a piece of cake. Principle B of the Binding Theory concerns pronouns. Unlike anaphors, pronouns will be familiar to you from traditional grammar instruction. They are words like *we*, *her*, *them*, *I* and so on.

The complementarity of pronouns and anaphors

We mentioned, in passing, one difference between pronouns and anaphors at the beginning of the previous section, when we were talking about discourse antecedents. It turns out that anaphors don't allow discourse antecedents, but pronouns do. So we have contrasts like the one in (51):

(51) I saw John$_i$ yesterday. He$_i$/*Himself$_i$ just quit his job.

If we look at more of the examples we discussed at the beginning of the anaphor section, we see more in the way of complementary distribution. That is, for any given syntactic position, if an anaphor can be there, then a pronoun can't be, and vice versa. (52a), for example, contrasts with (52b):

(52) a. John$_i$ likes himself$_i$.
b. *John$_i$ likes him$_i$.

Although I've mentioned this issue before, it's worth repeating. (52b) illustrates the importance of considering the indexation when looking at example sentences in a Binding Theory context. (52b) is only ungrammatical on the interpretation where *John* and *him* are intended to refer to the same person. Put another way, (52b) can't mean the same thing as (52a). On the other hand, (53) is fine:

(53) John$_i$ likes him$_j$.

The same complementarity seems to run through other examples as well:

(54) a. *John$_i$ thinks that Mary$_j$ likes himself$_i$.
b. John$_i$ thinks that Mary$_j$ likes him$_i$.

(55) a. *[John's$_i$ mother]$_j$ likes himself$_i$.
b. [John's$_i$ mother]$_j$ likes him$_i$.

(56) a. Mary$_i$ believes herself$_i$ to be smart.
b. *Mary$_i$ believes her$_i$ to be smart.

(57) a. *John$_i$ liked Mary's$_j$ picture of himself$_i$.
b. John$_i$ liked Mary's$_j$ picture of him$_i$.

(58) a. John$_i$ thinks that a picture of himself$_i$ is on sale.
b. *John$_i$ thinks that a picture of him$_i$ is on sale.

So, if anaphors and pronouns are in complementary distribution, then it would seem that we only need to make a minor modification to Principle A in order to get Principle B, the principle that deals with pronouns. If an anaphor must be bound in its governing category, then a pronoun must be *not* bound (i.e. 'free') in its governing category:

(59) *Principle B of the Binding Theory*
A pronoun must be free in its governing category (where 'free' = 'not bound').

Therefore, in precisely the same domain in which an anaphor *must* have an antecedent, a pronoun *cannot*, and we build in the complementary distribution which seems to hold between pronouns and anaphors.

The important thing to notice about (59) is that it doesn't *require* that a pronoun have an antecedent. It just says that, if it does, then the antecedent has to be outside the pronoun's governing category. So (60), for example, is ambiguous between two readings: one where *him* refers to John, and another where *him* refers to a third, unmentioned person:

(60) John$_i$ thinks that Mary$_j$ likes him$_{i/k}$.

The governing category for the pronoun will be the IP *Mary likes him*. Therefore, if *him* has an antecedent, then it must be outside *him's* governing category. This allows *him* to have index 'i', referring back to *John*, since *John* is not within the IP *Mary likes him*. If, on the other hand, *him* has index 'k', referring to someone who is neither *John* nor *Mary*, then *him* is also free in its governing category (and in fact free everywhere). It's not bound by a c-commanding antecedent in its governing category because there *is* no c-commanding antecedent. So, pronouns can actually satisfy Principle B of the Binding Theory in two ways, either (1) by having a c-commanding antecedent which

is outside of the pronoun's governing category or (2) by not having a c-commanding antecedent at all.

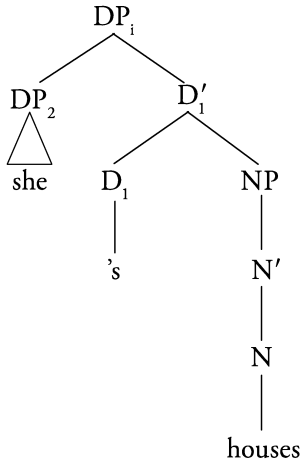
Governing categories and the analysis of possessive pronouns

Back in Chapter 3, we proposed what was perhaps a slightly unintuitive way of accounting for possessive pronouns such as *her*, *our* and *their*. We noticed that, in a possessive construction in English, the DP which is the possessor seems to be always marked by *'s* when it precedes the noun which is possessed.

(61) Mary's houses/that person's houses/two friend's houses

In order to bring out the parallelism with other DPs, and provide a uniform X′-Theory analysis of possessive pronouns, we suggested that the morphology of English was obscuring an underlying similarity. As far as the syntax was concerned, *her houses*, for example, is exactly the same as *Mary's houses*, because *her houses* is really *she's houses*, with the pronoun *she* as the specifier of the DP, and the *'s* as the determiner head.

(62)

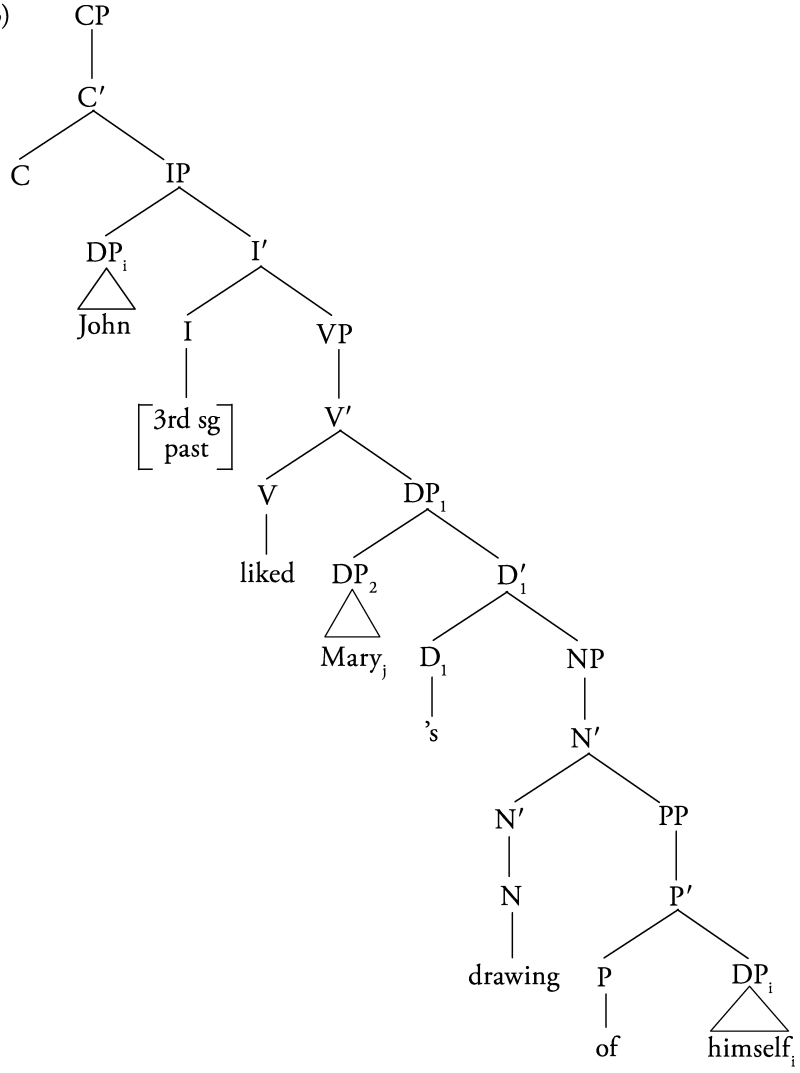


Then, the *'s* cliticizes onto the pronoun, and a low-level rule of English morphology 'spells out' *she's* as 'her' (and *we's* as 'our', and *they's* as 'their', etc.). Although the argument in Chapter 3 was based on parallelism and theory-neatness, Binding Theory actually provides us with an empirical argument that possessive pronouns are treated by the syntax as if they were 'broken down' in this way.

In Section 5.1 we talked about the notion 'subject' as it relates to the definition of governing category. In this context, any XP that has a DP specifier (immediately dominated by XP and a sister to X′) has a subject. Specifiers of

IP are the most common, but DPs can also have a subject when the DP itself has another DP as its specifier, as in (37), which I've repeated below:

(63)



DP$_2$, *Mary* is the subject of DP$_1$, *Mary's drawings of himself*, and so the governing category for the anaphor is restricted to the DP, rather than being the whole IP.

If we're right to analyse DPs containing possessive pronouns in the same way, then we should expect that Binding Theory treats these DPs as having subjects, and that looks to be the case. Possessive pronouns seem to pattern precisely with overt DP:

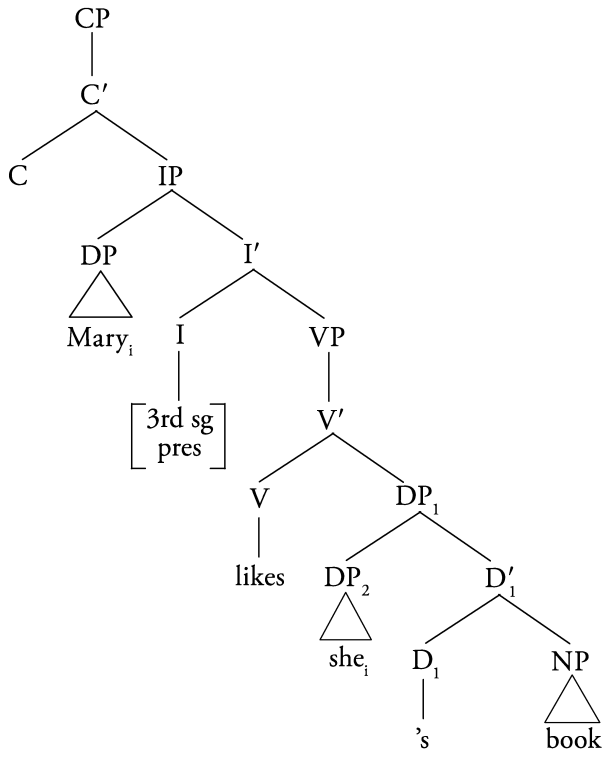
(64) a. John$_i$ liked Mary$_j$'s drawing of *himself$_i$/him$_i$.
b. John$_i$ liked her$_j$ drawing of *himself$_i$/him$_i$.

If *her* is really *she's*, then the two DPs are identical in terms of syntactic structure. The tree for (64b) is identical to the tree in (63), except for the fact that *Mary* is replaced by *she*.

Interestingly, this analysis also straightforwardly explains the grammaticality of sentences like (65), which under different assumptions about the structure of possessive DPs have been a problem:

(65) Mary_i likes her_i book.

(66)



Just as in the previous cases, the DP *her book* contains a subject for Binding Theory purposes, and so the governing category for the pronoun *she* is the DP *her [she's] book*. Therefore, *Mary* is outside the pronoun's governing category, and Principle B is satisfied.

5.3 R-expressions and Principle C

Having looked at both anaphors and pronouns, it's time now to move on to the last of the three kinds of DPs constrained by Binding Theory, which are called R-expressions. Proper names like *John* and *Mary* are the canonical examples of R-expressions. Not only do they not need an antecedent, they also pick out a specific person or thing in the universe of discourse. If we've been discussing

our friends Bob, John and Mary, and I say 'John left', you know exactly who I'm talking about. If, on the other hand, I use a pronoun, and start by saying 'He left', you don't know which of the two male friends I have in mind. This is where the 'R' in 'R-expressions' comes from, by the way. It stands for 'referring'. R-expressions are 'referring expressions'.

At first glance, R-expressions look like they're going to pattern with pronouns. Both (67a) and (67b) are ungrammatical.

(67) a. *John$_i$ likes him$_i$.
b. *John$_i$ likes John$_i$.

With respect to (67b) in particular, it's important to bear in mind that we're talking about the normal way to express in English the idea that John likes himself. (67b) becomes a lot better if one *John* or the other is contrastively stressed, but something different is going on there and we'll put that interpretation off to the side. We're interested in the neutral, unmarked way to express 'John self-likes', which is not (67a) or (67b).

When we look a little further, it becomes clear that R-expressions are subject to a slightly stronger requirement. In particular, it doesn't look like moving the binding antecedent to a position outside of the governing category does any good:

(68) *John$_i$ thinks that Mary$_i$ likes John$_i$.

Just as we saw with the previous examples involving pronouns, the governing category for the R-expression *John* in direct object position of the embedded clause would be the embedded IP *Mary likes John*. The antecedent for *John* is not within the embedded IP. It would therefore seem that R-expressions must be different from pronouns with respect to the Binding Theory. Notice as well that the effect is identical when the antecedent is a pronoun rather than another R-expression, so what is wrong with (68) has nothing to do with repeating a previously mentioned R-expression or something:

(69) a. *He$_i$ likes John$_i$.
b. *He$_i$ thinks that Mary$_j$ likes John$_i$.

One possible approach might be to see whether R-expressions just have a wider governing category than anaphors or pronouns. However, it quickly becomes apparent that that solution won't work. It looks like you can have as many IPs in the sentence as you like and the sentence remains ungrammatical:

(70) a. *John$_i$ believes that Mary$_j$ said that Alice$_k$ likes John$_i$.
b. *John$_i$ believes that Mary$_j$ said that Alice$_k$ wondered whether Jill$_l$ liked John$_i$.

However, it does look as though c-command is still playing a role. If the R-expression isn't c-commanded by its antecedent, the sentence improves considerably:

(71) [John's$_i$ mother]$_j$ likes John$_i$.

So, it looks like binding is relevant, but not governing categories. Nothing hard there. All we have to say is (72):

(72) *Principle C of the Binding Theory*
An R-expression must be free (again, 'free' = 'not bound').

This way, any configuration where an R-expression is c-commanded by a coindexed antecedent will be ruled out.

In addition to proper names, like *John*, there are other kinds of DPs which seem to function as R-expressions with respect to the Binding Theory. For example, DPs of the kind which in the philosophy literature would be called definite descriptions also function as R-expressions. An example of this would be a DP like *the man I saw yesterday*, where there is a specific person in mind:

(73) *He$_i$ told me that [the man I saw yesterday]$_i$ was a local politician.

Compare (73) with (74):

(74) [The man I saw yesterday]$_i$ told me that he$_i$ was a local politician.

In (74) the R-expression is not c-commanded by the pronoun and so no Principle C violation arises.

However, my favourite R-expressions are a class of elements called *epithets*. These are DPs like *the fool, the bastard* or *the president*, which can be used as a way of descriptively referring back to some previously mentioned person:

(75) I saw John$_i$ yesterday. The bastard$_i$ didn't have the 10,000 pesetas he owes me.

These elements might feel more like pronouns, in that they pick out a specific person only when they have an antecedent. However, (76) shows that epithets obey Principle C of the Binding Theory:

(76) *John$_i$ said that the bastard$_i$ was going to the store.

The epithet can't be interpreted as co-referential with *John*. In other words, (76) can't be interpreted as the grammatical (77):

(77) John$_i$ said that he$_i$ was going to the store.

Similarly, sentences like (78) are ill formed:

(78) *He$_i$ said that the president$_i$ would be addressing the nation tonight.

Just as with (73) and (74), reversing the pronoun and the epithet R-expression solves the problem:

(79) The president$_i$ said that he$_i$ would be addressing the nation tonight.

Summary

In this chapter, we've argued that in order to properly account for what you know about your language we need to assume that there are three different kinds of DP: anaphors, pronouns and R-expressions. The distribution of these three different kinds of DP is governed by Binding Theory, one of the most written about topics in Chomskyian linguistics.

Anaphors include reflexives, such as *herself* and *ourselves*, and reciprocals, such as *each other*. These DPs cannot independently refer, instead requiring that there be an antecedent for them within the appropriate domain. This requirement is expressed in Principle A of the Binding Theory:

(46) *Principle A of the Binding Theory*
An anaphor must be bound in its governing category.

(47) α *binds* β if α c-commands and is coindexed with β.

(48) The *governing category* for an element α is the minimal XP containing
(a) the element α,
(b) a governor for α, and
(c) a subject which is accessible to α.

(49) α is a subject which is accessible to β if and only if (hypothetical) coindexation between the anaphor and the subject violates no grammatical principle.

Pronouns, such as *she* and *us*, seem to be in complementary distribution with anaphors. Wherever you can have an anaphor, you can't have a pronoun and vice versa. This complementarity with anaphors is built into Principle B of the Binding Theory:

(59) *Principle B of the Binding Theory*
A pronoun must be free in its governing category (where 'free' = 'not bound').

R-expressions include proper names, definite descriptions and anaphoric epithets. Unlike anaphors and pronouns, they do not have a governing category of any kind. R-expressions may not co-refer with a DP that c-commands them, no matter how far away it is. This requirement is expressed in Principle C:

(72) *Principle C of the Binding Theory*
An R-expression must be free (again, 'free' = 'not bound').

Open issue: non-complementarity of pronouns and anaphors

As we mentioned above in discussing Principles A and B of the Binding Theory, pronouns and anaphors look like they're in complementary distribution, and this complementarity is built directly into Binding Theory. Anaphors must be bound in their governing category, and pronouns must be free in their governing category. The situation couldn't be more black and white than that. The problem is that, even in English, anaphors and pronouns are... not... quite... fully... exactly... 100 percent... in complementary distribution. It's a painful thing to admit. Trust me.

Unfortunately, the relevant cases are not even very complicated either. Consider the contrast (or lack thereof) in (80):

(80) a. They$_i$ think that pictures of each other$_i$ are on sale.
b. They$_i$ think that pictures of them$_i$ are on sale.

It looks like here we've got exactly what we didn't want to see. *Each other* is an anaphor, subject to Principle A, and *them* is a pronoun, subject to Principle B. Whatever the structure of the sentences in (80) is, we don't predict that you can get a pronoun and an anaphor in the same position. One of them should be grammatical and the other ungrammatical. But that's not what we see.

EXERCISE 2

Which of (80a) or (80b) is predicted to be ungrammatical by Binding Theory as we've currently formulated it? Why?

There have been a number of attempts to come to grips with the problem represented by (80) and related problems in other languages. Huang (1983) takes essentially a syntactic approach. He argues that we do need to relax the complementarity that Binding Theory enforces, but we don't need to modify Principles A and B *per se*. Rather, the solution is to assume that governing categories are defined differently for pronouns and anaphors.

Huang (1983) starts by noticing that, in addition to the empirical problem posed by (80), we also have a potential conceptual problem when it comes to the notion of 'governing category' and, in particular, the 'accessible subject' requirement that we saw in Section 5.1. He observes that 'being a subject which is accessible to a' just seems to be another way of saying 'is capable of serving as an antecedent for a'. It makes sense to include a requirement like this for the governing category of anaphors. As Huang puts it (1983: 556): 'since anaphors need antecedents, what the Binding Theory says is that they must be bound in the minimal domain whose [subject] is a possible antecedent (i.e. accessible).' However, it's precisely what makes the accessible subject requirement look right for anaphors that makes it look weird for pronouns. Pronouns can be free, as we've seen above. They don't ever *require* an antecedent. So in determining the

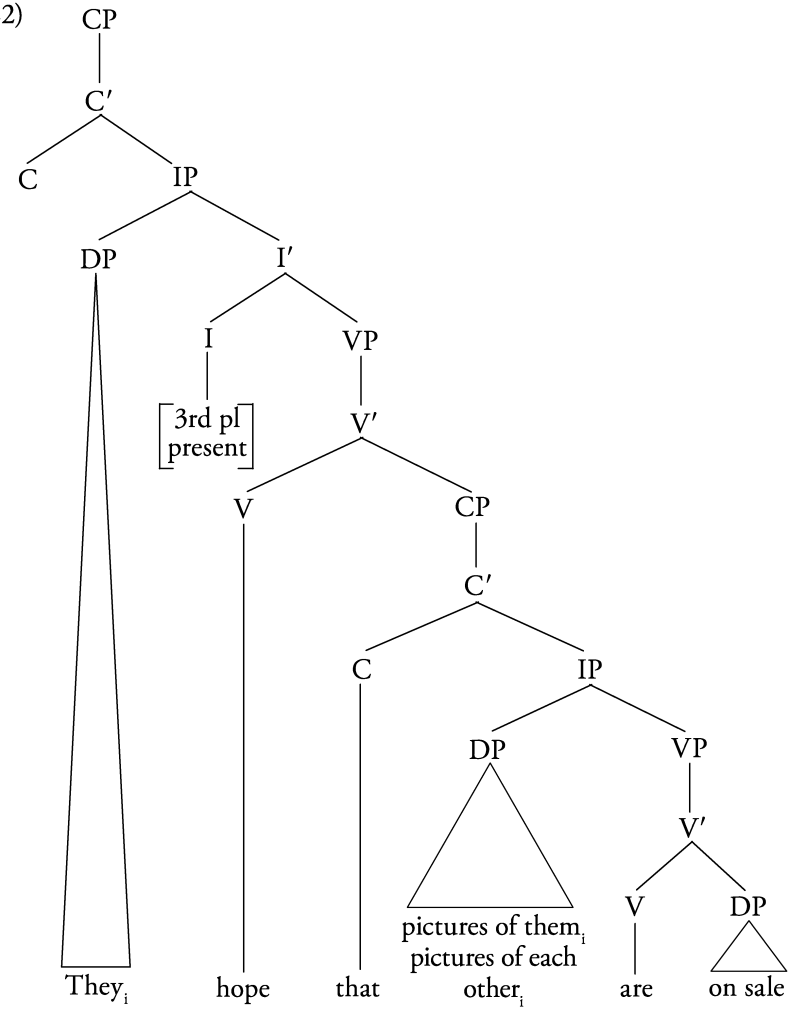
governing category for a pronoun, why should finding a possible antecedent matter? Why should it matter whether the subject is accessible or not?

Huang claims that if we eliminate this conceptual weirdness, we also explain why both sentences in (80) are grammatical. Specifically, Huang suggests changing the definition of governing category from (48) to (81) (I've put the change in italics):

(81) The governing category for an element α is the minimal XP containing
(a) the element α,
(b) a governor for a, and
(c) a subject which, *if* α is *an anaphor*, is accessible to α.

To see how (81) solves the problem, let's take a closer look at the sentences in (80).

(82)



Let's start with the anaphor case in (80a). In this case, the situation is exactly the same as *John thinks that a picture of himself is on sale*, which we discussed in Section 5.1. The governor for the anaphor is the P *of*. However, neither PP nor NP nor DP has a subject. Therefore the minimal XP which contains the anaphor, the anaphor's governor and a subject is the IP *pictures of each other are on sale*. However, since *each other* is an anaphor, we need to check whether this subject is an accessible one, and it turns out that it's not. If we were to hypothetically coindex the subject *pictures of each other* with the anaphor *each other*, we would create an i-within-i configuration. The next nearest category which has a subject is the matrix IP *They think that pictures of each other are on sale*. This subject, *they*, is accessible to the anaphor *each other*. Therefore, the governing category is the matrix IP, an antecedent for *each other* can be found within that domain, and the sentence is grammatical.

We proceed in exactly the same way with (80b)/(82) to determine the governing category for the pronoun *them*. The first likely governing category is the embedded IP *pictures of them are on sale*. This IP has a subject, but, since *them* is a pronoun, we don't need to check to see if the subject is accessible. Therefore, the governing category for the pronoun *them* is the embedded IP. Now, as Principle B requires, the pronoun *them* is free in its governing category, and (80b) is now correctly predicted to be grammatical.

FURTHER READING

The general topic of anaphora (the conditions under which two DPs can co-refer) has long been a topic of investigation within Chomskyian generative grammar. The original discussion of Binding Theory using the terms that we've used in this chapter can be found in Chomsky (1981), but work in this area dates back many years prior to that. Chomsky (1981) can be *extremely* difficult to read, however, so I don't suggest that people start there. There are discussions of Binding Theory more or less as we have it here in most standard Chomskyian textbooks. I think the best discussions are in Lasnik and Uriagereka (1988) and Haegeman (1994). Webelhuth (1995a) also has a chapter on Binding Theory, but that discussion won't really be accessible until Chapter 9 has been completed. Also worth checking out is a collection of articles on anaphora by Howard Lasnik (1989), some written as early as 1973. It also contains an excellent historical overview of Binding Theory.

ANSWERS TO IN-TEXT EXERCISES

Exercise 1

I saw [John and Mary]$_i$ yesterday. *I opened the door for [each other]$_i$.

Exercise 2

(80) a. They$_i$ think that pictures of each other$_i$ are on sale.
b. They$_i$ think that pictures of them$_i$ are on sale.

(80b), with a pronoun, is the one which is predicted to be ungrammatical. The governing category for the pronoun will be the entire IP *They think that pictures of them are on sale*. The minimal category containing the pronoun, the pronoun's governor, *of*, and a subject is the embedded subject *pictures of them*, but this is not an accessible subject for the pronoun. Hypothetical coindexation between the pronoun *them* and the subject *pictures of them* creates an i-within-i violation. The next higher category which contains a subject is the main clause. Coindexation between the main clause subject *they and* the pronoun *them* creates no i-within-i violation. Therefore, the minimal category containing the pronoun, a governor for the pronoun and a subject which is accessible to the pronoun is the entire IP *they think that pictures of them are on sale*. Since the pronoun is bound in its governing category, Principle B is violated.

FURTHER EXERCISES

Exercise 1

For the following sentences:

(a) Draw the correct trees.
(b) Identify all anaphors, pronouns and R-expressions.
(c) Show how Principles A, B and C of the Binding Theory make the indicated predictions.

1. John bought their$_i$ books about each other$_i$.
2. Sue$_i$ likes her$_j$.
3. *Bill Clinton$_i$ expected the discussion about the president$_i$ to be brief.
4. *Bob$_i$ hopes that Mary's$_j$ pictures of himself$_i$ will sell well.
5. Mary$_i$ wondered whether she$_i$ would like the pictures of herself$_i$.
6. *John$_i$ thought that Bill$_j$ said that Mary$_k$ liked John$_i$.
7. The problems with John's$_i$ car surprised his$_i$ friends.
8. For Mary$_i$ to support John$_j$ would help him$_j$.
9. *For Mary$_i$ to support John$_j$ would help herself$_i$.
10. [The employee of John's$_i$ firm]$_j$ deposited his$_{i/j}$ cheque.

Exercise 2

How does the sentence in (1) provide an argument for the existence of PRO, which we introduced in the last chapter?

(1) To behave oneself in public is every citizen's duty.

Exercise 3

Jackendoff et al. (1993) make some interesting observations about the interpretation of the word 'home', as illustrated in (1):

(1) John went home.

What are the interpretive possibilities for (2) and (3)? Put another way, whose 'home' is the destination?

(2) John said that Mary went home.
(3) John's friend went home.

What do these facts suggest to you about how to understand the interpretation of 'home' in the above examples? Be sure to discuss (2) and (3) in detail and draw the trees for them.

Now consider (4):

(4) John and Mary went home.

Is the interpretation of 'home' in this sentence problematic for your proposal? Give your reasons. (If your judgment regarding the interpretation of 'home' in (4) is unclear, discuss the *possible* interpretations and which ones (if any) *would* be a problem for your proposal.)

Movement and Chains

6

In the previous chapters, your attention has been deliberately directed towards certain kinds of examples. We argued in Chapter 4, for example, that a verb like *devour* must take a DP direct object. It can only be inserted in a phrase-structure tree under a V node which has a DP sister. By imposing this requirement, we explained contrasts like the one seen in (1):

(1) a. John devoured his dinner.
b. *John devoured.

However, this requirement seems manifestly violated in cases like (2):

(2) What did John devour?

In this case, it seems as though *devour* does not have the requisite DP sister, yet the sentence is grammatical. And if we try to add one, the sentence becomes ungrammatical:

(3) *What did John devour the pizza?

We have a similar kind of mystery with respect to Binding Theory. In the previous chapter, we concluded, among other things, that an anaphor must be bound in its governing category by its antecedent, and in order to have binding we have to have c-command. This was intended to explain cases like (4) and (5):

(4) *Himself$_i$ likes John$_i$.

(5) *John$_i$'s likes himself$_i$.

However, in the course of our discussion we neglected to mention sentences like (6):

(6) Which picture of himself$_i$ does John$_i$ like?

Assuming that *John* is in Spec IP as usual, it surely doesn't c-command the anaphor *himself*, wherever it is that the anaphor is actually sitting. Contrary to expectations though, (6) is grammatical.

The problems posed by the cases mentioned above aren't trivial. We seem to have evidence both that *devour* requires a DP sister and that it doesn't require one, and both that there *is* a c-command requirement on anaphors and that there *isn't*. In this chapter, we will address sentences like (2) and (6), suggesting that, despite outward appearances, they, in fact, support the theory that we've been developing so far, with one important additional assumption.

6.1 Transformations: an introduction

Consider for a moment the contrast between (1a) and (2):

(1)a. John devoured his dinner.

(2) What did John devour?

(2) presents a serious problem for the notion 'direct object' as we've formulated it so far. In this book, we're taking a configurational approach to terms from traditional grammar like 'subject' and 'direct object'. In other words, in the particular case at hand, we've said that something qualifies as a 'direct object' in virtue of being a sister to the verb. In (2), however, we have something different. *What* bears the same relationship to *devour* in (2) that *his dinner* does in (la). They're both assigned the Patient θ-role. *What* is interpreted as a direct object, but it's not a sister to the verb.

One thing we could do is modify our configurational approach to direct objects. We could say that direct objects can either be generated as sisters to the verb (what we've seen until now) or in a sentence-initial position, which we'll assume for the moment is the specifier of CP.

However, it's not clear that this solution is going to buy us much in the long term. To begin with, it's not simply direct objects which can appear in Spec CP. Essentially, anything can – arguments, adjuncts or even whole PPs. We would have to totally abandon configurational definitions of anything.

(7) a. *What* did John buy at the store on Fridays?
b. *Why* did John put the book on the table?
c. *In which park* did Charles Mingus play for free last year?

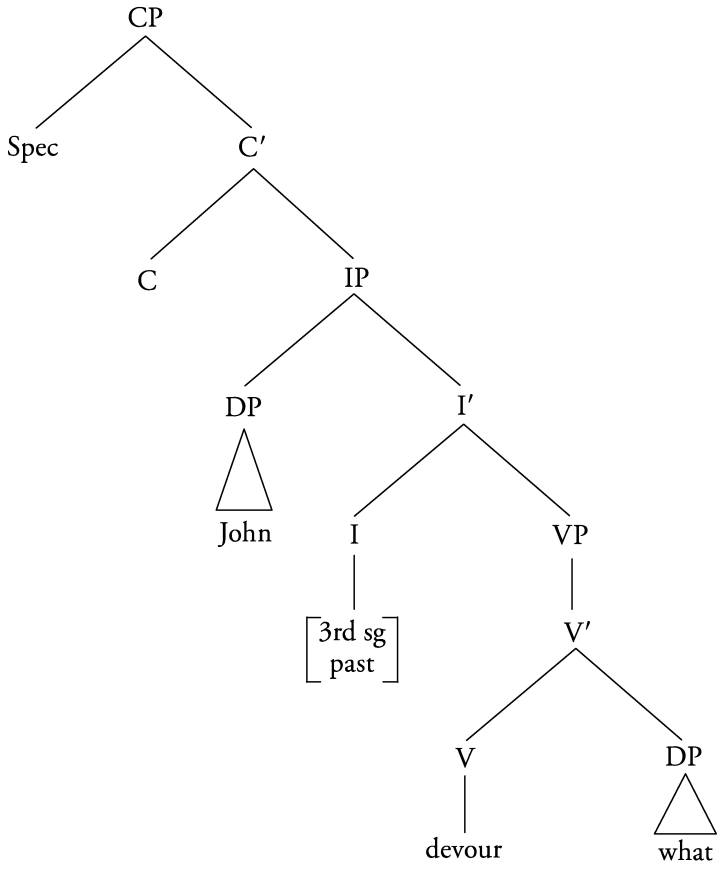
Furthermore, allowing subjects, direct objects, adjuncts and so on to be base-generated in Spec CP would undermine much of what we achieved in the previous chapters regarding the assignment of Case, for example. Spec CP is certainly not a position which is governed by V. We would therefore be forced to give up the requirement that Case can only be assigned under government.

There is an alternative, however. Let's assume instead that the order that the words are actually pronounced in isn't the whole story. Suppose that our phrase-structure rules generate an underlying structure, which is then transformed by

a rule into what we actually pronounce and perceive. Put another way, we have two kinds of operations: structure-building operations, as expressed by our phrase-structure rules, and structure-changing operations, which we'll call 'transformations'.

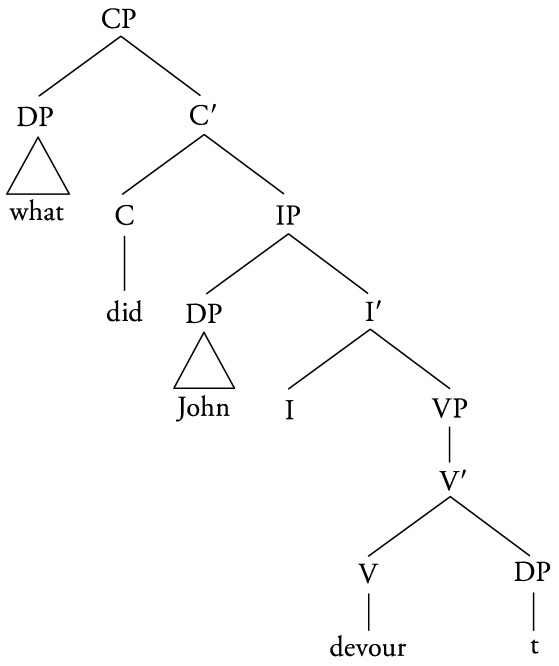
Intuitively, this approach has a certain appeal. For one thing, the idea that there are transformations allows us to account for the fact that, in (2), *what*, although in sentence-initial position, is interpreted as the direct object of *devour*. In an underlying representation, one which doesn't show up on the surface, it *was* the direct object. Using the terminology that we'll adopt, (2) has as its *Deep Structure* the representation in (8):

(8)



It is at this level of Deep Structure (usually abbreviated 'D-Structure') where θ-role assignment takes place. *What* is assigned the Patient θ-role under government from *devour*. The transformational rule of *Wh-movement* takes *what* and moves it to Spec CP, leaving an unpronounced place-holder element, called a 'trace' (abbreviated 't'), to mark the original position. We thus have the Surface Structure (or 'S-Structure') representation in (9):[1]

(9)



It should be clear from (9) why we want to say that Spec CP is the landing site for wh-words. The combination of word order and phrase structure almost forces that conclusion on us. Keeping our configurational definition of *subject*, (9) tells us that the subject must be in Spec IP. There are two words which must precede the subject, *what* and *did*. No other word orders are possible:

(10) a. *John what did buy?
b. *What John did buy?
c. *Did John what buy? etc.

The wh-word itself is a maximal projection (a DP, in fact, as it can surface in canonical DP positions like subject and direct object in echo questions (e.g. *John bought what?*)) and *do* is surely a head (presumably an I°). As required, the CP projection provides exactly one host of each type (an XP position, Spec CP and an X° position, C) in the right order (the XP position precedes the X° position). Therefore we seem justified in concluding that the landing site of Wh-movement is Spec CP. To introduce some terminology which will be useful in the chapters to come, positions where arguments of a predicate are generated at D-Structure (like the specifier of IP (subjects) or sister to the verb (direct objects)) are called argument positions or *A-positions*. Positions like the specifier of CP, where arguments of a predicate are not generated at D-Structure, are called *A′-positions* (read as 'A-bar', like X′).

This idea, that there's a D-Structure representation which is modified by transformations to give us what we see on the surface (S-Structure), would seem to

explain the contrast between (1) and (2). The reason that *devour* does not take a DP sister in (2) is that it already has a DP sister which has been moved to Spec CP. However, being good scientists, we need to make sure that transformations are not just an ad hoc device for getting us out of this one problem. To that end, let's take a look at a couple of other problems that transformations seem to solve for us.

Further evidence for transformations

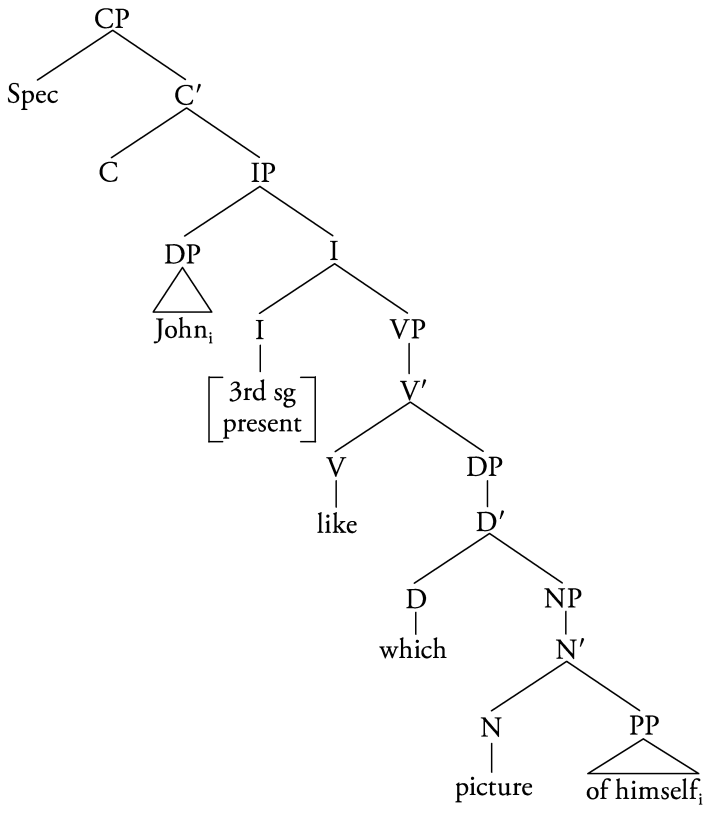
First, transformations do seem to get us out of that problem with respect to Binding Theory. We brought up sentences like (6) (repeated below), which we omitted in our discussion of Binding Theory in Chapter 5.

(6) Which picture of himself$_i$ does John$_i$ like t?

If a sentence like (6) is base-generated (i.e. created by our structure-building phrase-structure rules rather than by a structure-changing transformation), then we need to throw out most of what we claimed about Principle A. It seems unlikely that *John* is somehow able to c-command *himself* in (6). The first branching node dominating the subject DP is IP, and Spec CP is not dominated by IP.

However, the assumption that (6) is derived by the same transformational rule of Wh-movement, and that Principle A applies to the D-Structure representation, not the S-Structure one, would seem to solve the problem. The S-Structure (6) would have (11) as its D-Structure representation:

(11)



Here, the anaphor *himself* is bound in its governing category, just as Principle A requires.

Transformations also seem to be a good candidate for explaining certain subject-verb agreement facts. As we saw in Chapter 3, the I° node under Spec-Head agreement picks up the person and number features of the subject which are ultimately responsible for subject-verb agreement. With that in mind, consider the contrast in (12) and (13):

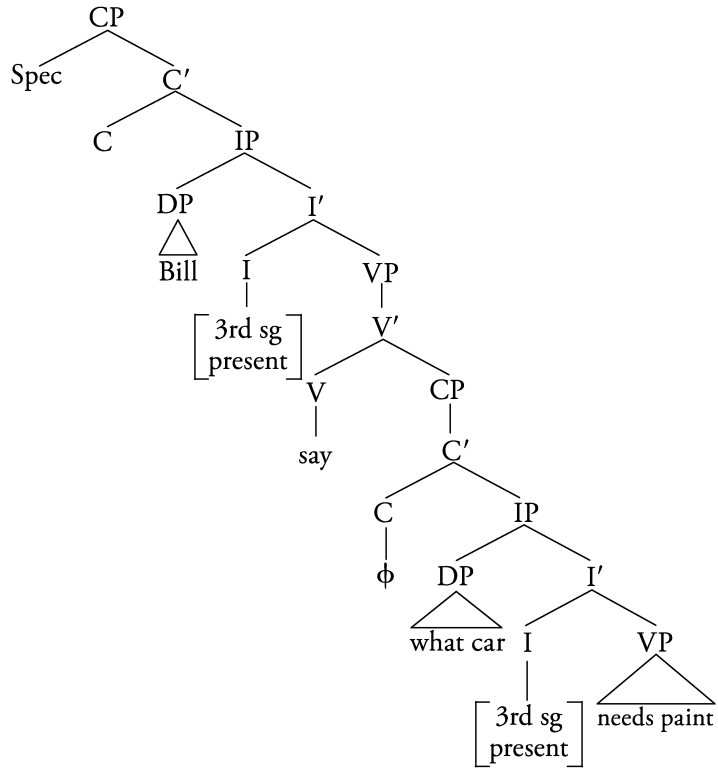
(12) a. Which car did Bill say needs paint?
b. *Which car did Bill say need paint?

(13) a. Which cars did Bill say need paint?
b. *Which cars did Bill say needs paint?

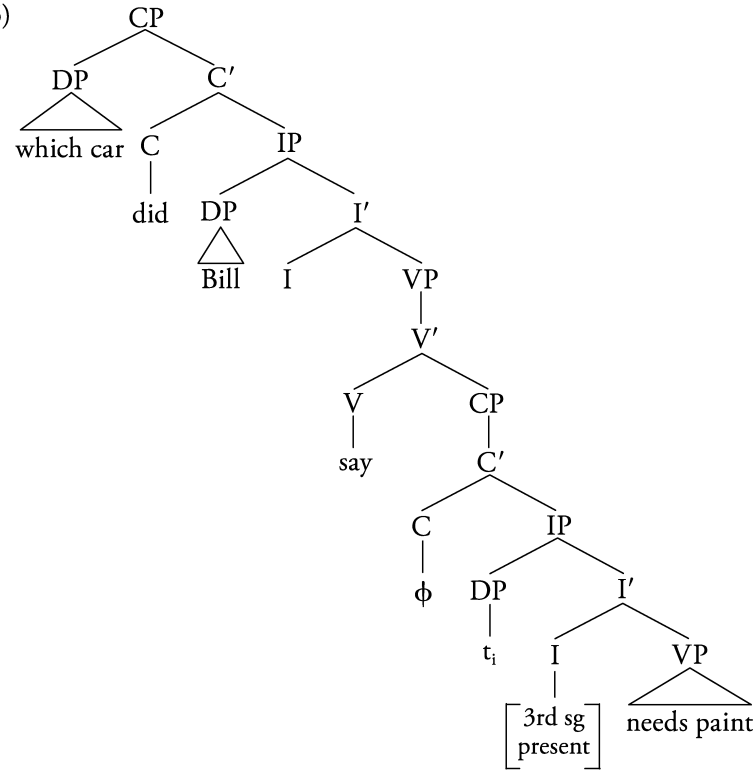
Assuming that the (a) sentences are directly generated by phrase structure rules (i.e. base-generated), the choice of *needs* versus *need* is hard to explain. Spec-Head agreement takes place between Spec IP and I, but in the (a) sentences there is no subject in the embedded Spec IP. One might expect that, since there's no subject to agree with, either (a) or (b) might be possible, or maybe that both would be ungrammatical. However, speakers of English have the unshakable intuition that the (a) sentences are grammatical and that the (b) sentences are not.

Transformations, on the other hand, allow us to say that, while there is no subject at *S-Structure*, there is one at *D-Structure*:

(14)



(15)



Subject-verb agreement takes place at D-Structure (as represented in (14)), and then Wh-movement moves the wh-phrase to Spec CP, creating the S-Structure representation in (15).

6.2 Wh-movement: subjacency and successive cyclicity

So far, we've seen examples of wh-movement which involved either one clause (as in (16)) or two clauses (as in (17)):

(16) $[_{CP}$ What $_i$ did $[_{IP}$ John devour $t_i]]$?

(17) $[_{CP}$ Which car$_i$ did $[_{IP}$ Bill say $[_{CP}$ $[_{IP}$ t_i needs paint$]]]]$?

In this section, we're going to take a look at wh-movement more closely, examining in more detail how it works and what its limits are.

On the basis of (16) and (17) a natural question is whether wh-movement can move a wh-word through any number of clauses, or whether there is some constraint on how far the wh-word can move.

BEFORE CONTINUING

Provide examples of wh-movement out of three clauses, four clauses and five clauses respectively.

An examination of cases similar to (16) and (17) might suggest that wh-movement can move a DP out of quite a few clauses:

(18) $[_{CP}$ Who$_i$ does $[_{IP}$ Alice think $[_{CP}$ that $[_{IP}$ Mary said $[_{CP}$ that $[_{IP}$ Bill liked $t_i]]]]]]$?

(19) $[_{CP}$ Who$_i$ did $[_{IP}$ Eunice say $[_{CP}$ that $[_{IP}$ John believed $[_{CP}$ that $[_{IP}$ Bill said $[_{CP}$ that $[_{IP}$ Peter liked $t_i]]]]]]]]$?

Examples like this can be continued ad infinitum.

There is, of course, a practical limit to the number of clauses out of which wh-movement can take place, as a result of short-term memory limitations. If the sentence has 500 clauses, it's likely that you'll forget half of it before you get to the end. If you tried to move the wh-word to the beginning of a sentence with 500 trillion trillion clauses, you'd die before you finished saying it. However, we will assume that the fact that humans have a finite lifespan, or a fixed short-term memory, is a *performance-related* effect. As far as your linguistic *competence*, the underlying system of knowledge that you possess, is concerned, wh-movement in a sentence with 500 trillion trillion clauses is in principle possible.

However, it turns out that this can't quite be the whole story. Consider the sentences in (20):

(20) a. John wondered where Bill saw *Midnight Cowboy*.
b. John wondered why Bill saw *Midnight Cowboy*.
c. John wondered who saw *Midnight Cowboy*.
d. John wondered which film Bill saw.

The embedded clauses in (20) seem to have a lot in common with wh-questions:

(21) a. Where did Bill see *Midnight Cowboy*?
b. Why did Bill see *Midnight Cowboy*?
c. Who saw *Midnight Cowboy*?
d. Which film did Bill see?

Although the sentences in (20) aren't questions, we still want to say that wh-movement is taking place in the embedded clause, for all the same reasons that we would give for the sentences in (21).

EXERCISE 1

Using Binding Theory and the subject-verb agreement facts discussed in Section 6.1 above, provide arguments that wh-movement takes place in sentences like (20).

Let's take a closer look now at (20b). If we were to question the object of *see*, asking what the thing was such that John was wondering about why Bill saw

it, we get (22):

(22) *What$_i$ did John wonder why$_j$ Bill saw t_i t_j?

(22) is quite strongly ungrammatical, which comes perhaps as something of a surprise given what we saw with (18) and (19) above. (22) has fewer words than either one of those sentences, and seems as though it should be more or less on a par with (23):

(23) What$_i$ did John say that he saw t_i?

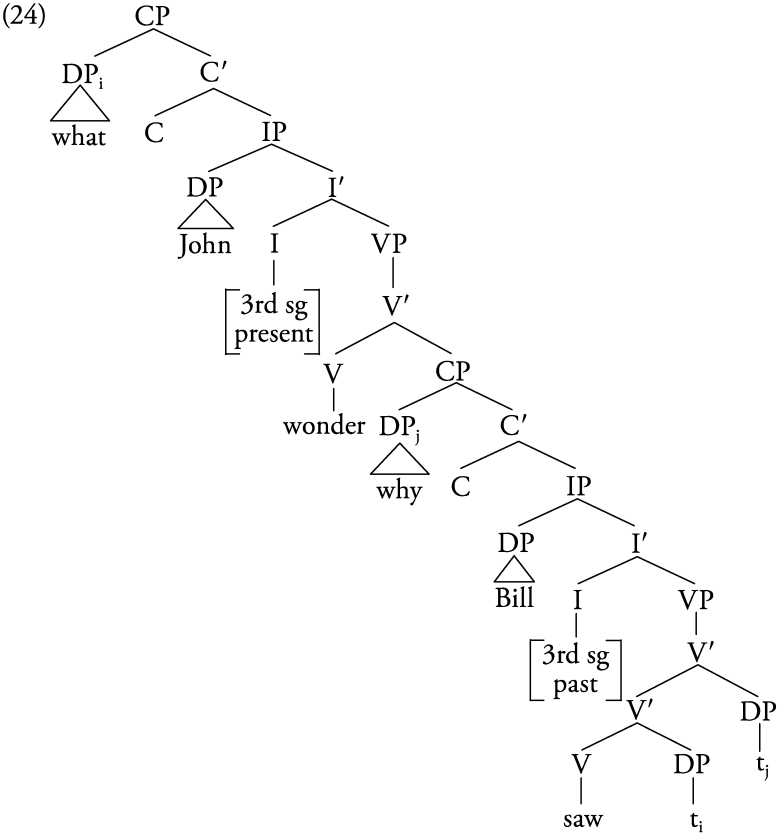
It doesn't on the surface seem as though (23) is any more 'complicated' than (22), so there must be some other principle at work.

BEFORE CONTINUING

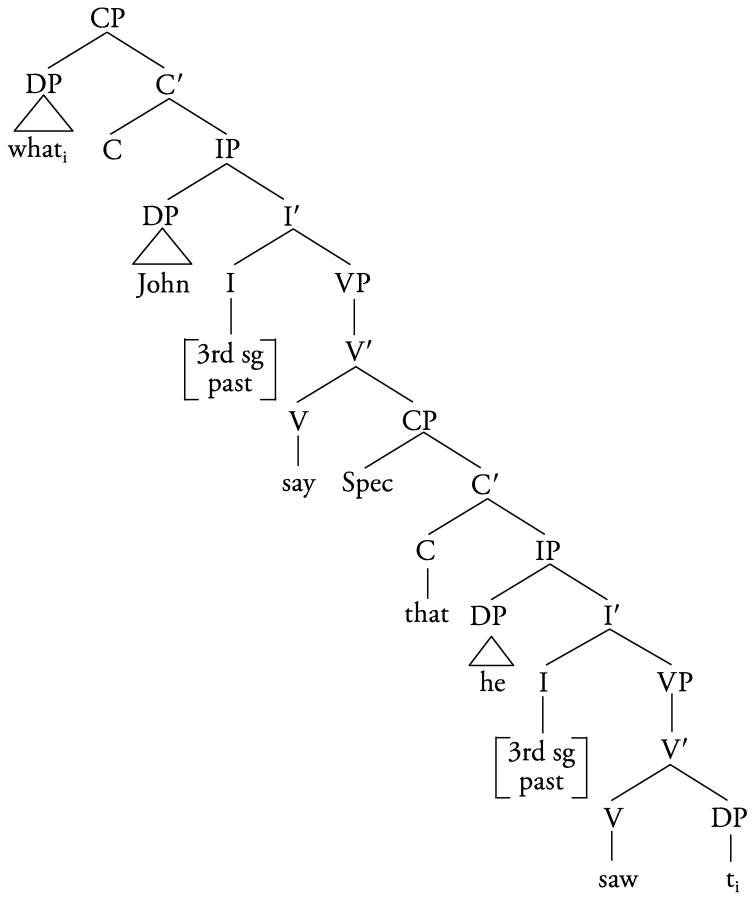
Draw the tree for (23) and then (22).

As usual, the first step is to see if there is any structural difference between the two sentences which might provide an explanation. Comparing the trees for (22) and (23), we have (24) and (25):

(24)



(25)



Although it is true that (22) and (23) have roughly the same number of words, notice that there is a crucial difference with respect to their structures. In the ungrammatical (22), there are *two* instances of wh-movement, with the result that *why* is occupying the Spec CP position of the embedded clause. On the other hand, the grammatical (23) simply has complementizer *that* in the head position C and only one instance of wh-movement.

This difference suggests a way to explain why (23) is grammatical but (22) is not: wh-words usually surface in Spec CP positions, whether it's the matrix (topmost) Spec CP as in (23) or an embedded one as in (20) above. Perhaps it's the case that wh-words don't just move in one fell swoop to the matrix Spec CP. Instead, perhaps there is some principle which forces them to move in a series of shorter steps *via intermediate Spec CP positions*. When one of these positions is already filled and the wh-word tries to 'skip' it, then this principle is violated, and the result is ungrammaticality. Let's refer to this principle as the *Successive Cyclic Hypothesis*: Wh-movement proceeds in a series of cycles, first moving to the Spec CP of the clause in which it is base-generated, and then from Spec CP to Spec CP until it reaches the matrix Spec CP. We could formulate such a

principle informally as (26):

(26) *Successive Cyclic Hypothesis*
'Long-distance' wh-movement must proceed via all intermediate Spec CP positions.

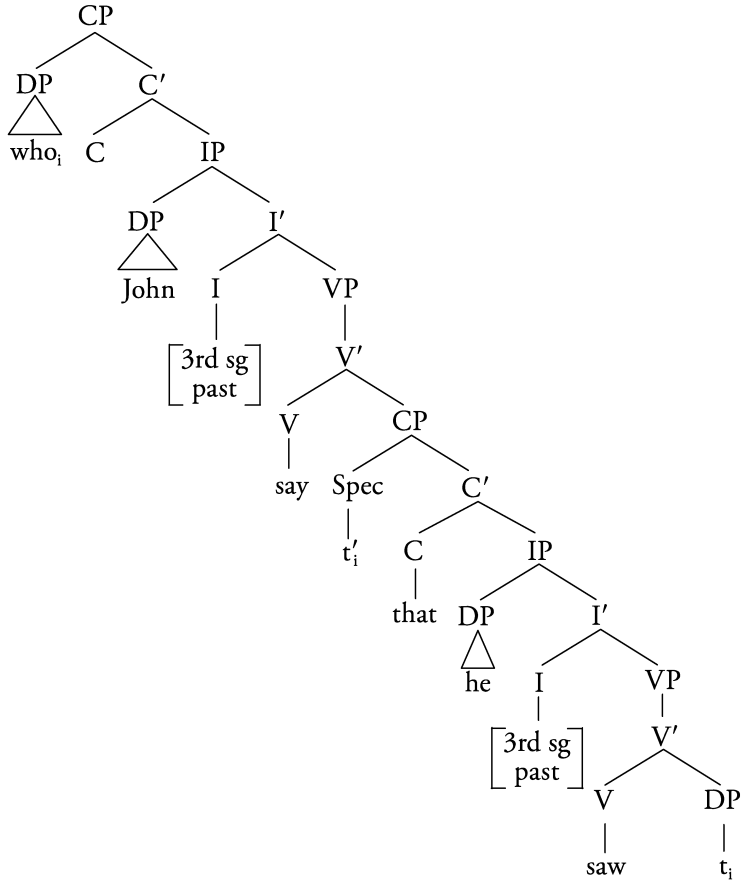
(26) will correctly account for the contrast between (22) and (23), and also correctly predict the grammaticality of the other long-distance' cases (18) and (19) which we saw above:

(18) Who does Alice think that Mary said that Bill liked t?

(19) Who did Eunice tell Mary that John believed that Bill liked t?

In both cases, the intermediate Spec CP positions are all open, and therefore *who* can use them to move successive-cyclically to the matrix Spec CP. When this happens, *who* not only leaves a trace in its D-structure position, but also an 'intermediate' trace in the Spec CP that it has passed through. So, technically, the tree in (25) needs to be changed slightly to (27), so that it has an intermediate trace in the embedded Spec CP.

(27)



EXERCISE 2

How many applications of wh-movement do we need for (18) and (19) if we accept (26)? Draw the S-Structure tree for (18).

So it seems that we're beginning to get a handle on the question 'How far is too far?' with respect to wh-movement. Under the Successive Cyclic Hypothesis, it looks as though the farthest that an individual application of wh-movement can go is from one Spec CP to the next highest Spec CP. This means that it's not the wh-movement operation itself which operates 'long distance'. Rather, the apparent distance is covered bit by bit as a series of shorter operations which are linked together.

Having formulated the Successive Cyclic Hypothesis, the next question is whether we've hit an irreducible axiom of the system, or whether the Successive Cyclic Hypothesis itself follows from some deeper principle. The traditional intuition in the literature is that there is indeed a deeper principle which underlies the Successive Cyclic Hypothesis. Specifically, the assumption is that certain nodes in the tree constitute *bounding nodes* for wh-movement. Since wh-movement can only proceed successive-cyclically from clause to clause, it must be the case that some node in the clause constitutes a bounding node for wh-movement. Under this assumption, we could then formulate a principle, which is referred to in the literature as *Subjacency*, along the following lines:

(28) *The Subjacency Condition*
A single instance of wh-movement (whether from D-Structure position to Spec CP or from one Spec CP to another) can cross only *one* bounding node.

In order to be clear about what we mean by 'crossing' a node, let's give an explicit definition. This is important here because 'crossing' sounds like a linear (left-to-right) concept, as if you 'cross' a node by moving to the left of it, but this isn't what we mean. Because phrase-structure trees encode hierarchical information, what we need is a hierarchical definition, like (29):

(29) An instance of movement *crosses* a node X if X dominates the departure site but does not dominate the landing site.

Now we're in a position to see how a Subjacency Principle could force movement to be successive cyclic: it must be the case that when an intermediate Spec CP is filled, *two* bounding nodes are crossed in order to get to the nearest *empty* Spec CP, and the result is ungrammaticality. Therefore, a single instance of wh-movement can only move you to the Spec CP of the clause that you're in, or from one Spec CP to the immediately next higher Spec CP.

If the Subjacency Principle is on the right track, the next task is to figure out which node or nodes might be the bounding nodes. Which node or nodes look like good possibilities? There aren't too many. Even in very simple clauses, the effect can still be seen. For example, consider (30), with some structure indicated:

(30) *[$_{CP}$ What$_i$ did [$_{IP}$ John [$_{VP}$ wonder [$_{CP}$ why$_j$ [$_{IP}$ Bill [$_{VP}$ hit t_i t_j]]]]]]?

We're assuming now that *what* crosses at least two bounding nodes in moving from direct object of *hit* to the matrix Spec CP, which is why the sentence is ungrammatical. (The intermediate Spec CP position is filled up by *why*.) Under this assumption, looking at the structure, the only real possibilities for bounding nodes are VP, IP and CP. We can eliminate CP immediately. In moving from direct object position of *hit* to the matrix Spec CP, we cross only one CP, not two. (The matrix CP won't be crossed when you move to the matrix Spec CP because it dominates *both* the departure site *and* the landing site.) VP is a possibility, since there are at least two of them in (30). However, if we consider a sentence like (31), with a 'VP-shell' structure of the sort we discussed in Chapter 3, wh-movement still appears possible:

(31) Who$_i$ should [$_{IP}$ John [$_{VP}$ have [$_{VP}$ been [$_{VP}$ speaking to t_i]]]]?

Therefore, we seem to be left with IP as the only plausible candidate. Two IP nodes are crossed in (30), the embedded IP *Bill hit* t_i t_j and the IP of the main clause. (They both dominate the departure site and both fail to dominate the landing site.) However, in (31), there's only one IP, so only one IP is crossed. If you think about in general moving successive-cyclically from one Spec CP to the next, only one IP will be crossed.

Let's therefore introduce the definition of IP as a bounding node into the definition of Subjacency:

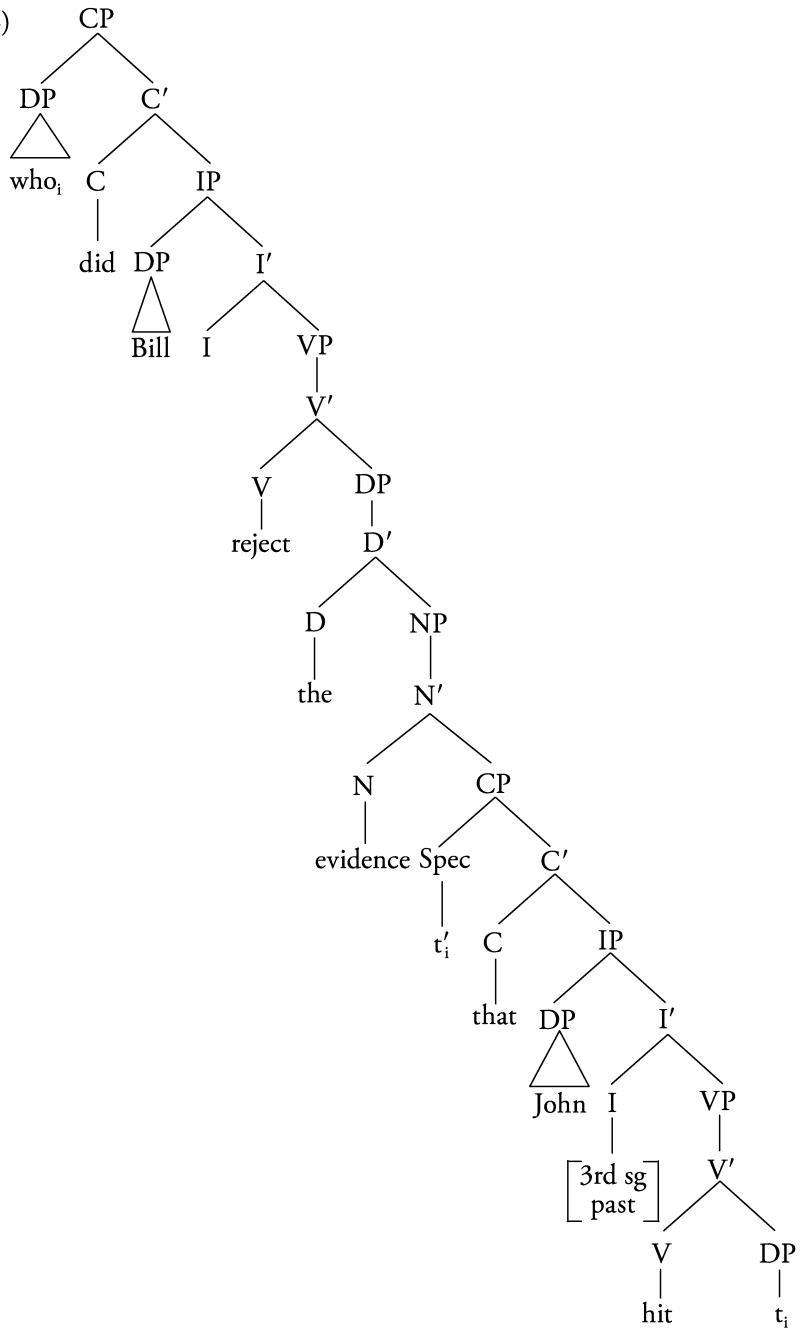
(32) *The Subjacency Condition* (amended)
A single instance of wh-movement can cross only one bounding node, where IP is a bounding node.

Now you might be wondering why we've gone to all this trouble with respect to the Subjacency Condition when we already have the Successive Cyclic Hypothesis. Given the cases we've seen so far, the Subjacency Condition just seems to be a different way of saying the same thing. However, we're about to see a case where the Successive Cyclic Hypothesis makes the wrong prediction, but the Subjacency Condition, with a slight addition, makes the right one.

Consider (33), with the tree in (34):

(33) *Who$_i$ did Bill reject the evidence that John hit t_i?

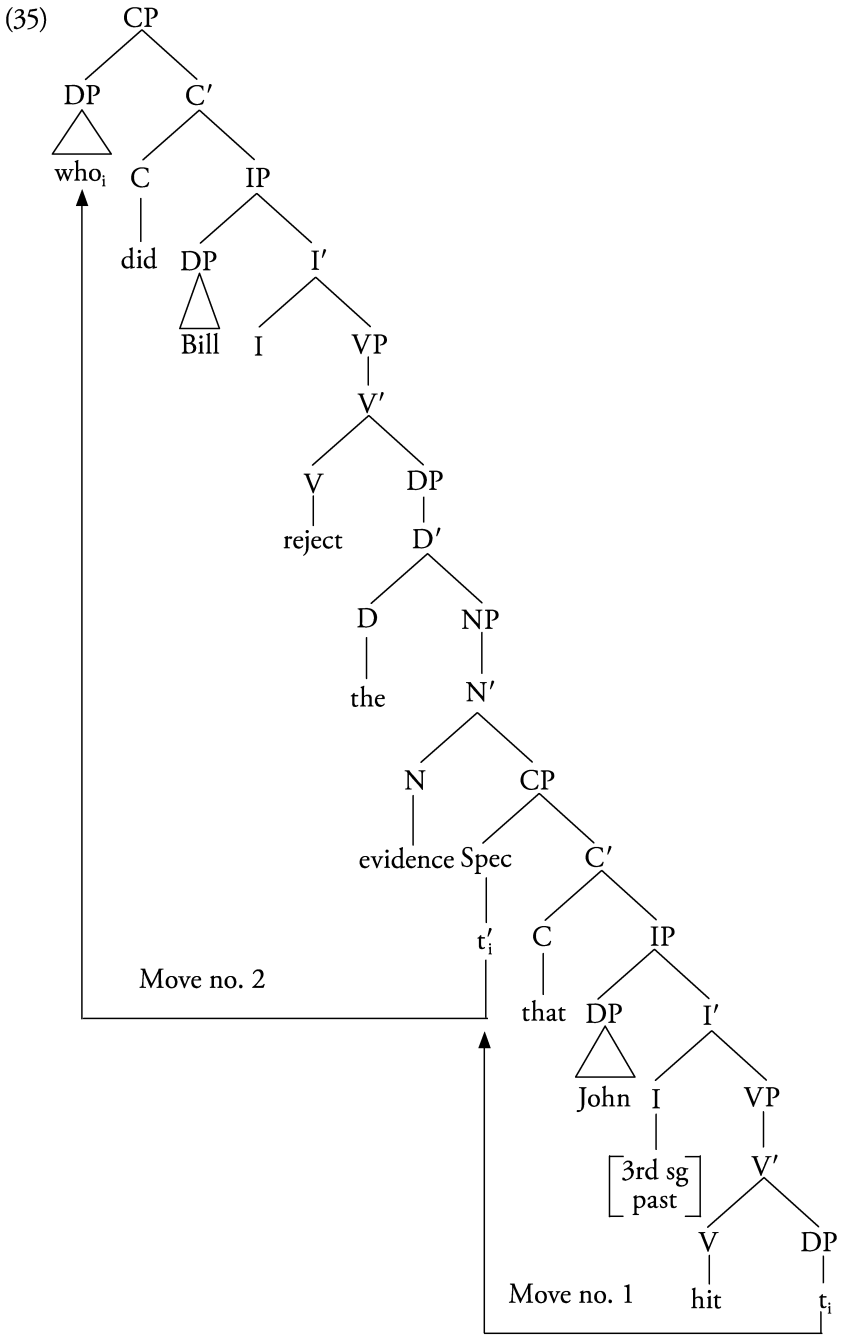
(34)



The Successive Cyclic Hypothesis, and the Subjacency Principle as we've formulated it so far, predict that (33) should be grammatical. The CP complement of the noun *evidence* has its specifier position open, and so the derivation

should proceed in the following way: *who* first moves from its D-Structure position to the specifier of the CP complement, crossing one bounding node (the embedded IP), and then can move from there to the matrix Spec CP, crossing only one bounding node again, the matrix IP:

(35)



What the ungrammaticality of (33) would seem to indicate is that there's an extra bounding node that gets crossed somewhere along the line as part of one of those two moves. To try to figure out where that mystery bounding node might be, let's take a closer look at the derivation of (33). The two moves of *who* are illustrated by the numbered arrows in (35).

It doesn't seem as though move no. 1 is the one that's causing the problem. That move is just like the ordinary, garden-variety wh-movement seen in (36):

(36) Who did John hit t?

It doesn't seem as though there could be a mystery bounding node that's somewhere in (35) but not in (36). Therefore, the violation must be triggered during move no. 2, when the wh-word moves from the embedded Spec CP to the matrix Spec CP.

Looking at move no. 2, there is an obvious possible candidate for our mystery bounding node: the DP. We know that movement from a lower clause to a higher clause is in general possible. It must be the fact that what we have in (35) is a clause that's within a DP. The CP in (35) is a complement of an N, rather than a V.

Under the assumption that DP is a bounding node in addition to IP, move no. 2 in (35) from the embedded Spec CP to the matrix Spec CP will now cross two bounding nodes, DP and the matrix IP, and the ungrammaticality of (33) will be explained. So let's amend our definition of Subjacency to include DP among the class of bounding nodes:

(37) *The Subjacency Condition* (final version)
A single instance of wh-movement can cross only one bounding node, where IP and DP are bounding nodes.

Hopefully, it's now clear why we needed to replace the Successive Cyclic Hypothesis. There's no obvious modification that we could make to the Successive Cyclic Hypothesis to ensure that move no. 2 in (35) caused a violation. It's important to understand though that by replacing the Successive Cyclic Hypothesis by the Subjacency Condition, we're not denying that wh-movement takes place from Spec CP to Spec CP. It does, and we'll even continue to use the term 'successive cyclicity' to refer to the phenomenon. But it's understood that this is just a shorthand. Successive cyclicity is forced, not by the Successive Cyclic Hypothesis, but more indirectly by the Subjacency Condition.

6.3 More evidence for successive cyclicity

In the previous section, we argued that 'long-distance' wh-movement did not proceed in one fell swoop. Instead, we argued that wh-movement proceeds in a successively cyclic fashion, via intermediate Spec CPs. However, the evidence that we saw for successive cyclicity was only indirect: it is only by moving from Spec CP to Spec CP that a wh-element can satisfy the Subjacency

Principle. English doesn't provide any direct, visible evidence for successive cyclicity.

However, once we move away from English, we see that other languages provide more direct support for the hypothesis that wh-movement proceeds successive-cyclically. Subject inversion in Spanish, and the choice of complementizer in languages such as Modern Irish, seem to confirm that wh-movement takes place in a clause-by-clause (CP by CP) fashion. In this way, we can confirm that Subjacency is a universal principle, which constrains wh-movement not just in English, but in all languages.

Subject inversion in Spanish

Torrego (1984) notes that Spanish, like many 'null subject' languages, allows 'free subject inversion', as illustrated in (38). (38a) represents the canonical subject-verb-object word order of Spanish. In this respect, Spanish is exactly like English. However, under certain discourse conditions, the subject can appear in a rightward position, as in (38b):

(38) a. Juan contestó la pregunta.
Juan answered the question.

b. Contestó la pregunta Juan.
'Juan answered the question.'

Interestingly, this subject inversion is in fact obligatory when certain wh-phrases appear in Spec CP, as illustrated in (39):

(39) a. Qué querían esos dos?
What wanted those two

b. *Qué esos dos querían?
'What did those two want?'

(40) a. Con quién vendrá Juan hoy?
With whom will come Juan today?

b. *Con quién Juan vendrá hoy?
'Who will Juan come with today?'

Obligatory subject inversion in Spanish takes place both in main and embedded clauses, as illustrated in (41) and (42):

(41) a. No sabía qué querían esos dos.
Not knew-1sg what wanted-3Pl those two

b. *No sabía qué esos dos querían.
'I didn't know what those two wanted.'

(42) a. Es impredecible con quién vendrá Juan hoy.
Is not predictable with whom will come-3sg Juan today

b. *Es impredecible con quién Juan vendrá hoy.
'It's impossible to predict who Juan will come with today.'

So, generalizing from the data in (39) to (42), it seems as though subject inversion is obligatory in any clause when Spec CP contains a wh-word of the appropriate type.

Now, let's look at cases involving 'long distance' wh-movement. Consider the following declarative sentence (with the subjects italicized):

(43) *Juan* pensaba que *Pedro* le había dicho que *la revista* había publicado
Juan thought that Pedro him had said that the journal had published
ya el artículo.
already the article
'Juan thought that Pedro had told him that the journal had published the article already.'

When forming a wh-question about *el artículo*, *all* subjects obligatorily undergo subject inversion, not just the matrix one:

(44) a. Qué pensaba *Juan* que le había dicho *Pedro* que había publicado
What thought Juan that him had said Pedro that had published
t *la revista?*
the journal
'What did Juan think that Pedro had told him that the journal had published?'

b. *Qué pensaba *Juan* que *Pedro* le había dicho que *la revista* había publicado?

If, on the other hand, the wh-word originates in a higher clause, inversion in the subordinate clause is not required:

(45) A quién$_i$ prometió *Juan* [$_{CP}$ que Pedro se encargaría
To whom promised Juan that Pedro self would-take-upon
[$_{CP}$ de que la gente sacara las entradas a tiempo]] t$_i$?
that the people got the tickets on time
'To whom did Juan promise that Pedro would make sure that the people got the tickets on time?'

Under the assumption that a wh-word in Spec CP is what triggers subject inversion, the contrast seen in (44) and the one between (44b) and (45) is explained if wh-movement proceeds successive-cyclically. When it passes through each intermediate Spec CP, it triggers subject inversion *with each subject along the way*. In (45), since the wh-word originated in the matrix clause and never passed

through any of the embedded Spec CPs, the subjects of the embedded clauses do not undergo inversion.

Complementizers in Irish

Additional support for the Successive Cyclicity Hypothesis can be found in languages such as Modern Irish. In these languages, when a wh-word has 'stopped off' in an intermediate Spec CP position, the form of the complementizer changes.

Chung and McCloskey (1987) note that the normal subordinating particle in Modern Irish is *go*, as illustrated by (46):[2]

(46) Dúirt sé go dtiocfadh sé.
Said he that would come he
'He said that he would come.'

However, when wh-movement has taken place, the complementizer *go* is replaced by *aL*, as illustrated by (47):[3]

(47) an bhean$_i$ aL chuir t_i isteach air
the woman that put in on it
'the woman who applied for it'

Crucially for our purposes, when Subjacency would predict that successive cyclic movement must have taken place, it turns out that every complementizer along the way changes to *aL*:

(48) an rud$_i$ aL shíl mé aL dúirt tú aL dhéanfá t_i
the thing Comp thought I Comp said you Comp would do
'the thing that I thought you said that you would do'

This would seem to be further evidence that wh-movement (in this case, inside a relative clause) is proceeding in a successively cyclic manner.

6.4 Subjacency and the analysis of relative clauses

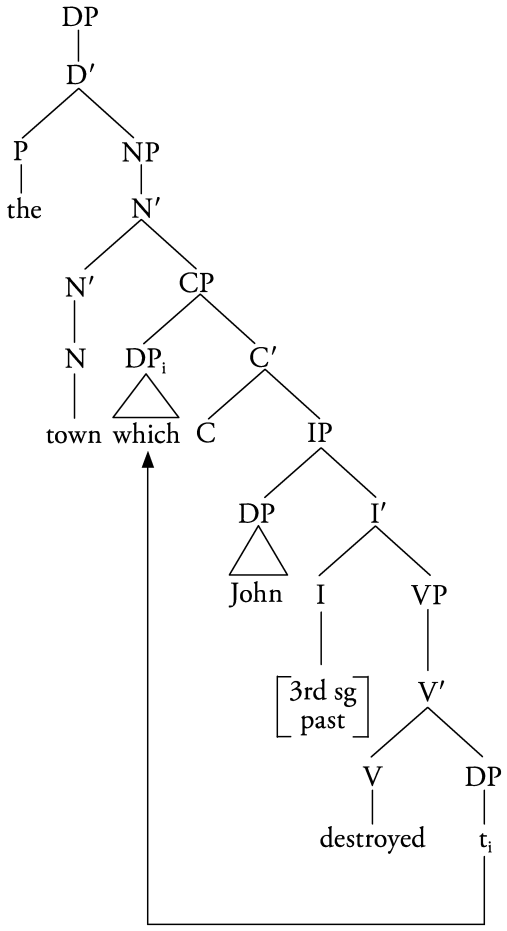
One of the interesting things about constraints on movement like Subjacency is that you can use them rather like sub-atomic particle detectors. You might not be able to see the particle, but the detector registers that a particle is there. We'll see this in a much more radical way in the next chapter when we talk about the concept of Logical Form, but for the moment we can see the process at work in relative clauses. Some of the trees we drew back in Section 6.2 sort of hinted at an analysis of relative clauses, but I now want to come back and discuss this more explicitly.

Let's look first at (49):

(49) The town which John destroyed

Given the discussion that we've just had about wh-movement, it will not surprise you to learn that we'll analyse (49) as containing an instance of wh-movement.

(50)



Just as we did with embedded wh-movement structures, we assume that *which* in (49) is base-generated as the direct object of *destroy*, and then moves to the specifier of a CP which is the sister to the N′ *town*.

However, relative clauses can be formed in different ways in English, and they don't always, at least on the surface, have a wh-word. Consider (51) and (52):

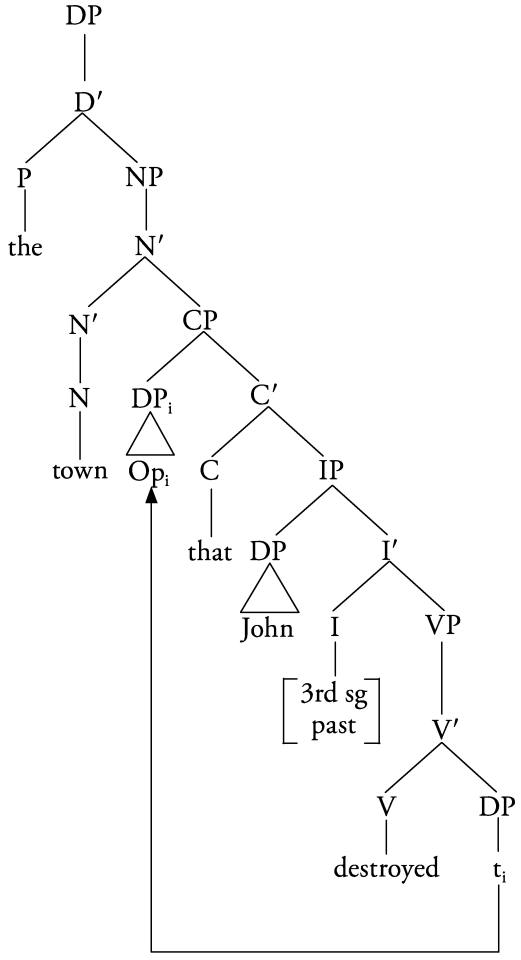
(51) The town that John destroyed

(52) The town John destroyed

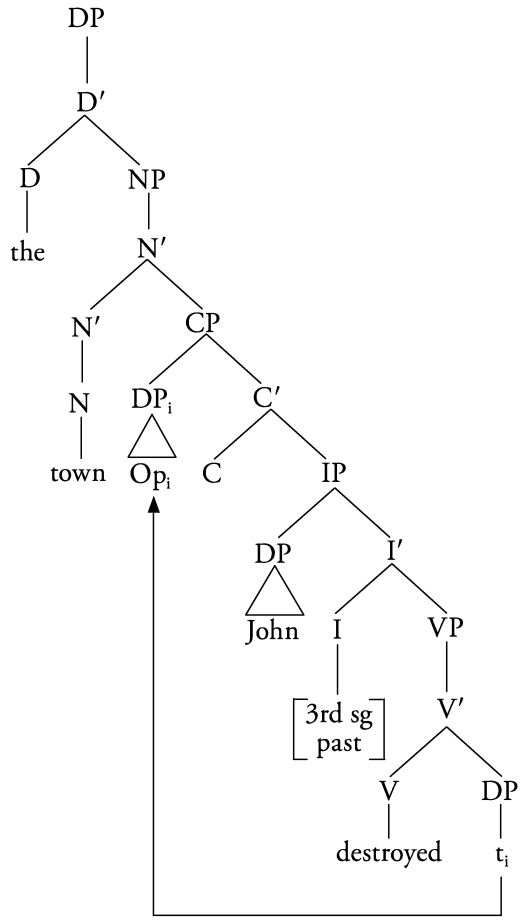
In my dialect of American English, both (51) and (52) are entirely equivalent to (49). I think there's some prescriptive rule about when you're supposed to use *which* and when you're supposed to use *that*, but I can never remember what it is, and I don't really care anyway. (Remember we're interested in *description*, not *prescription*.) What's interesting about (51) and (52) is that, in both cases, *destroy* seems to be lacking a direct object, but, unlike (49), there's nothing around that could obviously be it. (51) appears to have the complementizer *that* instead of a wh-word, and (52) has nothing at all.

The introduction to this section probably gives the game away, but the solution we're going to propose is that both (51) and (52) *do* have wh-movement. What's different is that it is movement of a null, that is an unpronounced, wh-element. For reasons that will become clearer in the next chapter, null wh-elements of this kind are often referred to as *null operators* and abbreviated *Op*. So the trees for (51) and (52) are really (53) and (54):

(53)



(54)



So how can we argue that an invisible wh-word is undergoing invisible wh-movement in relative clauses like (51) and (52)? We can use Subjacency. If wh-movement is *really* happening, then it ought to be constrained by the sort of things that constrain wh-movement. And what's really kind of cool is that that looks to be the case.

Let's start with a relative clause that *does* have a wh-word. Not surprisingly, it does look like relative clause formation is constrained by Subjacency in the way that you'd expect:

(55) *the town which$_i$ Bill wondered how$_j$ John destroyed t_i t_j

(55) seems to be more or less on a par with a 'standard' instance of Subjancency-violating wh-movement:

(56) *What$_i$ did Bill wonder how$_j$ John destroyed t_i t_j

In both cases, the wh-element needs to stop off in the specifier of the CP which is the sister of *wonder*, but can't because another wh-word has blocked it up. It is forced to move to the specifier of the matrix CP in one fell swoop, and thus violates Subjacency. For reasons which we'll discuss in Chapter 8, if you try to relativize the subject of the embedded clause, rather than the object, (55) and (56) get even worse:

(57) a. *The man who$_i$ Bill wondered how$_j$ t_i destroyed the town t_j
b. *Who$_i$ did Bill wonder how$_j$ t_i destroyed the town t_j?

What's really interesting is that you get exactly the same judgements even in relative clauses that don't contain an overt wh-element:

(58) a. *the town that Bill wondered how John destroyed
b. *the man that Bill wondered how destroyed the town

(59) a. *the town Bill wondered how John destroyed
b. *the man Bill wondered how destroyed the town

Whether the relative clause is introduced by complementizer *that* or by nothing at all you get results which are fully parallel to relative clauses with overt wh-movement. The (a) sentences are bad and the (b) sentences are a lot worse. Relative clause formation therefore seems to involve an element which undergoes wh-movement, even when that element is an unpronounced null operator. In other words, (58) and (59) really make it look like there *are* invisible wh-words (null operators) that *do* undergo 'invisible' wh-movement.

6.5 More transformations: passive and raising

So far in this chapter we've concentrated exclusively on wh-movement when discussing transformations. Movement of this kind is called *A'-movement* because it involves movement to an A'-position, Spec CP. However, there are other sentences in natural language which involve movement to the specifier of IP, an A-position. This kind of movement is called *A-movement* and we'll see that those phenomena, too, are best explained by transformations.

Passive

One familiar case of A-movement from traditional grammar instruction is that of passivization. Corresponding with an active sentence like (60), there is a passive counterpart (61):

(60) Big Business killed music.

(61) Music was killed by Big Business.

(60) and (61) are essentially synonymous. The active verb *kill* is replaced by its passive counterpart *be killed*, the direct object of the active verb appears as the subject of the passive, and the former subject of the active version appears in

a PP. The question is: what is the relation between an active sentence and its passive counterpart? Just as in the cases of wh-movement, there are in principle two options. It could be that passive sentences are directly base-generated by our phrase-structure rules. The alternative is that sentences like (61) are derived by transformation from an underlying D-Structure.

In order to decide between these two options, let's take a closer look at some of the less obvious properties of passive sentences. Recall from Chapter 4 that we discussed so-called *expletive* elements, elements which must not be assigned a θ-role. English has two expletives, *it* and *there*, depending on whether the associate of the expletive is a CP or a DP. Either one can appear in Spec IP when the verb has been passivized.

(62) a. There was a man killed.
b. It is believed that John is a fugitive from justice.

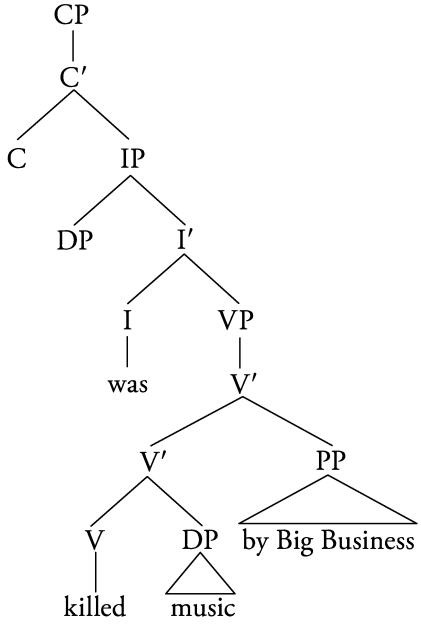
From the grammaticality of (62a) and (62b), we conclude that Spec IP is not a θ-position in passives. That is, the subject of a passive verb is not a position to which a θ-role is assigned.

But if that's the case, that would seem to be an argument against the base-generation hypothesis for passives. The θ-criterion tells us that all DPs need to be assigned a θ-role, but if *music* is base-generated in the subject position in (61), it won't receive one and we'd expect (61) to be ungrammatical, on a par with (63):

(63) *John seems that Bill is here

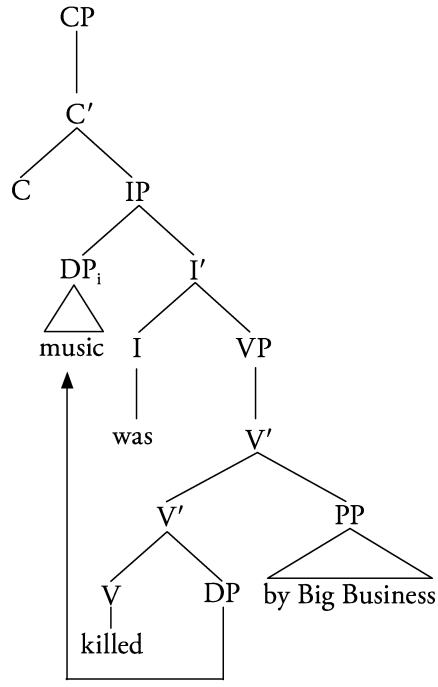
If, on the other hand, we assume that passive sentences are transformationally derived, we don't have this problem. A sentence like (61) can be derived from a D-Structure like (64):

(64)



Since *music* is base-generated as the direct object of the verb, that's the θ-role that it receives.[4] The passive transformation then moves the direct object to the empty subject position, generating the order that we see on the surface:

(65)



However, there's one more thing we need to say about passives. On the basis of the discussion so far, we don't have an explanation for (66):

(66) *There was killed music by Big Business

It seems as though the base-generated object can't just stay in its D-Structure position. It must move to subject position:[5]

(67) Music was killed by Big Business.

A seemingly natural way to explain this is to say that, in addition to not assigning a θ-role to their subjects, passive verbs also don't assign Accusative Case to their objects. (In fact, in what's come to be known as 'Burzio's Generalization', Luigi Burzio (1986) suggested these two properties are intrinsically connected. We'll come back to this in Chapter 9.)

In this way, the following picture of passives emerges: the direct object, base-generated in a position that receives a θ-role, but not Case, needs to move to a position to which Case is assigned. By the θ-criterion, this Case position can't be a θ-position, since the object's already got its one θ-role. Fortunately, the

subject position in passives provides just such a position (Case, but no θ-role), and so the direct object moves there.

Raising

In addition to passive verbs, there's another class of verbs (and adjectives, actually) which also have properties which seem to be best explained by a transformational analysis. These predicates are called *raising* predicates. A canonical raising predicate is the verb *seem*. In Chapter 4, we saw *seem* in its ordinary 'I take a CP complement' guise, as illustrated by (68):

(68) It seems [$_{CP}$ that six executives like Guided by Voices].

However, *seem* has another guise, in which it takes an IP infinitival complement, as exemplified by sentences like (69):

(69) Six executives seem [$_{IP}$ to like Guided by Voices].

Just as it does when it takes a CP complement, the subject position of *seem* is a non-θ-position when it takes an IP complement:

(70) There seem [$_{IP}$ to be six executives here].

We therefore conclude that the DP *six executives* cannot have been base-generated in subject positions in sentences like (69). Instead, the DP is base-generated as the subject of the infinitive, where it receives its θ-role as the subject of *like* and is subsequently 'raised' to the subject position of *seem*.

(71) Six executives$_i$ seem [$_{IP}$ t$_i$ to like Guided by Voices].

An example of a raising adjective is *likely*. Just as with raising verbs, raising adjectives can take a finite CP complement:

(72) It is likely that six executives enjoy Guided by Voices.

(72) can alternate with an infinitival IP complement, as in (73):

(73) Six executives are likely [$_{IP}$ to enjoy Guided by Voices].

The subject is still a non-θ-position:

(74) There are likely to be six executives here.

We therefore conclude that the subject *six executives* in (73) was raised from the subject position of the infinitive, as illustrated by (75):

(75) Six executives$_i$, are likely [$_{IP}$ t$_i$ to enjoy Guided by Voices].

EXERCISE 3

Consider the following four sentences:

(a) John promised to go.
(b) John appeared to go.
(c) John is certain to go.
(d) John intends to go.

In each case, *John* is interpreted (i) as the promisor, appearer and so on, and (ii) as the goer. Using the θ-criterion and expletives as a test, discuss which of the four are raising predicates and which are not.

6.6 Levels of application of principles

We've been exploring the idea that phrase-structure rules generate D-Structures which are then operated on by transformations to give S-Structure. Given this revision of our model, there's a question we need to ask: now that we have two levels, D-Structure and S-Structure, which level or levels do Case marking, Binding Theory, θ-marking and so on apply to? This question didn't arise before, since in previous chapters we were dealing only with one level, S-Structure.

We've mentioned a couple of cases in passing already. It was examples involving Binding Theory that we used to motivate the existence of transformations in the first place. Remember (6)?

(6) Which picture of himself$_i$ does John$_i$ like t?

At S-Structure, *John*, the antecedent of the reflexive *himself*, fails to c-command it. (The first branching node dominating *John* is the matrix IP, which does not dominate Spec CP.) It is only at D-Structure that the relevant c-command relationship exists between *John* and *himself*. It must therefore be the case that Principle A of the Binding Theory applies at D-Structure.

θ-marking is another process which must take place at D-Structure rather than S-Structure. As we saw in Section 6.5, passive and raising move elements into positions which are not assigned a θ-role, as evidenced by the fact that these positions allow expletive *it* and *there*:

(76) There were three men killed.

(77) It is likely that John is here.

If θ-marking took place at S-Structure, we would expect sentences like (78) and (79) to violate the θ-criterion:

(78) Three men were killed t.

(79) John is likely t to be here.

At S-Structure, *three men* and *John* are not in positions which are assigned a θ-role, and would be predicted to be ungrammatical. We therefore conclude that θ-marking must take place at D-Structure.

What about subject-verb agreement? This presents an interesting case, and one which will lead us directly into the next section, where we'll clarify the relationship between moved elements and their traces. We argued in Section 6.1 that certain facts about subject-verb agreement provided an argument for transformations. Recall the contrast in (13):

(13) a. Which car did Bill say needs paint?
b. *Which car did Bill say need paint?

It is only at D-Structure that the requisite Specifier-Head relationship exists between the wh-phrase which is in the lower Spec IP position and I. We therefore concluded that, in order to account for choice of *needs* over *need*, subject-verb agreement had to take place at D-Structure.

However, if we take a close look at the cases of passive and raising that we saw in Section 6.5, they seem to argue for the opposite conclusion. The D-Structure representations for (78) and (79) are of course (80) and (81):

(80) were killed three men

(81) is likely John to be here

Unlike the contrast seen in (13), in (80) and (81) the requisite Specifier-Head relationship doesn't exist at D-Structure. It only exists at S-Structure.

So we seem to have a paradox: some instances of subject-verb agreement must take place at D-Structure, while others must take place at S-Structure. However, this paradox can be resolved by taking a closer look at the relationship between a moved object and its trace.

6.7 Chains

In this section we're going to introduce the idea of a *chain*, a technical concept concerning the relation between a moved element and its trace which plays a central role in the theory of movement. For purposes of discussion, consider the raising structure in (82):

(82) John_i seems t_i to be here.

John is a DP. The trace too must be a DP, given its distribution (i.e. it appears only in DP positions: Spec IP, sister to V, sister to P, etc.). However, the relation between an antecedent and its trace is not just an ordinary relationship between two coindexed DPs, like the one in (83):

(83) John$_i$ said that he$_i$ was late.

Syntactic requirements which hold of DPs treat the moved element and its trace as though they were a single DP. For this reason, we say that the moved element and its trace form *a chain*. In (82) the chain is created transformationally by the raising of *John*. The chain has two members, *John* and the trace. In the case of a single DP which simply remains in its base-generated position, we say that that DP forms a *one-membered chain*.

Notice that this will require that the θ-criterion and the Case Filter be redefined in terms of chains, but this is easily done. Chomsky (1986: 97) provides the following chain-based definition of the θ-criterion:

(84) *The θ-criterion*
Each argument A appears in a chain containing a unique ... θ-position P, and each θ-position P is ... a chain containing a unique argument A.

And the Case Filter can be revised as follows:

(85) *The Case Filter*
Each chain must be assigned Case.

The concept of the chain, and the reformulation of the θ-criterion to apply to chains, explains a multitude of otherwise odd facts. It first explains why, as we saw above, *John* can (and indeed must) move to a position that doesn't assign a θ-role. *John* itself doesn't need an independent θ-role in (82) because it's linked up with a trace which already has one. Additionally, we also explain the contrast in (86):

(86) a. John$_i$ seems t$_i$ to be here.
b. *John$_i$ believes t$_i$ to be here.

If *John* already has a θ-role by virtue of its association with the trace, (86b) is straightforwardly ruled out as a θ-criterion violation, defined in terms of chains, as the chain (*John*, t) receives *two* θ-roles.

We also resolve the paradox that we mentioned at the end of Section 6.5 regarding subject-verb agreement. In (13) (repeated below), which involved wh-movement, it seemed as though subject-verb agreement took place at D-Structure:

(13) a. Which car does Bill say needs paint?
b. *Which car does Bill say need paint?

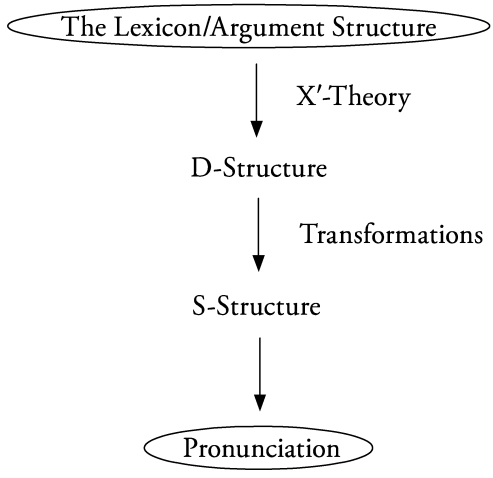
In the case of passive and raising, it seemed as though subject-verb agreement must take place *after* passive or raising, that is, at S-Structure:

(87) a. $John_i$ is likely t_i to be here.
b. *$John_i$ are likely t_i to be here.
c. *[John and Mary]$_i$ is likely t_i to be here.
d. [John and Mary]$_i$ are likely t_i to be here.

However, the notion of the chain allows us to have elements be in two places at once, in effect, at S-Structure. We can then say that subject-verb agreement takes place at S-Structure and that agreement can take place either with the head of the chain (the moved element) or the tail of the chain (the trace). In the case of wh-movement, it's the tail of the chain that is in the embedded Spec IP, but the embedded INFL agrees by virtue of the trace being in a chain with *which car*. In (87), by contrast, the matrix INFL is agreeing with the head of the chain (John, t), which is *John*.

Summary

This chapter has motivated one of the most significant additions so far to our model of what speakers of a language know. One thing that speakers seem to know about their language is that certain sentences are not simply base-generated. Rather, they are derived from an underlying D-Structure representation. Transformations apply to this D-Structure representation, generating the S-Structure representation, which is what actually undergoes pronunciation. Schematically, we have the following model:



Notice that the facts we've discussed – facts which are intuitively obvious to any speaker of a language – again constitute clear examples of Plato's Problem. It appears to be the case that in order to account for what we know when we know a language, we need to say that speakers don't pay attention to just what there is on the surface. Rather, speakers know that sentences have multiple representations. There's a D-Structure representation, which in certain constructions (like raising) never appear on the surface. Then there are also transformations which map these 'invisible' D-Structures into the S-Structures that are pronounced.

In particular, we looked at several different transformations, including wh-movement, passive and raising. We saw that wh-movement in particular is subject to the Subjacency Principle:

(37) *The Subjacency Condition*
A single instance of wh-movement can cross only one bounding node, where IP and DP are bounding nodes.

This principle forces wh-movement to proceed successive-cyclically, moving in stages from Spec CP to Spec CP. Although English doesn't provide any direct evidence for successive cyclicity, other languages like Spanish and Irish do.

It also turned out that the relationship between an element moved by transformation and the trace it leaves behind was unlike anything we've seen before. It seemed as though the θ-criterion and the Case Filter treat the moved DP and its trace as a single DP. As a result we introduced the notion of *a chain* and were led to modify our definitions of the θ-criterion and the Case Filter from Chapter 4:

(84) *The θ-criterion*
Each argument A appears in a chain containing a unique ... θ-position P, and each θ-position P is ... a chain containing a unique argument A.

(85) *The Case Filter*
Each chain must be assigned Case.

Open issue: parametric variation and bounding nodes in Italian

In Section 6.2 we motivated the Subjacency Condition:

(37) *The Subjacency Condition* (final version)
A single instance of wh-movement can cross only one bounding node, where IP and DP are bounding nodes.

However, Rizzi (1982) noticed that certain Romance languages, such as Italian, seem to present a problem for the Subjacency Condition as we've formulated it.

Consider the following sentence:

(88) Tuo fratello, [$_{CP}$ a cui$_i$ [$_{IP}$ mi domando [$_{CP}$ che storie [$_{IP}$ abbiano
Your brother, to whom myself I-ask which stories have-3pl
racontato t_i]]]], era molto preoccupato.
told was very worried
'Your brother, to whom I wonder which stories they have told, was very worried.'

In particular, look at the subject DP, *your brother, to whom I wonder which stories they have told*. The English translation violates Subjacency – just like example (89), which we saw earlier:

(89) *What $_i$ did [$_{IP}$ John wonder [$_{CP}$ why [$_{IP}$ Bill hit t_i]]]?

The movement crosses two bounding nodes (in this case, two IPs), but, unexpectedly, the result is grammatical in Italian.

There are a couple of different options we could pursue at this point. We could simply abandon the idea that Subjacency is a universal principle and content ourselves with the work that it does for us in English and other languages. However, a brief look at some other data in Italian shows that this would be the wrong decision. Although the kind of Subjacency violations we saw in (89) don't seem to exist, violations involving 'complex' DPs like (33) above do:

(90) *Tuo fratello, [$_{CP}$ a cui$_i$ [$_{IP}$ temo [$_{DP}$ la [$_{NP}$ possibilità
Your brother, to whom fear-1sg the possibility
[$_{CP}$ che [$_{IP}$ abbianoraccontato tutto t_i, ...]]]]]
that have-3pl told everything, ...
'Your brother, to whom I fear the possibility that they have told everything, ...'

In (90) the wh-phrase *a cui* has crossed both an IP node and a DP node, and, as Subjacency would predict, the result is ungrammatical. So it seems clear that Subjacency needs to be modified rather than completely abandoned.

One obvious way to modify Subjacency would be to say that IP isn't a bounding node in Italian. What really matters in Italian is crossing a DP, and crossing only one bounding node will result in ungrammaticality:

(91) *Italian Subjacency (proposed)*
Movement cannot cross even one bounding node (where the bounding nodes are DP only).

(91) makes the correct predictions about (88) vs (90). Movement crosses no DP nodes in (88) and so is correctly predicted to be grammatical. Movement crosses one DP node in (90) and so is correctly predicted to be ungrammatical.

However, (91) also predicts that movement can cross as many IPs as you like, provided that no DP is crossed. That prediction isn't correct. Consider the following subject NP:

(92) *Francesco, [$_{CPx}$ che$_i$ [$_{IPx}$ non immagino [$_{CPy}$ quanta gente [$_{IPy}$ sappia
Francesco, who not imagine-1sg how-many people know
[$_{CPz}$ dove [$_{IPz}$ hanno mandato t_i, ...]]]]]]
where have-3P1 sent
'Francesco, who I can't imagine how many people know where they have sent, ...'

In (92), *who* has had to move in one fell swoop from its position as the direct object of *send* to Spec CP_x, because the intermediate Specs CP_y and CP_z are filled by *where* and *how many people* respectively. There's no DP node crossed here, only IPs, so (91) predicts that (92) should be grammatical. And that's not the right result.

So we're left with something of a mystery: what is it that distinguishes (88) from (90) and (92)? A closer look at the structure indicated in (88) suggests the answer. Notice that, while two IP nodes were crossed by *a cui*, only one CP node was crossed. (Remember the highest CP node isn't *crossed*, since it dominates both the departure site (sister to V′) and the landing site (Spec CP).) In the ungrammatical (92), *two* CP nodes are crossed (CP_y and CP_z).

With this insight, Rizzi observed that Subjacency in Italian can be accounted for with a minimum amount of fuss. The only difference between English and Italian has to do with which nodes are bounding nodes. Thus, we have (93):

(93) *Italian Subjacency (final)*
Movement cannot cross more than one bounding node (where the bounding nodes are DP and CP).

So Subjacency seems to be subject to parametric variation. For a language like English, the bounding nodes are set to DP and IP, but for other Romance languages like Italian, the bounding nodes are DP and CP. This suggests that, while Subjacency is indeed a universal principle, the precise details are subject to parametric variation.

FURTHER READING

The idea that there are structure-changing operations like transformations is one of the oldest ideas in Chomskyian generative linguistics, going back to Chomsky (1957) (although similar ideas can be found in other traditions hundreds of years earlier). The first observations about what we've been calling 'Subjacency' effects go back to John Robert Ross's (1967) dissertation, where he identified what he called 'islands'. As with some of the other topics we've discussed, any textbook on Government and Binding Theory will have discussion of these ideas, including Lasnik and Uriagereka (1988) and Haegeman (1994).

ANSWERS TO IN-TEXT EXERCISES

Exercise 1

We can construct sentences which are entirely parallel with those in Section 6.1 to show that the sentences in (20) cannot be base-generated, but instead must be derived by transformation. Consider (i):

(i) John wondered [which picture of herself]$_i$ Mary liked t_i.

Although (i) is grammatical, the required c-command relationship would not exist between *Mary* and the anaphor *herself* if the sentences were base-generated. On the other hand, if (i) is transformationally derived from the D-Structure in (ii), then Binding Theory can be satisfied. At D-Structure, *Mary* does c-command *herself*:

(ii) John wondered [$_{CP}$ Mary liked which picture of herself].

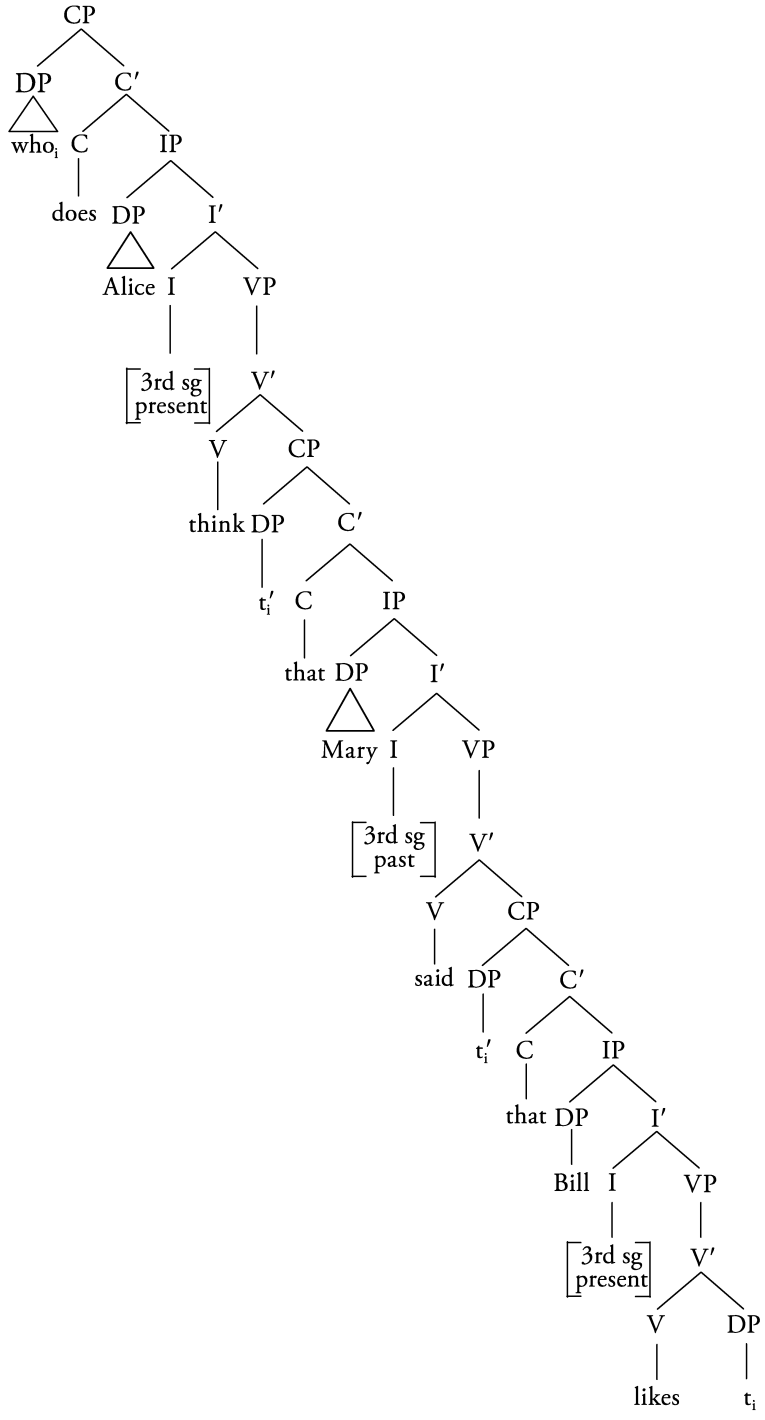
A similar situation exists for subject-verb agreement:

(iii) John wondered which car Bill thinks needs paint.
(iv) *John wondered which car Bill thinks need paint.

Unless sentences like (20) are transformationally derived and not base-generated, it's not clear why there should be a difference between (iii) and (iv).

Exercise 2

Clearly, if we accept (26), we're going to need one application of wh-movement for each CP. (26) says that wh-movement first moves an element to the nearest Spec CP, and then from Spec CP to Spec CP. Therefore, in (18) we need three applications of wh-movement and in (19) we need four.


CP
DP
C′
who$_i$
C
IP
does
DP
I′
Alice
I
VP
[3rd sg present]
V′
V
CP
think
DP
C′
t$_i$′
C
IP
that
DP
I′
Mary
I
VP
[3rd sg past]
V′
V
CP
said
DP
C′
t$_i$′
C
IP
that
DP
I′
Bill
I
VP
[3rd sg present]
V′
V
DP
likes
t$_i$

Exercise 3

If the subject position of the verb allows an expletive, then it does not assign a θ-role. Therefore, by the θ-criterion, the predicate must be a raising predicate, with the subject getting its θ-role as the subject of the lower verb. Therefore, we have:

(ia) *It promises that John will leave.
(ib) *There promises to be a man here.

Conclusion: *promise* is not a raising predicate.

(iia) It appears that John will leave.
(iib) There appears to be a man here.

Conclusion: *appear* is a raising predicate.

(iiia) It is certain that John will leave.
(iiib) There is certain to be a man here.

Conclusion: *certain* is a raising predicate.

(iva) *It intends that John will leave.
(ivb) *There intends to be a man here.

Conclusion: *intend* is not a raising predicate.

FURTHER EXERCISES

Exercise 1

Draw the D-Structure and S-Structure trees for the following sentences:

1. Which book about linguistics did Mary write?
2. Why did John think that Bill left the room?
3. What did John believe Bill said that Mary had written?
4. On which postcard has John drawn a picture of himself?
5. For John to appear to be lying would be a political disaster.
6. Toby intended to be seen at the party.

Exercise 2

Put traces in the appropriate places and identify the chains in the following sentences. How does each DP in the sentence satisfy the revised θ-criterion and Case Filter?

1. John seems to have left the building.
2. The duck salesmen were attacked in the streets yesterday.

3. What did John say that he put on the shelf?
4. I believe Bob to have destroyed Carthage.
5. Who was killed?
6. Who does John seem to like?

Exercise 3

There's a kind of construction which we mentioned in Chapter 2, but haven't looked at explicitly in this chapter, called Topicalization. It's exemplified by sentences like (1) and (2):

1. This book, John bought yesterday.
2. Beatles records, Bill likes.

Do you think that these sentences are directly generated by our phrase-structure rules or derived by a transformation?

Exercise 4

One useful thing that Subjacency can do is provide a diagnostic test for phrase structure. Consider the following extraction from a sentential subject:

(1) *What$_i$ did that Brian recorded t_i yesterday bother Mike?

Examine the derivation. Does any application of wh-movement lead to a Subjacency violation? How could we modify our analysis of the phrase structure of sentential subjects in order to make the correct prediction with respect to (1)?

Exercise 5

We argued in this chapter that Binding Theory is checked at D-Structure on the basis of sentences like (1):

(1) Which picture of himself does John like?

However, consider the interpretive possibilities for *himself* in a sentence like (2):

(2) Which picture of himself did John say that Bill liked?

Can *himself* refer to *John*, *Bill* or either one? What problems, if any, do the co-reference possibilities present? Consider the derivation and whether *himself* would be bound in its Governing Category by its antecedent at D-Structure.

Exercise 6

The examples which show that Binding Theory applies at D-Structure have so far all come from wh-movement. However, consider the following data from raising verbs:

(1) John$_i$ seems to me to like himself$_i$.
(2) They$_i$ seem to each other$_i$ to be good players.

What impact do (1) and (2) have on the claim that Binding Theory is checked at D-Structure? How can the concept of the chain help us here?

Exercise 7

Consider the following instances of long-distance passive and raising:

(1) John was believed to have been attacked.
(2) Bill seems to be likely to lose the race.

Do passive and raising proceed successive-cyclically? Are they subject to Subjacency too? Try to construct sentences in which 'one fell swoop' passivization and raising are forced because a potential intermediate 'stopping off' point is filled and see what the result is.

Logical Form

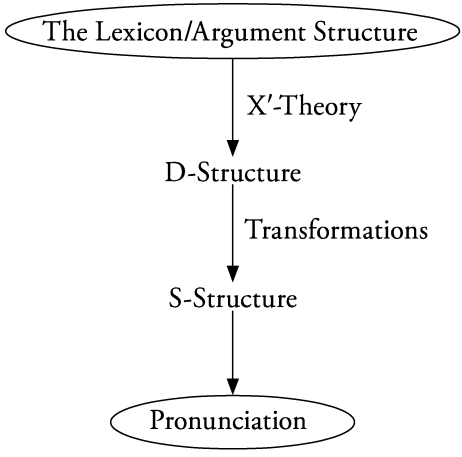
7

In the previous chapter we introduced the idea of transformations and discovered that part of what you know about your language includes the fact that sentences have two kinds of representations: S-Structures, in which the words appear in the order in which they're pronounced, and D-Structures, which interface with the lexicon and encode information about the argument structure of the predicates. S-Structures and D-Structures are linked by transformations.

This might have seemed like a pretty radical step. What we claimed is that part of what you know about your language includes information about D-Structure representations – representations which *never* appear on the surface in many cases. And even S-Structure representations, in which the word order corresponds to the 'surface' word order, often have traces in them – elements which are present but which are *never* pronounced.

In this chapter, we're going to take this idea one step further. Not only are there *representations* which you never 'see', but there are also *transformations* whose effects you never 'see'. That is, there are transformations which are not concerned with deriving the surface order that is pronounced. In introducing this idea, we're going to modify our assumptions about the architecture of the grammar once again. (Don't worry! This is the last time.) So far, we have (1):

(1)

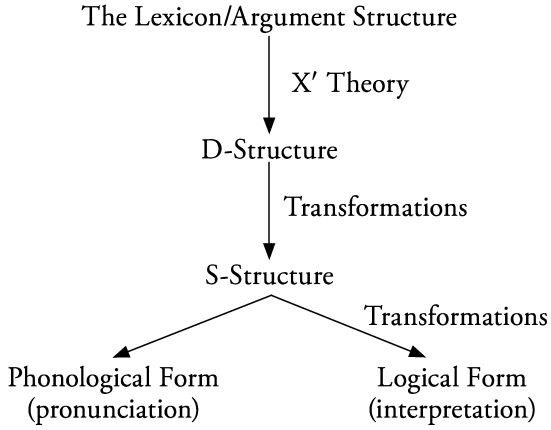


The structure-building operation, X′-Theory, together with the lexicon and argument structure of the words involved, creates D-Structure. D-Structure is then modified by transformations (if appropriate) to create S-Structure. S-Structure is then used as the input to phonology and pronunciation.

What we're going to claim in this chapter is that, after S-Structure, the computation 'splits', rather than going directly to pronunciation. As before, the phonological information is stripped out of the S-Structure representation and sent to the phonological component for processing. In the phonological component, intonational contours are assigned and other phonological operations happen, ultimately yielding a representation of pronunciation. What's new in this chapter is the idea that the *syntactic* computation (minus phonological features) actually continues, and specifically that there are other transformations which continue to manipulate the structure. When those transformations have been completed, we arrive at a level known as *Logical Form*, which is then passed on to the semantic component for semantic interpretation. *Logical Form* (LF) therefore serves as the 'interface' between the syntax and the semantics, just as S-Structure serves as the interface between the syntax and the phonology.

Therefore, instead of (1), what we really have is (2):

(2)



This is known as the 'Y' model of grammar, because the shape of it looks something like an upside-down Y.

In Section 7.1, we'll take a look at Chinese and Japanese, particularly as they compare with English. Chinese and Japanese are languages in which wh-words *appear* not to undergo movement to Spec CP. That is, unlike in English, there are no wh-movement effects visible in pronunciation. There is no *overt* movement. However, we'll argue that in those languages wh-words *do* undergo wh-movement, but they do so *covertly*, after S-Structure in the LF component. We'll then go back to English and argue that English has cases of LF wh-movement too. In Section 7.2, once the idea of LF movement is more familiar, we'll talk a little more generally about Logical Form as a level of representation, in the context of another LF transformation, quantifier raising.

Both LF wh-movement and quantifier raising will turn out to be instances of covert A′-movement, so in Section 7.3 we'll turn our attention to LF

A-movement, arguing that there's an LF-counterpart to raising in English. We'll claim that in structures of the form *there ... DP*, the DP raises to the position of *there* in the LF component.

7.1 Wh-movement in the LF component

In this section, we're going to take a look at two cases of LF wh-movement; that is, wh-movement which takes place but which you don't 'see'. In languages like Chinese and Japanese, following a line of analysis started by Huang (1982), it turns out that *all* wh-movement happens in the LF component. The idea of LF-movement differs from the null-operator analysis of relative clauses that we proposed in the previous chapter, because in these cases the wh-words aren't 'invisible'. They're there. They just don't look like they've moved anywhere. But we're going to argue that they do.

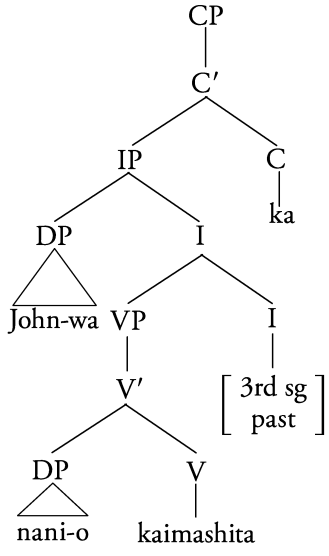
LF wh-movement in Chinese and Japanese

Chinese and Japanese represent two languages in which, on the surface, wh-elements like *what* and *why* do not undergo wh-movement. Consider the following sentence from Japanese:

(3) John-wa nani-o kaimashita ka?
John-Top what-Acc bought
'What did John buy?'

As you'll recall from Chapter 3, Japanese is a head-final language. Therefore, *nani-o* 'what' is in a standard direct object position as a sister to V, as illustrated by (4):

(4)



However, notice that *nani* (a wh-word in Japanese) has not undergone any wh-movement in (3), despite (3) being the normal way to ask a wh-question in Japanese. The wh-word remains *in situ*, that is in its D-Structure position. We can see this when we replace *nani* with a non-wh-word, as in (5):

(5) John-wa hon-o kaimashita.
John-Top book-Ace bought
'John bought a book.'

In this section, we'll argue that, despite this apparent lack of *overt* wh-movement, Chinese and Japanese have *covert* wh-movement to Spec CP. Put another way, Chinese and Japanese are essentially identical to English with regard to their wh-questions. The difference between them is simply a matter of whether wh-movement takes place between D-Structure and S-Structure (overtly) or between S-Structure and LF (covertly).

There are two kinds of arguments that we'll look at to support the idea of LF wh-movement. The first will be a more conceptual, cross-linguistic kind. In looking at other languages, we'll see that they have overt movements which correspond to the covert ones that we're postulating (as in English, for instance). We'll also try to make some of the same arguments for LF wh-movement in Chinese and Japanese as we made for *overt* wh-movement in English in the last chapter. However, we'll then try to go one better. We'll see that, even though the elements in question don't seem to move anywhere, the elements are behaving as if they did move and are subject to some familiar constraints on movement.

An argument from selection

By way of introduction, let's take a quick look at an English sentence of a type we discussed in the previous chapter:

(6) John wondered what$_i$ Bill bought t_i.

Remember that we want to say that wh-movement is taking place in (6) even though (6) as a whole does not have the semantics of a question. It's just a lexical fact about *wonder* that it requires a wh-word in the specifier position of the CP that is its sister. That's just part of the meaning of that verb. Other verbs which can take a +wh CP all fall into roughly the same semantic area, involving states of knowledge.

(7) a. I asked what$_i$ John bought t_i.
b. I know what$_i$ John bought t_i.

Let's now contrast this situation in English with the situation in a language like Chinese. The Chinese translation of (6) is (8) (with the proper names changed to appropriate ones for the language, as is commonly done):

(8) Zhangsan xiang-zhidao [$_{CP}$ [$_{IP}$ Lisi mai-le shenme]].
Zhangsan wonder Lisi bought what
'Zhangsan wonders what Lisi bought.'

The obvious difference between (6) and (8) is that wh-movement does not appear to have taken place in (8), just like the Japanese case that we saw in (4) above. In the S-Structure in (8), *shenme* is in the same position that it's in at D-Structure. It has not moved to Spec CP.

How do we account for this difference between Chinese and English? We could say that the verb *xiang-zhidao* in Chinese is different from *wonder* in English. *Wonder* simply has a semantic requirement that *xiang-zhidao* lacks. *Wonder* requires a wh-word in its Spec CP, but *xiang-zhidao* does not.

The problem with that solution is twofold: first, it seems a little odd to say that *wonder* and *xiang-zhidao* differ to such a radical extent. Whatever precise semantic property it is which requires *wonder* to select a +wh CP is surely some part of its core meaning, which is not something that we'd expect to differ across languages. However, more concretely, the hypothesis that verbs like *xiang-zhidao* in Chinese don't select a +wh CP also seems wrong when you consider how (8) contrasts with (9) and (10). (8) is a statement, (9) is a question and (10) is ambiguous between a statement and a question:

(9) Zhangsan yiwei Lisi mai-le shenme?
Zhangsan think Lisi bought what
'What does Zhangsan think that Lisi bought?'

(10) Zhangsan jide Lisi mai-le shenme?
Zhangsan remember Lisi bought what
'Zhangsan remembers what Lisi bought.' (statement meaning)
or 'What does Zhangsan remember that Lisi bought?' (question meaning)

How can we explain the different interpretive possibilities in (8), (9) and (10)? In every case, *shenme* is in the same place. Why is (8) a statement, (9) a question and (10) ambiguous between the two?

Well, as we saw in the last chapter, we need to go beyond what we see on the surface. Notice that with respect to the three verbs in English, *wonder* requires that its CP sister is +wh, *think* requires that its CP sister is –wh, and *remember* can take either a +wh CP sister or a –wh one:

(11) a. I wonder [what John bought]. +wh
b. *I wonder [that John bought a car]. –wh

(12) a. *I think [what John bought]. +wh
b. I think [that John bought a car]. –wh

(13) a. I remember [what John bought]. +wh
b. I remember [that John bought a car]. –wh

This looks like the key to understanding (8), (9) and (10). Let's say that *xiang-zhidao*, *yiwei*, and *jide* all have exactly the same semantics as English *wonder*, *think* and *remember*. They *must*, *cannot* and *may* select a +wh CP respectively.

Let's now postulate that there's another level of representation, Logical Form, which is the input to semantic interpretation. Let's also suppose that LF is related to S-Structure via transformations, and that wh-words in Chinese move to the relevant Spec CP between S-Structure and LF, rather than between D-Structure and S-Structure. On these assumptions, it looks like we've got everything explained.

In (8), *xiang-zhidao* 'wonder' requires that a wh-word be in its sister's specifier at LF since it selects a +wh CP, so *shenme* can only move to the embedded Spec CP, resulting in the 'statement' interpretation. In (9), by contrast, *shenme* cannot move to the embedded Spec CP, as *yiwei* 'think' does not allow its sister CP to be +wh. Therefore, at LF *shenme* moves to the main clause Spec CP, resulting in a wh-question interpretation for the sentence. Finally, in (10) *jide* 'remember' optionally may or may not take a +wh CP. If it does, then *shenme* must move to the embedded CP at LF, resulting in the statement interpretation. If it does not, then *shenme* moves to the matrix Spec CP, and a wh-question interpretation results. Since there is no morphological marking as to whether a given Spec CP is +wh or not, and there is no *overt* wh-movement to Spec CP, (10) is ambiguous.

EXERCISE 1

Given the preceding discussion, would you expect the following Chinese sentence to be a statement, a question or ambiguous?

Zhangsan zhidao Lisi mai-le shenme
Zhangsan knows Lisi bought what

An argument from Subjacency

In addition to the argument from selection which we've just seen, there is also another kind of argument supporting the idea that Chinese wh-phrases undergo covert movement. In the previous chapter we saw that *overt* movement was constrained by the Subjacency Condition:

(14) *The Subjacency Condition*
A single instance of wh-movement can cross only one bounding node, where IP and DP are bounding nodes.

If we could show that Subjacency effects seem to obtain even when it looked like no overt movement was taking place, we'd have a strong argument that there were transformational operations going on between S-Structure and our new hypothesized level of Logical Form.

To begin with, consider the following string of words:

(15) Why did John say Mary had left?

(15) is actually ambiguous between two different interpretations. On one interpretation (with falling intonation), (15) is a question about why John said something. A possible answer would be 'because he thought he saw her car drive away'. On the other interpretation (with rising intonation), (15) is a question about why Mary left. A possible answer would be 'because she needed to get back home by 7 p.m.'.

This ambiguity arises because the S-Structure in (15) can have either the D-Structure in (16a) or the D-Structure in (16b) as its source:

(16) a. John said [Mary had left] why
b. John said [Mary had left why]

(16a) is the D-Structure associated with the 'why John said something' interpretation, since in this D-Structure *why* is a sister to, and therefore modifies, the V′ *said Mary had left*. (16b) is the D-Structure associated with the 'why Mary left' interpretation, because *why* is a sister to the V′ *left*.

With this discussion in mind, consider now (17):

(17) Why do you like the man that bought the books?

Unlike the example in (15), (17) is not ambiguous. (17) cannot be a question about why the man bought the books. It can only be a question about why you like the man. A possible answer to (17) would be 'because he seems like a nice person' (an answer to why you like the man). 'Because he needed them for work' (an answer to why the man bought the books) is not a possible response to the question posed by (17).

The reason why the S-Structure in (17) is not ambiguous is because, unlike (15), there are not two possible D-Structure sources. This is because the D-Structure in which *why* is a sister to the V′ *bought the books* requires a Subjacency violation to move *why* to Spec CP:

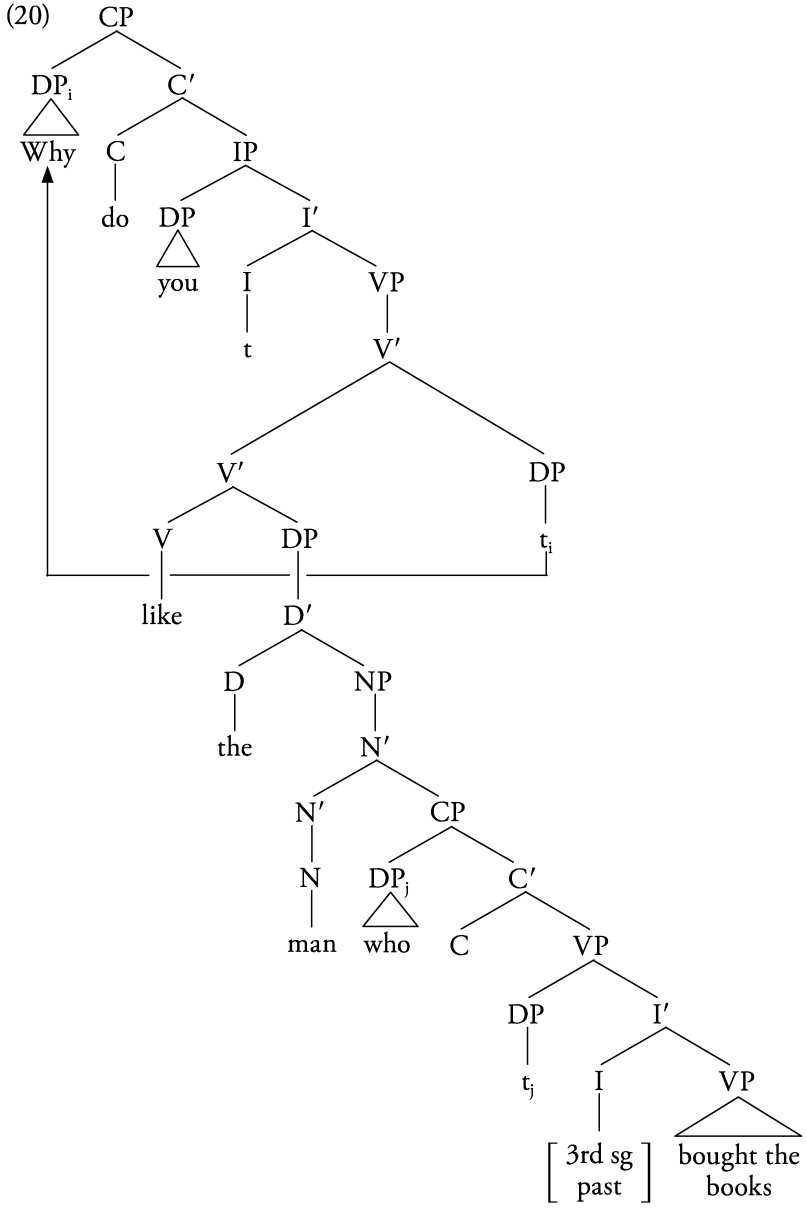
(18) *Why$_i$ do [$_{IP}$ you like [$_{DP}$ the man [$_{CP}$ t$_i'$ that bought the books t$_i$]]]?
Move no. 2 Move no. 1

The wh-movement proceeds successive-cyclically, but the movement from the embedded Spec CP to the matrix Spec CP crosses a DP and an IP bounding node. Since this D-Structure is not available as a source for (17), (17) cannot be a question about why the man bought the books.

On the other hand, if *why* is a sister to the V′ *like [the man who bought the books]*, then no Subjacency violation occurs in moving to the matrix Spec CP:

(19) Why$_i$ do you like [$_{DP}$ the man [$_{CP}$ who bought the books]] t$_i$?

Only a single bounding node (the matrix IP) is crossed by the movement in (19).



In this way, the string of words in (17) actually provides evidence for the Subjacency Condition, because the only interpretation possible is the one in which there is no Subjacency violation. A different interpretation for the D-Structure position of *why*, (18), which would involve a violation of Subjacency, isn't available.

With this in mind, consider the following Chinese sentences cited in Huang (1995):

(21) *ni zui xihuan [$_{DP}$ [$_{IP}$ weishenme mai shu] de ren]?
you most like why buy book Comp person
'Why$_i$ do you like [the man who bought the books t_i]?'

Since Chinese is a language in which wh-movement occurs at LF, rather than at S-Structure, there is not the same potential for ambiguity that we saw in (17) above. At S-Structure, the wh-word *weishenme* is in its D-Structure position and is therefore clearly seen to be modifying the VP *bought the books*. If we assume that wh-words must move at LF to Spec CP of the matrix clause, mirroring the overt wh-movement that we see in English, and we assume that Subjacency applies equally to LF movement as to S-Structure movement, then we can account naturally for the ungrammaticality of (21). In moving from its D-Structure (and S-Structure) position, it crosses two bounding nodes: DP and IP. As we expect, if we move *weishenme* to a position in which it modifies *like*, (22) becomes perfectly grammatical, with the interpretation that it's a question about why you like the man:

(22) ni weishenme zui xihuan [$_{DP}$ mai shu de ren]?
you why most like buy book Comp person
'Why$_i$ do you like [the man who bought the books] t_i?'

C-Command in Japanese

Turning now from Chinese to Japanese, we can see another case in which overt movement (in English) and LF movement (in Japanese) seem to be subject to the same requirements.

In addition to Subjacency, there is also a constraint against 'downward' movement. A moved element must always c-command its trace. To see an example, consider the S-Structure in (23):

(23) *t_i wonder [$_{CP}$ who$_i$ [$_{IP}$ John likes Bill]]

Wonder has its usual requirement that it selects a +wh CP, but in this case we've tried to satisfy that requirement by lowering *who* from the matrix subject position to the embedded Spec CP.

Given our theory so far, it's not at all clear what's wrong with (23). Subjacency is not violated, neither is the θ-criterion or the Case Filter. The chain {t_i, who$_i$} has both Case and a θ-role from the matrix subject position. However, under the assumption that antecedents must always c-command their traces, we can rule out the ungrammatical (23). This requirement is often referred to as the Proper Binding Condition.

For the moment, however, all we need to observe is that the Proper Binding Condition seems also to hold in cases where no overt movement takes place. This suggests that there *is* movement taking place, but that the movement isn't 'visible'. Consider the contrast in (24) (from Harada 1972, cited in Saito

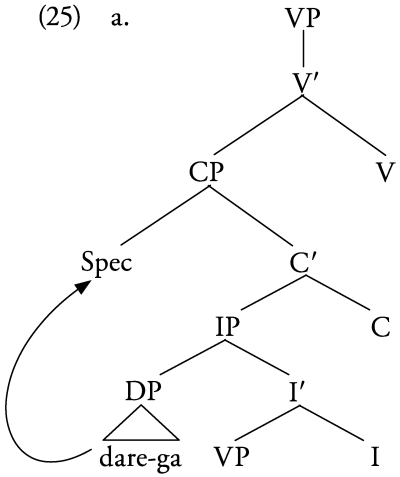
1989). *ka* is a 'scope marker', which indicates which Comp the wh-phrase is to be interpreted in:

(24) a. John-ga [$_{CP}$ [$_{IP}$ dare-ga sono hon-o katta] ka]
John-Nom who-Nom that book-Acc bought Q
siritagatteiru koto
want-to-know fact
'the fact that John wants to know who bought that book'

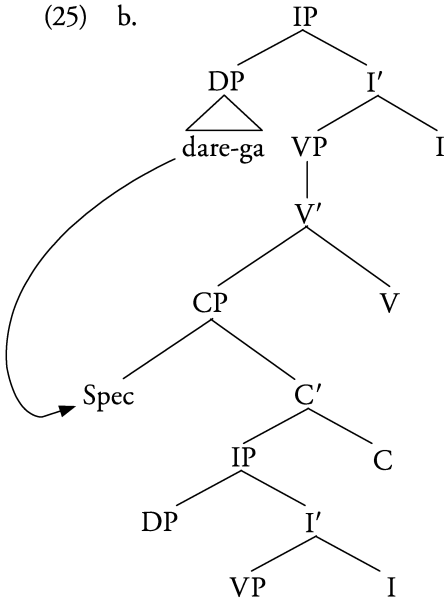
b. *dare-ga [$_{CP}$ [$_{IP}$ John-ga sono hon-o katta] ka]
who-Nom John-Nom that book-Acc bought Q
siritagatteiru koto
want-to-know fact
'the fact that who wants to know who John bought that book'

If we assume that the wh-element *dare* undergoes movement between S-Structure and LF, the Proper Binding Condition immediately gives us an account of the contrast, as can be seen from the partial structures in (25):

(25) a.



(25) b.



In (24a), the scope marker *ka* indicates that *dare* 'who' moves at LF *up* to a *higher* position from which it c-commands its trace. By contrast, in (24b) the scope marker indicates that *dare* is to move 'downward' at LF, to a non-c-commanding position. The Proper Binding Condition would seem to explain the contrast between the two, and we therefore have more evidence for the 'reality' of covert LF movement. To the extent that proposed LF movements seem to be subject to the same constraints that apply to overt movements, we have a strong argument for their existence.

LF wh-movement in English

Having discussed LF wh-movement in a couple of different languages, we're now in a better position to come back to English. Contrary to what we've been implying so far, English actually does have instances of LF wh-movement. It's just that it only happens when you have *two* wh-words. In that case, only one of the wh-words moves to Spec CP, the other stays *in situ*. Consider the question in (26), focusing in on the wh-word *what*:

(26) Who told Mary about what?

With rising intonation, (26) is interpreted as an 'echo' question, where *what* essentially means 'I didn't hear whatever words came after *about*'. (This can also be used to indicate surprise or disbelief, if *what* is heavily stressed.) However, with falling intonation, (26) is a double wh-question, where both *who* and *what* are being questioned. An expected answer would be a set consisting of pairs of people and what they told Mary about (e.g. *John told Mary about linguistics*, *Alice told Mary about history*, etc.).

EXERCISE 2

Assuming the 'double question' interpretation, give two answers each to the following questions:

(a) Who gave the book to John when?
(b) What did Bill put where?

It is the double wh-question interpretation which we're interested in here in (26). In order to account for the fact that it's *both* people *and* things which are being questioned, we want to say that *who* and *what* have *both* moved to Spec CP. Since we don't see any movement at S-Structure, however, it must be the case that *what* is moving at LF. (In order not to completely undermine Subjacency, though, we'll have to say that these multiply-filled Spec CPs are allowable only at LF, not at S-Structure, otherwise there would be no Subjacency violations ever.)

We can see even more clearly how moving to the main clause Spec CP means that a wh-word is part of the question if we take a look at a famous example, first discussed by Baker (1970). Consider (27):

(27) Who_i t_i is wondering where_j we bought what t_j?

Notice that (27) is ambiguous. Under one interpretation, it is a single question asking who the person is who is wondering something, namely where we bought what. An answer would be: 'Mary is. Mary's wondering where we

bought what.' However, (27) can also be interpreted as a double question, asking for a list of people paired with objects. An answer to *that* question would be: 'Mary is wondering where we bought the car. John is wondering where we bought the violin', and so on.

If we assume that *what* is undergoing wh-movement at LF, we get a nice explanation for this ambiguity. There are two possible Spec CPs that *what* could be moving to, and those correspond to the two possible interpretations. First, *what* could be moving at LF to the embedded Spec CP:

(28) $[_{CP}$ Who$_i$ $[_{IP}$ t_i is wondering $[_{CP}$. what$_k$ where$_j$ $[_{IP}$ we bought t_k $t_j]]]]$?

If *what* moves at LF to the embedded CP, then the only wh-word in the matrix Spec CP is *who*. We therefore get a single question interpretation. We want to know who the person is who is wondering something. An answer is 'Mary is wondering where we bought what'.

The other possibility is that *what* is moving at LF to the matrix Spec CP:

(29) $[_{CP}$ what$_k$ who$_i$ $[_{IP}$ t_i is wondering $[_{CP}$ where$_j$ $[_{IP}$ we bought t_k $t_j]]]]$.

If *what* moves at LF to the matrix Spec CP, then it's a question involving *who* and *what*, and we get the multiple wh interpretation. The question we're asking has *two* elements to the answer. We want to know about pairs of (1) people and (2) things that they're wondering where we bought. Thus an expected answer is 'Mary is wondering where we bought the car. John is wondering where we bought the violin', and so on. (We'll address the potential Subjacency violation in the next section.)

In a language like Polish, we can actually see multiple wh-movement at S-Structure. So, in Polish, all wh-movement takes place overtly between D-Structure and S-Structure. (Thus the ban on multiply-filled Spec CPs at S-Structure which holds in English does not hold in Polish.)[1] Consider the following example from Lasnik and Saito (1984):

(30) a. Zastanawiam sie *kto co* przyniesie.
Wonder-lsg who what will bring
'I wonder who will bring what.'
b. *Zastanawiam sie *kto* przyniesie *co*.
Wonder-lsg who will bring what

In (30) both wh-phrases must move to Spec CP at S-Structure. This is unlike the English multiple questions which we've seen, where one wh-word moves at S-Structure and the others wait until LF.

This Polish data seems to further confirm the general theory. What we see is that all multiple wh-questions have the same structure at LF, but there are variations as to the timing of wh-movement. All wh-movement happens overtly in Polish. In English, one wh-word moves overtly and any others move covertly, and in Chinese or Japanese all wh-words move covertly.

English LF wh-movement and Subjacency

So far, we've seen some arguments that English does in fact have LF wh-movement in the case of multiple wh-questions. Not only does it seem to account for the multiple-question interpretation that multiple wh-sentences can have, but also we've seen a language (Polish) where it appears as though what happens at LF in English happens at S-Structure in this language. What we now want to do, paralleling our discussion of wh-movement in Chinese and Japanese, is to look at parallels between overt movement and LF movement. To the extent that our postulated LF movements pattern with similar overt movements, then we'll have some good evidence that movement that we don't see really is taking place.

In particular, we saw earlier that we could use Subjacency as a diagnostic for LF wh-movement in Chinese, and it turns out that we can use it again in English. As usual, the logic of the argument is that wherever we're postulating covert LF movement, we should check to see if this 'movement' obeys constraints that we know hold of overt movement.

In that context, consider (31), from Huang (1995):

(31) *Who likes the man [who fixed the car how]?

We might expect that (31) should be a perfectly well-formed double question, asking for pairs of people and ways in which the men that they liked fixed the car. However, (31) is ungrammatical. We can explain this in the same way as we did in Section 7.1. We're assuming that *how* must undergo movement at LF to the matrix Spec CP, but we have again a classic 'complex DP' Subjacency configuration. *How* can't move to the matrix Spec CP without crossing an IP and a DP node, and the ungrammaticality of (31) now follows.

However, this argument must remain slightly tentative. As the eagle-eyed among you may have noticed, the claim that Subjacency constrains LF movement as well as overt movement causes problems for (27), which we looked at in the previous section:

(27) Who is wondering where we bought what?

We said a moment ago that (27) can be interpreted as a multiple question, asking for pairs of people and things that they're wondering where we bought. Therefore *what* must be able to move to the matrix Spec CP. But, given what we've just said, that should create a Subjacency violation just like (31), since again Spec CP is blocked in precisely the same way. Movement from the position of the direct object of *buy* to Spec CP should cross two IP nodes.

Unfortunately, we'll have to leave this as an outstanding problem for the time being. In Chapter 8, when we look at constraints on movement in more detail, we'll have a way of dealing with this contrast. It'll turn out that adjuncts are subject to slightly different principles than arguments when extracted from certain kinds of Subjacency configurations. It is also possible that the ideas that we'll discuss in the 'Open Issue' section at the end of this chapter could be relevant.

7.2 Quantifier raising: another LF transformation

Now that you're getting used to the idea of covert LF movement, I want to step back for a minute and talk a little more about Logical Form as a level of representation. As we mentioned briefly at the start of the chapter, it is assumed that the language faculty interfaces with two other systems: the articulatory-perceptual system (phonetics/phonology) and the conceptual-intensional system (semantics). This is intended to encode the traditional intuition that what's at the heart of language are sound/meaning pairs. LF is the level which provides the interface between the syntax and the conceptual-intensional system. What's represented at LF are those aspects of semantic interpretation which are determined by the syntax. In this section, we're going to look at 'quantificational' DPs and argue that they too undergo LF movement in English.

Introduction to quantifiers

Quantifiers are words in a language like *every*, *some*, *all* and *many*. The effect that they have is to create DPs which do not pick out a specific entity in the world. Rather, what these quantificational DPs do is delimit a range of possible entities and then focus in on some (quantified) part of that range.[2] This contrast can be easily seen if you compare a good, old-fashioned R-expression to a quantificational expression:

(32) a. *John* likes The Beatles.
b. *Every sane person* likes The Beatles.

In (32a), a particular person in the universe of discourse is picked out, namely *John*, and it is claimed that it is true of him that he likes The Beatles. On the other hand, when I assert (32b), I'm not referring to a specific person and making a claim about him or her. Rather I'm picking out a group or range of people whoever they might be (in this case the class of people who are sane) and saying that it is true that every person who falls into that group likes The Beatles. As first argued for by the philosopher Bertrand Russell (1905), if we were to express the claim made in (32b) using logical notation, we would express it using the following formula:

(33) $(\forall x: x \text{ is a sane person})\ (x \text{ likes The Beatles})$

(33) is read 'for every x, x a sane person, x likes The Beatles'.

Notice that quantifiers seem to have a lot in common with wh-words and phrases. Wh-words/phrases are also DPs which don't pick out specific entities, but rather draw our attention to groups or ranges of entities. And they have the same logical structure in (33). Consider the S-Structure in (34):

(34) [Which eighties pop group]$_i$ does John admire t_i?

Just as with quantifiers, there's no specific eighties pop group which is being referred to by the wh-phrase. Rather, what the wh-word does is identify a range of things (eighties pop groups) and asks which one within that range is the one such that John admires it:

(35) (which x: x is an eighties pop group) (John admires x)

In fact, it's commonly assumed that wh-words consist of an interrogative feature together with a quantifier like *some*.

We'll use the term *operator-variable structures* to refer to structures like (33) and (35). The *operator* portion expresses the range and what subpart of it we're dealing with (whether it's every member of the range, most members of the range, some member(s) of the range, etc.). The *variable* portion tells us where in the corresponding sentence we would plug in the individual member of the range to see if the sentence is true or false. So, for example, assuming that the range of sane people is {Maria, Erich, John}, when I assert something like *every sane person likes the Beatles*, I'm saying that if you plug Maria, Erich and John in place of the x, the sentences *Maria likes the Beatles*, *Erich likes the Beatles* and *John likes the Beatles* will all be true. In (33) and (35), the operators are the quantifier and the wh-phrase respectively. The variable in both cases is the trace.

However, quantifiers and wh-words differ in that, in the case of wh-words, their operator-variable structure is represented in the syntax, at least in English. The wh-word/phrase, which corresponds to the operator part, moves to Spec CP. This operation leaves behind a trace, which corresponds to the variable part. As we saw in Section 7.1, languages which don't have overt wh-movement like Chinese and Japanese still have covert wh-movement at LF . Therefore, at the level of LF, the operator-variable structure of wh-questions is always represented.

This is not too surprising, though, when you think about it. The level of LF is the interface with the conceptual-intensional systems, which contain things like semantics and real-world knowledge. Presumably the operator-variable nature of questions is the same in all languages.

But if quantifiers also have operator-variable structure, then we might expect that this would also be represented at LF. Unlike wh-questions, though, we don't have any overt movement of the quantifier in English. It remains in its D-Structure position (or S-Structure position if it's undergone raising or passive). However, the operator-variable structure could still be represented at LF if quantifiers, in fact, undergo a process of quantifier raising (QR). This gives us a type of conceptual argument for saying that there should be a transformation in the LF component which targets quantifiers.

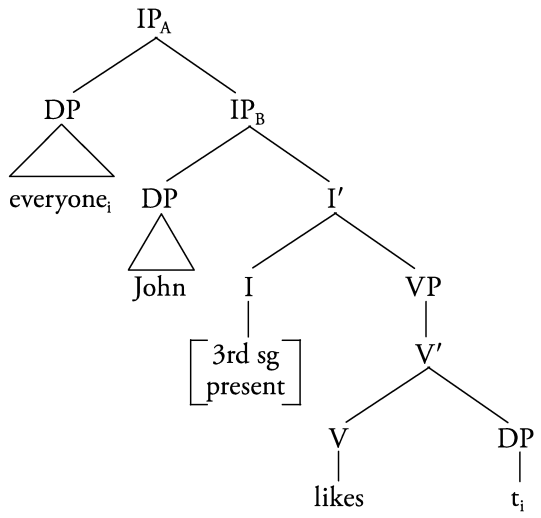
The landing site of QR and 'scope'

If quantifiers undergo movement at LF, we need to know what the landing site for the movement is. What is generally assumed is that QR 'adjoins' the

quantifier to IP, so that the S-Structure in (36) has the LF in (37):

(36) $[_{CP}\ [_{IP}$ John likes everyone]].

(37)



The process of adjunction targets the maximal projection and 'splits' it into two *segments*. IP_A and IP_B in (37) are *not* two maximal projections. They together count as only one. We therefore have no violation of X′-Theory, as there is still one head and one maximal projection. (The process of *Topicalization*, which we mentioned back in Chapter 2, is a transformation in the overt syntax which also adjoins an element to IP.)

A traditional idea, which will be particularly important in the next section, is that quantifiers have *scope*. In other words, they have a domain in which they have effects. Within our system, the scope of a quantifier is its c-command domain at LF. Therefore, the scope of the quantifier *everyone* in (36) is everything dominated by IP_B. If quantifier A asymmetrically c-commands quantifier B at LF (in other words, quantifier A c-commands quantifier B, but not vice versa), then quantifier A is said to have *wide* scope relative to quantifier B. Quantifier B on the other hand is said to have *narrow scope* relative to quantifier A.

One question which we'll have to have a little more to say about in just a minute is why QR targets IP, rather than VP or CP, for example. For the moment, there is an answer that comes from traditional logic and semantics. Basically, there's a long history of work in these fields which has argued that quantifiers have scope over propositions. Within the syntactic system that we're developing here, it's IP that corresponds to 'proposition'. So, from that point of view, the representation in (37) encodes this idea. The quantifier *everyone* in (37) c-commands IP_B and therefore has scope over the proposition that it represents.

EXERCISE 3

Draw the tree for the LF representation of

(a) Everyone saw Bill
(b) John gave the book to every student

Scope ambiguities and QR

We've seen that quantificational DPs seem to undergo a process of quantifier raising which adjoins them to IP. This process happens in the LF component in many languages, including English, but can in fact be seen overtly in languages like Polish and Hungarian.

However, we've only seen cases involving one quantifier so far. When there are two (or more) quantifiers in a clause, they can interact in interesting ways to create what's known as *scope ambiguity*. By way of illustration, consider (38):

(38) Every politician avoided some journalist.

(For some speakers, (38) is more natural with *some* replaced by the indefinite article *a*, which can also function as a quantifier.) Notice that there are two quite different situations under which (38) could be true. First, it could be the case that for every politician, there was some journalist or other that he or she avoided (a different journalist for every politician). However, (38) is also true if there's one specific journalist and every politician avoided him or her. To make this a bit clearer, look at the following diagrams which each illustrate one of the readings:

(39a) Politician no. 1 → Journalist no. 1
Politician no. 2 → Journalist no. 2
Politician no. 3 → Journalist no. 3
Politician no. 4 → Journalist no. 4
Politician no. 5 → Journalist no. 5
and so on

(39b) Politician no. 1 ↘
Politician no. 2 ↘
Politician no. 3 → Journalist no. 1
Politician no. 4 ↗
Politician no. 5 ↗
and so on

Interpretation (39a), in which there are multiple journalists, is referred to as the reading in which *every* has wide scope relative to *some* (for every politician x, there is some journalist y, such that x avoided y). Interpretation (39b), in which there is a single journalist, is referred to as the reading in which *some* has wide scope relative to *every* (there is some journalist y, such that for every politician x, x avoided y).

Taking the scope of a quantifier to be its c-command domain at LF, and assuming that we can apply QR to the quantified expressions in either order, we can explain the interpretational ambiguity of sentences like (38).

Let's start by looking at the S-Structure for (38):

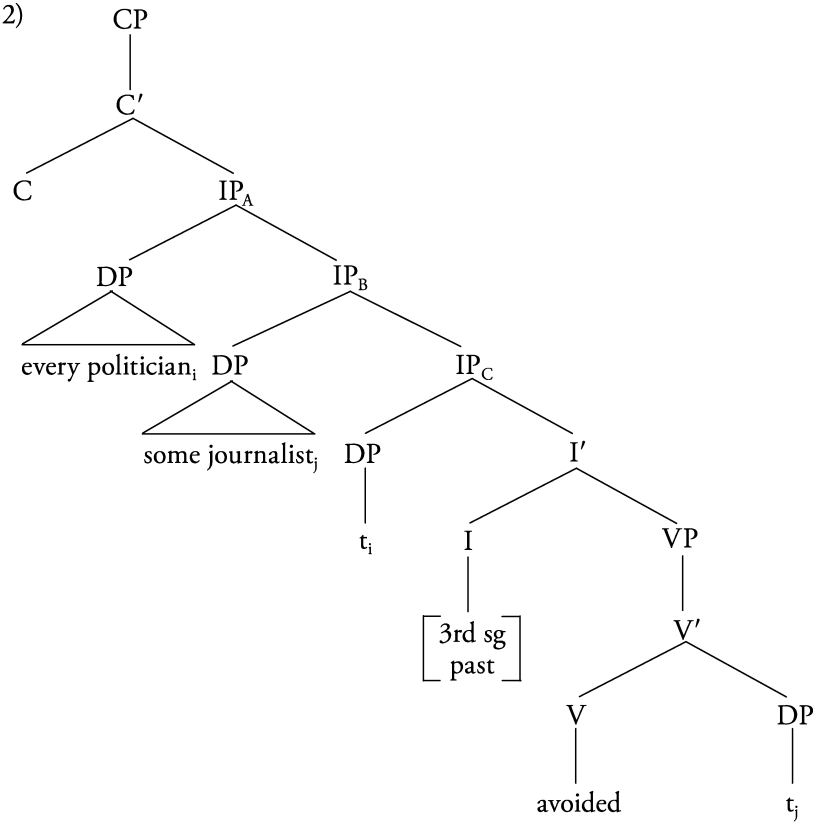
(40) $[_{CP}\ [_{IP}$ every politician$_i$ $[_{VP}$ avoided some journalist$_j$]]]]

We know that both quantified phrases *every politician* and *some journalist* must undergo QR at LF because one of the main functions of LFs is to represent quantifier scope. If we apply QR to *some journalist* first, we get (41):

(41) $[_{CP}\ [_{IPA}$ some journalist$_j$ $[_{IPB}$ every politician$_i$ $[_{VP}$ avoided t_j]]]]

We then QR *every politician*, giving us (42) (which I've done as a tree to make it easier to see):

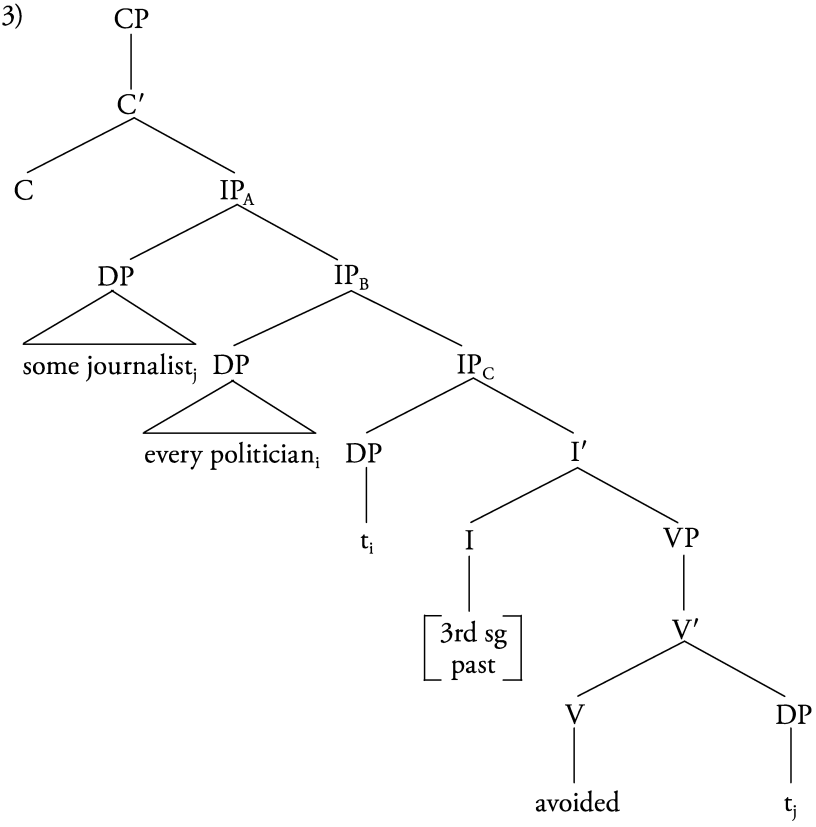
(42)



In (42), since the DP *every politician* has the quantifier *some journalist* within its scope at LF, we can encode the fact that *every politician* is interpreted as having wider scope than *some journalist*, resulting in interpretation (39a).

If, on the other hand, QR applies to *every politician* first and then to *some journalist*, the LF representation in (43) will result. This corresponds to the interpretation in (39b), in which there is some specific journalist such that every politician avoided him or her.

(43)



EXERCISE 4

Discuss the scope ambiguities present in the following sentences:

(a) Some student leaves home every day
(b) John showed three books to every customer

Quantifier raising and constraints on movement

Now that we've talked a little bit about quantifiers and how QR works, and given some conceptual arguments for the LF movement of quantifiers, it's time to turn to the arguments based on constraints on movement. As with LF wh-movement in Section 7.1, if we can show that QR obeys the same constraints that overt A′-movement shows, we'll have a strong argument that covert A′-movement is in fact taking place.

Quantifier raising and Subjacency

As with LF wh-movement in the previous section, our first instinct should be to check Subjacency. If QR looks like it's subject to Subjacency, we'll have a good argument that movement is really taking place, even if we don't 'see' it. Unfortunately, the facts about Subjacency and QR are ultimately inconclusive for independent reasons. I want to discuss the issue, though, as an example of how you can get a seemingly 'negative' result which *isn't* a problem.

The core of the problem is that Subjacency constrains long-distance movement, but QR can't move quantifiers in a long-distance fashion. QR can adjoin them to the closest IP which dominates them, but can go no higher. Consider (44):

(44) $[_{IPA}$ some student thinks that $[_{IPB}$ John saw every band]]

If *some student* adjoins to IP_A and *every band* adjoins to IP_B, we have the LF in (45):

(45) $[_{IPA}$ some student$_i$ $[_{IPA}$ t_i thinks that $[_{IPB}$ every band$_j$ $[_{IPB}$ John saw $t_j]]]]$

Some student c-commands *every band*. Therefore *some student* has wide scope relative to *every band* and the interpretation is that there is a single student who thinks that John saw every band. If, on the other hand, *every band* could adjoin to the higher IP, IP_A, we should be able to generate the LF in (46):

(46) $[_{IPA}$ every band$_j$ $[_{IPA}$ some student$_i$ $[_{IPA}$ t_i thinks that $[_{IPB}$ John saw $t_j]]]]$

But this LF, in which *every band* has wide scope relative to *some student*, is interpreted as meaning that there is a multiplicity of students and bands. For every band, there is some student who thinks that John saw them: Mary thinks John saw Guided by Voices, Bill thinks that John saw Baby Ray, and so on.

But (44) is not ambiguous between these two readings. *Some* can have wide scope relative to *every*, but not vice versa. There can only be a single student who has some thought, namely that John saw every band. There can't be a multiplicity of students who each think that John saw a different band.

Notice that we can't use the impossibility of the reading represented by (46) as an argument for QR. There's no Subjacency problem there. IP_B is crossed but not IP_A since all the 'segments' of IP_A don't dominate the landing site. It just seems to be the case that, for whatever reason, quantifiers must adjoin to the closest IP and go no further. Therefore, we can't use Subjacency to argue for LF movement of quantifiers. It simply turns out not to be relevant.

Inverse linking and c-command

Although we have problems with Subjacency, we can now turn to a phenomenon called 'inverse linking', originally discussed by May (1977) and one of the most persuasive arguments I know for QR and for the idea of LF (covert) transformations as a whole. It's essentially the Proper Binding Condition from

Section 7.1. The traces that would be left behind if you assumed covert QR was taking place look like they need to be c-commanded by the QRed element, just like the traces of overt movement. This suggests that movement really is taking place.

Earlier we saw how two quantifiers can create ambiguity, depending on which one undergoes QR first (and thus has narrow scope). With that discussion in mind, consider the following sentence, which also has two quantifiers:

(47) Every man in some Italian city likes opera.

BEFORE CONTINUING

What scope relations are possible in (47)? Is (47) ambiguous?

Taking the discussion in the previous section, we might expect that (47) should have two possible readings, just like the other cases we saw with two quantifiers, since QR can apply to either one of the quantifiers first. (There's only one IP, and so we don't have the problems that we saw with example (44) in the previous section.) If *every* has wide scope with respect to *some*, then (47) should mean that for every man who loves opera, there is some Italian city in which he resides. Alternatively, if *some* has wider scope, then (47) should mean that there is some single Italian city such that every man in that city likes opera. These two readings are represented graphically in (48) and (49).

(48) Man no. 1 → City no. 1
Man no. 2 → City no. 2
Man no. 3 → City no. 3
Man no. 4 → City no. 4
Man no. 5 → City no. 5
and so on

every > some

(49) Man no. 1 ↘
Man no. 2 ↘
Man no. 3 → City no. 1
Man no. 4 ↗
Man no. 5 ↗
and so on

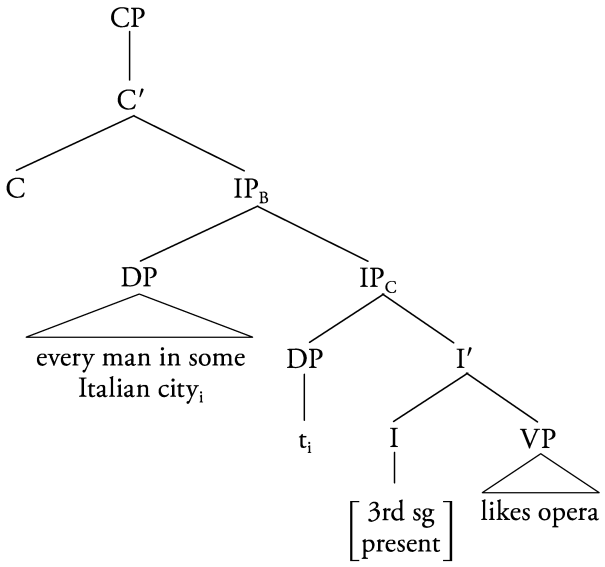
some > every

Interpreting sentences with multiple quantifiers is a notoriously difficult thing to do, but if you're having particular difficulty getting the reading in (48), in which *every* has wider scope than *some*, there's a reason. Unlike the cases in (38) that we saw in the previous section, (47) has two quantifiers but is not ambiguous! The only reading possible for (47) is the one in which *some* has wider scope than *every*. (47) can only mean that there is some specific Italian city and every man in that city likes opera. This is really quite surprising. What could it be that allows two quantifiers to interact to create ambiguity in (38) but not in (47), when one quantified expression is contained within another?

May suggested that an answer could be found by looking at what the two readings would look like if we assumed that QR was taking place in the LF

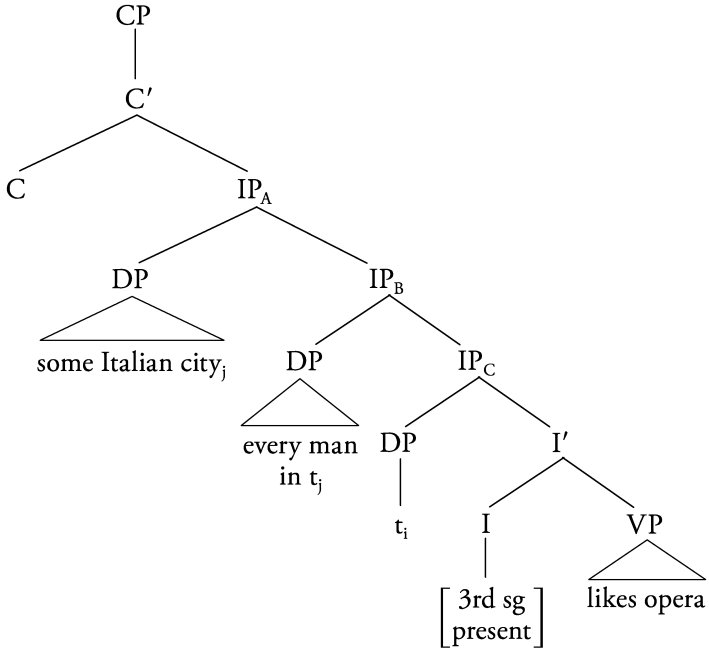
component. Let's first take a look at the reading in which *some* has wide scope – the permitted reading. Since the last quantifier to undergo QR is the one which has the widest scope, we want to begin by targeting the DP which is headed by *every*. That DP is the DP *every man in some Italian city*. Applying QR to that DP gives us the representation in (49):

(49)



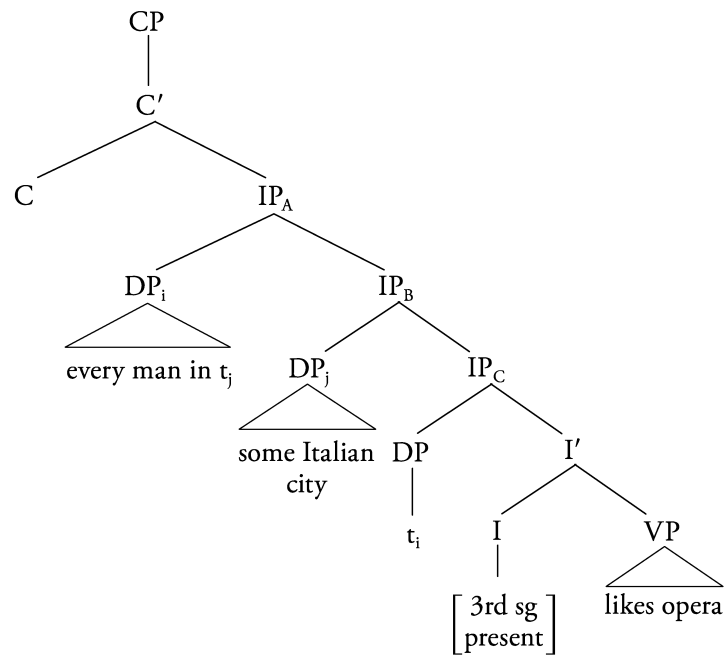
We then QR the DP *some Italian city* resulting in the LF representation in (50):

(50)



However, let's take a look at what happens if we try to derive the impossible reading – the one in which *every* has wider scope than *some*. First we QR *some Italian city* from within the subject DP and then the DP *every man in t*, resulting in the representation in (51):

(51)



Notice now what's happened in the course of the derivation. When we QRed *every man in t*, it took the trace of *some Italian city* with it. This has the result that the DP_j *some Italian city* now no longer c-commands its trace t_j. The first branching node dominating DP_j is IP_B, and IP_B doesn't dominate DP_i.

But we saw in Section 7.1 that covert LF movement is constrained by the Proper Binding Condition; that is, traces of LF movement, just like traces of overt movement, must be c-commanded by their antecedents. That was one of the arguments we gave for saying that an LF wh-movement was taking place in Japanese. And we can use it again here. Even though there are two quantifiers in (47), there aren't two possibilities for applying QR. If we QR *every man in some Italian city* followed by *some Italian city*, nothing goes wrong. Therefore, as predicted, the wide scope reading of *every* is available in (47). On the other hand, to QR *some Italian city* followed by *every man in t*, resulting in (51), would violate a general principle which we know to hold of transformations – The Proper Binding Condition. We therefore can't apply QR in that order, and

we correctly predict that (47) cannot be interpreted with *every* having wider scope than *some*.

QR and weak crossover

A final area that we'll look at is the behaviour of quantifiers in what are called 'weak crossover' configurations. It has long been noted that pronouns cannot have a 'bound-variable' interpretation after certain kinds of wh-movement have taken place. The interesting thing for our purposes turns out to be that quantifiers also can prohibit pronouns from taking on this bound-variable interpretation, even though no movement *appears* to be taking place. However, if there really is movement, in the form of QR at LF, then everything is explained.

But first, what exactly does it mean for a pronoun to have a 'bound-variable' interpretation? How is this different from a pronoun's normal interpretation? The easiest thing to do is take a sentence where this bound-variable interpretation is possible, but not necessary, and compare the two interpretations. Consider a sentence like (52):

(52) $\text{Who}_i\ t_i$ thinks that $\text{he}_{i/j}$ is smart?

The most natural interpretation for the pronoun in (52) is one in which it has index 'i', referring to the same person as the wh-word. This interpretation, where the pronoun refers back to a wh-phrase, is called the *bound-variable* interpretation. The pronoun is in effect acting as a second variable in addition to the wh-trace, both of them being bound by the operator *who*. Under the bound variable interpretation, (52) means something like 'who is the person that thinks that he or she is smart?' or, slightly awkwardly, 'who thinks himself smart?'. A possible answer would be 'John thinks that he is smart', meaning 'John thinks that John is smart'. The pronoun, unlike normal, does not refer to a specific person, but instead varies over the same individuals that the wh-phrase does.

In fact, the use of 'inclusive' or 'non-sexist' language can sometimes aid in bringing out the bound-variable interpretation. Masculine pronouns such as *he* or *his* are sometimes replaced by both masculine and feminine pronouns *he/she* or *his/her* when the actual gender of the person is not known or irrelevant. For example (53):

(53) As for the new employee, his/her duties will include...

In the context of a wh-word, this use of non-sexist language can force the bound variable interpretation. In rewriting (52) as (54), it makes it clearer that the bound-variable interpretation is intended:[3]

(54) $\text{Who}_i\ t_i$ thinks that he/she_i is smart?

By contrast, if the pronoun has index 'j', and is therefore not coindexed with the wh-word, then the pronoun is not acting as a bound variable. It instead

refers to some unspecified other person. For example, say we've been talking about our friend John, I could produce (55):

(55) Who$_i$ t$_i$ thinks that he$_j$ is smart?

In this case, the pronoun *he* is not acting as a bound variable and simply refers to *John*. A possible answer to the question in (55) might be '*Bob* thinks that he [i.e. John] is smart'.

Having explained what bound variables are, we now need to explain what 'weak crossover' is. So-called 'weak crossover' configurations are illustrated by sentences like (56):

(56) *Who$_i$ does [his$_i$ mother] like t$_i$?

The configuration is a 'crossover' configuration because the wh-word has 'crossed over' the pronoun in moving from its D-Structure position to the matrix Spec CP. It is a 'weak' crossover configuration because the pronoun does not c-command the wh-trace at S-Structure. (There are also 'strong crossover' configurations, in which the pronoun *does* c-command the wh-trace.)

The thing to notice about weak crossover configurations (and strong ones too, actually) is that the pronoun in (56) *cannot* act as a bound variable, coindexed with the wh-word *who*. If it did, it would mean 'who is the person x, such that x's mother likes x?'. In other words, it would be the same question as (57):

(57) [Whose mother]$_i$ t$_i$ likes him$_i$?

But (56) can't mean what (57) means. In (56) *his* must have a different index from *who*, meaning we've been talking about our friend John, and I'm asking 'who does his [i.e. John's] mother like?'.

Compare the question in (56) with the following question:

(58) Who$_i$ t$_i$ likes his$_i$ mother?

Here, the bound-variable interpretation for the pronoun *is* possible. (58) can be put as 'who likes his own mother?'. This is because in (58) the wh-word does not cross over the pronoun in moving to Spec CP, and so the bound-variable interpretation for the pronoun is perfectly possible.

If we use our 'inclusive language' test, we can see the difference between (56) and (58):

(59) a. ??Who does his/her mother like?
b. Who likes his/her mother?

The contrast between (59a) and (60) is clear. The strangeness of (59a) comes from the fact that (59a) represents a weak crossover configuration, in which the

bound-variable interpretation is prohibited. However, by using *his/her* we're trying to *force* a bound-variable interpretation. In (59b), on the other hand, there is not a potential crossover configuration. There is therefore no such strangeness when you try to force a bound-variable interpretation.

Now let's take what we've seen about crossover configurations with respect to wh-movement and see what happens when we replace the wh-word with a quantifier. Consider first (60):

(60) Everyone$_i$ loves his$_i$ mother.

Here there's no problem with a bound-variable interpretation for the pronoun. (60) can perfectly well mean 'for every x, x loves x's mother', and, in fact, that's the most natural interpretation of the sentence (although, of course, as always a non-bound-variable interpretation is possible as well).

Compare (60) with (61), though, and you see something interesting:

(61) *His$_i$ mother loves everyone$_i$.

In (61), the pronoun seems unable to act as a bound variable. (61) can't mean something like 'for every x, x's mother loves x', making (61) parallel with (60). Why should this be? It can't be that there's a general prohibition on pronouns acting as bound variables with things to their left, as (62) shows (from Reinhart 1983: 129):

(62) a. Near his$_i$ child's crib nobody$_i$ would keep matches.
b. For his$_i$ birthday, each of the employees$_i$ got a Mercedes.

However, if we assume that *everyone* undergoes QR, then in (61) we'll have a weak crossover configuration at LF:

(63) [$_{IP}$ everyone$_i$ [$_{IP}$ his$_i$ mother loves t$_i$]]

Therefore, under the assumption that *everyone* undergoes QR, we can explain why the bound-variable interpretation is impossible in (61). Again, the 'inclusive language' test also gets the same results as it did with wh-movement:

(64) a. Everyone loves his/her mother.
b. ??His/her mother loves everyone.

Although we haven't really specified what it is about weak crossover configurations that interferes with a bound-variable interpretation (unlike strong crossover, which we'll see in the next chapter is due to Principle C of the Binding Theory), what we do see is quantifiers patterning with wh-words. It seems not only that quantifiers semantically function as operators, but also that the operator-variable structure of quantifiers is ultimately syntactically represented at LF.

Scope ambiguities and the landing site of quantifier raising

Just to close this section, I now want to come back briefly, as promised, to the question of the landing site of QR, and argue that one can use the interpretation of sentences with multiple quantifiers to provide a sort of 'process of elimination' argument for IP as the landing site.

To start, consider the fact that, in a simple one-clause sentence, two quantifiers can interact from basically any position:

(65) Every politician avoided some journalist (subject, direct object)

(66) Every student showed John to some teacher (subject, indirect object)

(67) John showed every student to some teacher (object, indirect object)

In (65), as we've discussed above, the journalists can vary with the politicians (every > some) or there can be one particular journalist that every politician avoided (some > every). In (66) and (67), the teachers can vary (every > some) or not (some > every). That means that wherever these quantifiers are moving to at LF, it has to be possible for either quantifier to have scope over the other in (65)–(67).

This immediately rules out anything below IP as the landing site for QR. If, for example, the quantifier in object position were adjoined to VP, or the quantifier in indirect object position could adjoin to PP, there shouldn't be any ambiguity with a quantifier in subject position. Remember from the previous section that the Proper Binding Condition constrains QR, so QR can't move anything downward. As a result, the quantifier in subject position would have to at least move to IP, and from there it would always asymmetrically c-command a quantifier adjoined to VP or PP, and it would therefore always have wider scope.

BEFORE CONTINUING

We've just ruled out anything *lower* than IP as a possible landing site. How would you rule out something *higher*, like CP?

In order to rule out CP as a possible landing site for QR, we need to take a look at clauses which don't contain CP, but instead are 'bare' IPs. If a quantifier can 'escape' from that clause to interact with a quantifier higher up, then we have an argument that the landing site for QR should be CP. If, on the other hand, a quantifier in 'bare' IPs is trapped within that IP, then we have an argument that IP is where quantifiers adjoin to.

We've seen two instances so far of cases where a V takes an IP complement rather than a CP complement. In Chapter 6, we saw that raising verbs like *seem*

take an IP complement rather than a CP complement, and back in Chapter 4 we saw that the same is true of Exceptional Case Marking (ECM) verbs. So let's take a look at those contexts.

One thing we need to make sure about is that there's no possibility of 'cross-contamination', so when we pick examples to check, we want to make sure that we pick quantifiers which are *definitely* in one clause or the other. If we choose an example where the quantifier is in the lower clause at D-structure, but raises to the higher clause at some point, say, then our experiment is going to be messed up. To that end, consider the raising example in (68):

(68) John$_i$ seems to some student [$_{IP}$ t_i to have read every book].

Because the subject of *seems* was the subject of *to have read* at D-structure, we definitely don't want to put a quantifier in that position. However, the indirect object PP in the matrix clause was never in the lower clause at any point. Similarly, we can be pretty confident that the direct object of the embedded clause doesn't somehow sneak into the matrix clause at LF in a way that would mess up our experiment.

Interpretively, (68) seems to be unambiguous, and allows only the wide scope reading for *some student*. There is a single student, say Mary, to whom it seems that John has read the totality of the books. This suggests that *every book* is 'trapped' within the embedded clause, which would be the case if it were moved to the nearest IP by QR. That IP is [t_i to have read every book]. Then, *some student*, who would be moved by QR to adjoin to the matrix IP, would asymmetrically c-command *every book* at LF.

It looks like similar results obtain when we look at ECM constructions.

(69) Some student expected John to read every book.

Again, interpretively, it seems as though (69) can only be interpreted as meaning that there is a particular student who expected that John would read the totality of the books. There cannot be pairs of students and books. This would be accounted for if *every book* was forced to adjoin to the IP [John to read every book]. *Some student*, adjoined to the matrix IP, would always have wider scope.

Notice, though just in passing that when we look at ECM clauses, the subject of the main clause and the subject of the ECM clause actually do seem to interact, which is somewhat unexpected:

(70) Some student expected every professor to retire next year.

(70) can mean either that there is a specific student who expects that every professor will retire next year (some > every) or that there is a multiplicity of students, each of whom expects that a different professor will retire next year (every > some). This suggests that the ECM subject at LF must be slipping into the higher clause somehow, and from there it's QRed to the matrix IP, so that it can interact with a quantifier in the subject position of the matrix clause. We'll have to leave this as a mystery for the moment, but in Chapter 9, when we have

a look at accusative Case assignment again, we'll see an explanation for why it might be happening.

7.3 LF A-movement: expletive replacement

So far, we've seen various examples of LF A′-movement. The landing sites of the movements have all been places like Spec CP or adjoined to IP. These are A′-positions – not positions to which a grammatical function like subject or object is assigned. However, there's nothing in principle that says that LF movement is restricted to A′-movement, and in this section we'll examine expletive replacement, which is an instance of LF movement to an A-position – specifically, subject position.

One of the distinctive features of expletive *there* constructions in English is that the verb shows subject-verb agreement with the 'associate' of the expletive (as it's called) rather than the expletive itself:

(71) a. *There is two people in the garden.
b. There are two people in the garden.

This is a fairly straightforward challenge to what we've seen so far about agreement.[4] *There* is occupying the Spec IP position, and, if agreement always takes place in a specifier-head relationship at S-Structure, it's unclear why agreement with the associate happens. Clearly *two people* is not in the requisite specifier-head relationship with I°. How can we handle this situation?

Well, the title of this section probably gives the answer away. It looks as though we can kill two birds with one stone if we assume that *two people* moves to subject position and replaces the expletive. Since this movement is obviously something that you don't 'see', then it must be happening between S-Structure and LF. We can then say that although subject-verb agreement is present at S-Structure, it actually represents the LF configuration. It is whatever's in the specifier of IP *at LF* that the verb is actually agreeing with. If we examine the LF representation of (71b) under this expletive replacement assumption, we have:

(72) Two people$_i$ are [$_{PP}$ t_i in the garden].

All we've got to worry about is the chain {two people, t}; and the chain has Case (from the subject position) and a θ-role (from its D-Structure position). Again, if subject-verb agreement is determined by the LF representation, there will be no problems there either.

Expletive replacement and constraints on movement

As always, the next question is whether or not we can dig up any independent evidence for the expletive replacement hypothesis just presented, apart from the fact that it seems to provide a nice solution to several potentially thorny problems. Well, just as we saw with the other instances of LF movement that

we've discussed, expletive replacement also seems to share a lot of properties with overt movement. To begin with, notice that sentences in which you have expletive–associate relationships alternate with configurations in which overt A-movement takes place:

(73) a. There have been six records broken. [passive]
b. Six records$_i$ have been broken t_i.

(74) a. There seems to be a man in the garden. [raising]
b. A man$_i$ seems t_i to be in the garden.

However, we can make an even stronger argument by checking configurations in which overt A-movement *isn't* possible and seeing whether an expletive–associate relationship is possible. For example, as (74b) illustrates, raising is possible when the embedded clause is infinitival. However, raising is not possible when the embedded clause is tensed:

(75) *A man$_i$ seems that t_i is in the garden.

We'll see exactly what's wrong with (75) in the next chapter. However, the thing to notice about the configuration in (75) is that an expletive–associate relationship is also impossible.

(76) *There seems that a man is in the garden.

Chomsky (1995: 156–7) points out a number of other configurations in which both overt movement and expletive–associate relationships are disallowed.

(77) a. *A man$_i$ seems that John saw t_i.
b. *There seems that John saw a man.

(78) a. *A man$_i$ was thought that pictures of t_i were on sale.
b. *There were thought that pictures of a man were on sale.

We'll see in more detail in the next chapter exactly why these configurations should be ungrammatical. For the moment, it is enough to notice that these examples suggest that whatever principles are constraining overt movement are also constraining expletive–associate relationships. That's a good indication that there really is movement happening in (71b) even though it's covert rather than overt.

Summary

In this chapter, we've taken the ideas from the previous chapter and gone one step further. Not only are there levels of representation that you don't 'see', like D-Structure, but there are also transformations that you don't

'see' – those that happen between S-Structure and Logical Form. As a result, the computational system of the language faculty now looks like as shown in Figure 7.1.

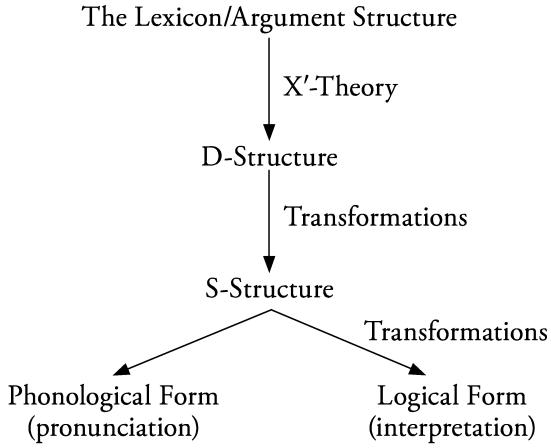


Figure 7.1 The computational system of the language faculty.

Every example and argument in this chapter has served to motivate and defend this claim about the existence of covert, post-S-Structure movement. In Section 7.1, we looked at the question of wh-movement and suggested that both in Japanese and in English there were wh-words which appeared *in situ*, that is in their D-Structure position, but which moved to Spec CP at LF. In Section 7.2, we made the same claim about quantificational DPs in English. In Section 7.3, we looked at structures of the *there... DP* form and argued that the DP raised at LF and replaced to the expletive.

What I hope to have convinced you of is that just because these transformations happen covertly doesn't mean that we can't find evidence for them. On the one hand, we saw a couple of instances where it looked as though transformations we were postulating at LF actually take place overtly in some languages. If semantically identical sentences in different languages are *structurally* identical at LF, then seeing a transformation which happens overtly *could* be evidence that it's happening at LF in another language where you don't 'see' anything. However, we also looked at evidence which to my mind is a little stronger. If the 'Y' model of grammar is right, and transformations are the same whether they happen between D-Structure and S-Structure or between S-Structure and LF, then they should be subject to the same constraints, like Subjacency or the Proper Binding Condition. To the extent that we find proposed LF movements patterning identically with their overt counterparts, we have strong evidence that the same thing is at work.

Open Issue: 'Superiority' and discourse-linking

Chomsky (1973) observed something which he called 'superiority' effects, which occur in multiple wh-constructions in English. A typical example is seen in the contrast in (79):[5]

(79) a. Who_i t_i bought what?
b. *What_i did who buy t_i?

Descriptively, when you have two wh-phrases either of which could move to Spec CP, it is the higher-up or 'superior' one which must move. Various people at more or less the same time (including Jaeggli 1982, Aoun et al. 1981 and Chomsky 1981) claimed that superiority effects provided another argument that there was LF wh-movement taking place in English multiple wh-constructions, which is the conclusion we were arguing for in Section 1 above.

However, Pesetsky (1987), citing observations by Chomsky and by Richard Kayne, notes that we don't see the superiority effects we'd expect when we turn from wh-words like *who* and *what* to lexical wh-phrases like *which book*. There are also some unexpected instances in Japanese where we don't see the Subjacency effects that we should.

To start off, Pesetsky considers the following contrast:

(80) a. Which man_i did you persuade t_i to read which book?
b. Which book_i did you persuade which man to read t_i?

Descriptively, (80b) seems to resemble the superiority configuration that we saw in (79b) above. However, while we had a contrast in (79), there's no contrast in (80). Why is that?

The problem is compounded by (81):

(81) Which man knows where which woman will live?

This looks like a real problem. (81) displays exactly the same ambiguity that we saw in (27). (81) can be interpreted either as a single question or as a double question. In the first case, there is only the one man whose identity is in question and the sentence is essentially synonymous with (82):

(82) Which man knows where each woman will live?

In the second interpretation of (81), there is a set of man-woman pairs that we need to find out about, with each man knowing where a different woman will live.

Just as we did when we talked about (27), that would seem to indicate that movement is really taking place, with the ambiguity residing in whether the movement is to the embedded Spec CP (single question) or the matrix Spec CP (double question). However, notice that *which woman* is not in a position from which overt movement is possible.

(83) *Who$_i$ does John know where t_i will live?

We seem to have a contradiction. We want, on the one hand, to say that *which woman* does undergo LF movement because of the ambiguity (and because that's the general position that we've been arguing for in this chapter). On the other hand, (83) makes it look like we want to say that *which woman* doesn't undergo movement, since the leading idea is that movement is movement, whether it happens at S-Structure or at LF, and the same constraints should apply. Therefore (83) and (81) should be the same.

We could just stipulate things and say that there's something different about wh-phrases. However, there's also a problem first noted by Lasnik and Saito (1984) concerning an absence of Subjacency effects in Japanese. Consider (84a) and its English translation in (84b):

(84) a. Mary-wa [$_{DP}$ [$_{CP}$ John-ni nani-o ageta] hito-ni] atta-no?
Mary-Top John-Dat what-Ace gave man-Dat met-Q
b. *What$_i$ did Mary meet [$_{DP}$ the man [$_{CP}$ who gave t_i to John]]?

The movement crosses two bounding nodes, DP and the embedded IP, and the result should be ungrammaticality. Unfortunately, we only see the ungrammaticality in English, not in Japanese.

There are various possible approaches to this problem. The standard one is that of Lasnik and Saito (1984, 1992), who say that Subjacency simply doesn't hold of LF movement, only of S-Structure movement. Notice that this would be a real problem for us, though. Many of our arguments for LF movement have assumed that Subjacency *does* constrain LF movement. Empirically, there's good evidence that it does. And conceptually, being able to say that all movement is the same, no matter where it happens, is also something we'd like to hang on to if we can.

Fortunately, Pesetsky (1987) has an approach which doesn't require saying that LF movement is any different from S-Structure movement. Instead, Pesetsky suggests that the difference is between two kinds of wh-words/phrases which appear *in situ*, and that this will address both the problem presented by (80)/(81) and (84). Pesetsky claims that some *in situ* wh-phrases are 'discourse-linked' (or 'D-linked') and others are not. *Which*-phrases are normally D-linked. That is to say, by using a which-question, the speaker is implying that there is some set of possible answers/objects which the speaker and the hearer (implicitly) share or have in mind. Thus, a question like *which band does John like?* is very odd without any context or background. One would expect that if that question were asked that there would have been a previous discourse about music or bands. On the other hand, 'pure' wh-words such as *who* or *what* are normally not D-linked. A question such as *what does John like?* does not contain any presupposition of sets or previous discussion and would be a perfectly appropriate way to open a new topic of conversation, for example.

The core of Pesetsky's suggestion is that non-D-linked wh-phrases have their scope decided by moving to a particular Spec CP at LF, which is what we've been assuming so far for all wh-phrases. On the other hand, those which are D-linked have their scope assigned by simple coindexation. Essentially what happens is that the +wh feature that's present in either the matrix or embedded CP has an index, like the ones we used for Binding Theory. All wh-words also have an index, and they're interpreted as being in whichever CP corresponds to the index they have. The crucial difference is that there's no actual movement of the wh-word under this indexation route.

This would explain all of the facts that we've seen so far. The reason why *which woman* in (81) is able to appear in a position from which movement isn't normally possible is that, being D-linked, it doesn't actually move at all. If the wh-word *nani* in the Japanese example (84) can also be shown to be D-linked, then we predict that it should not move and the lack of a Subjacency violation is expected. This is a potentially very interesting hypothesis, allowing us to retain the claim that all movement no matter where it happens is constrained by the same principles.

As Pesetsky observes, one way we could argue for this approach would be to take a normally D-linked wh-phrase and find a way to 'un-D-link' it somehow. If we could do this, we'd expect that it should then undergo movement and be subject to the constraints on movement that we've seen. Pesetsky suggests that phrases such as 'what the hell' or 'what on Earth' are ways to create an explicitly non-D-linked wh-element. In a question like *what the hell/what on Earth did you read that in?*, the point is to emphasize the surprise and imply a lack of connection to the previous discourse. *What the hell* can be also used colloquially as in (85), creating a wh-phrase which is, in Pesetsky's terms, 'aggressively non-D-linked':[6]

(85) What the hell book did you read that in?!?!

Interestingly, Japanese appears to have something which corresponds to *the hell* in English:

(86) Mary-wa John-ni ittai nani-o ageta-no?
Mary-Top John-Dat the hell what-Acc gave-Q
'What the hell did Mary give to John?'

Now if we were to go back and look at (84a) again, and this time use *ittai* to force a non-D-linked interpretation, we would predict that Subjacency effects would re-emerge, since the aggressively non-D-linked wh-phrase should undergo movement at LF. This seems to be the case:

(87) *Mary-wa [$_{DP}$ [$_{CP}$ John-ni ittai nani-o ageta]
Mary-Top John-Dat the hell what-Acc gave
hito-ni] atta-no?
man-Dat met-Q
'What the hell did Mary meet the man who gave to John?'

It's not that you can't use *ittai* in embedded clauses or something. (88), which isn't a Subjacency configuration, is perfect:

(88) Mary-wa [$_{CP}$ John-ga ittai nani-o y onda to] itta-no?
Mary-Top John-Nom the hell what-Acc read that said-Q
'What the hell did Mary say that John read?'

So it looks like there might be way in which we can maintain our assumption that S-Structure and LF movement both obey Subjacency, but still account for the lack of superiority effects in English that we saw in (80) and the lack of Subjacency effects in Japanese. It's not that there's any difference between S-Structure and LF movement. It's just that D-linked wh-phrases don't undergo movement.

FURTHER READING

Although the idea of Logical Form had been around in linguistics previously, the first systematic discussion of some of the ideas we've seen in this chapter is Robert May's (1977) dissertation, which was finally made more widely available in 1990. (It had originally been distributed only by the Indiana University Linguistics Club.) May was the first, for example, to propose and investigate Quantifier Raising as a transformation on a par with other overt movement. The idea that wh-movement is taking place between S-Structure and LF has been most thoroughly explored by James Huang since his dissertation in 1982. He has a good overview/summary in the Logical Form chapter of Webelhuth (1995a). Also, most Chomskyian syntax textbooks have discussions of LF and the ideas that we've been talking about. I particularly recommend Lasnik and Uriagereka (1988), but there's also good material in Haegeman (1994). There are also two books called simply *Logical Form* – one by May (1985) and one by Hornstein (1995). They've both got some interesting arguments, but can be a little difficult. I wouldn't recommend looking at them directly until after Chapter 8.

ANSWERS TO IN-TEXT EXERCISES

Exercise 1

Since we're operating under the assumption that the semantics of *know* and 'zhidao' are identical, we should see whether *know* requires, does not permit, or allows its sister to be a +wh CP in English. Depending on which of these is true, we would expect the sentence to be a statement, a question or ambiguous in Chinese, respectively.

(a) John knows [what Bill bought] +wh
(b) John knows [that Bill bought a car] –wh

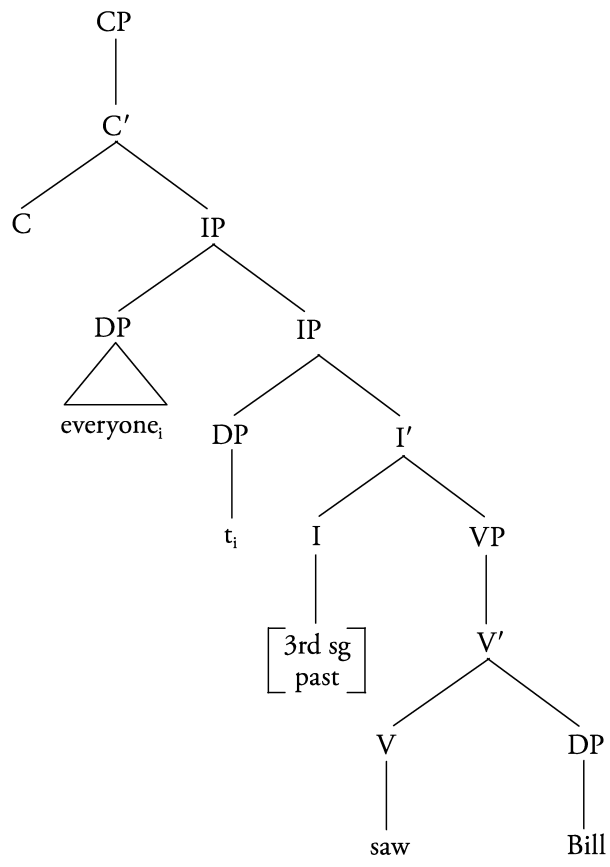
Since *know* appears to allow either a +wh CP or a –wh CP as its sister, *shenme* will move to either the embedded CP or the matrix Spec CP as appropriate. *Know* is therefore similar to *remember*, and we predict that the sentence in Chinese should be ambiguous between a statement and a question.

Exercise 2

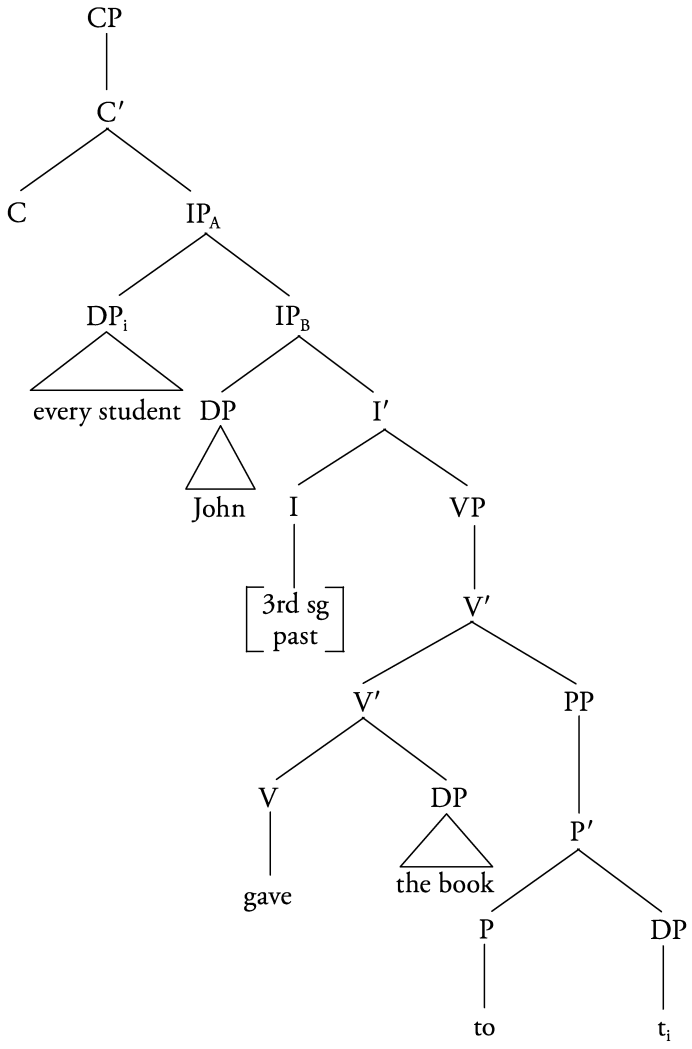
Under the 'double question' interpretation, (a) is asking for pairs of people and the times when they gave John the book. Two appropriate answers would be 'Bill gave the book to John yesterday' and 'Mary gave the book to John three days ago'. Example (b) is asking for pairs of things and locations where Bill put them. Two appropriate answers would therefore be 'Bill put the money in the drawer' and 'Bill put the keys on the tray'.

Exercise 3

(a)



(b)



Exercise 4

In the (a) sentence, either *some student* or *every day* may have wide scope relative to the other, creating a scope ambiguity. If *some student* has wide scope relative to *every day*, then the interpretation of the sentence is that there exists a single student who leaves home every day. If, on the other hand, *every day* has wide scope, then the sentence is interpreted as meaning that during every day there is some student or other who leaves home. In the (b) sentence, we see an ambiguity between the numeral quantifier phrase *three books* and the universal quantifier *every customer*. If *three books* has wide scope, then there are three specific books which John showed to every customer. If *every customer* has wide scope, then for every customer John showed three books to them and they need not be the same three each time.

FURTHER EXERCISES

Exercise 1

Draw the LF representation for the following sentences:

1. Everyone left yesterday.
2. John said that Bill believed some politician.
3. Zhangsan xiangzin [shei mai-le shu]? [Chinese]
   Zhangsan believe who bought book
   'Who does Zhangsan believe bought books?'

Exercise 2

As we mentioned in passing in this chapter, the indefinite article *a(n)* can also serve as an existential quantifier like *some*. Show that this is the case by constructing an argument using scope ambiguity like the one at pp. 207–11.

Exercise 3

Consider the following sentences:

1. Some student whispered that Bill read every book.
2. Some student believed that Bill read every book.

What are the relative scope possibilities for *some* and *every* in (1) and (2)? Does either one create a problem for our theory?

Exercise 4

Under the assumption that a pronoun must be c-commanded by an operator in order to serve as a bound variable, the following sentences can be used to argue for the existence of QR. How?

1. A report card about every student$_i$ was sent to his$_i$ parents.
2. *The woman who loved every man$_i$ decided to leave him$_i$.

Exercise 5

We assumed in Section 7.3 that, at LF, expletive *there* was replaced at LF by its associate, so that a sentence like 1 has the LF in 2:

1. There is a man in the room.
2. A man$_i$ is t$_i$ in the room.

What problem for this idea is created by the following pair of sentences?

3. There weren't many books on the shelf.
4. Many books weren't on the shelf.

Exercise 6

In section 7.3, we saw that subject-verb agreement had to be checked at LF, rather than D-Structure, in order to account for sentences like 1:

1. There were six people in the office.

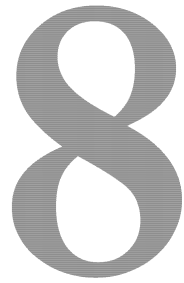
Does the level at which Binding Theory applies need to be similarly changed? What relevance do the following examples have?

2. $John_i$ bought every picture of $himself_i$.
3. *He_i liked every picture that $John_i$ took.

Is your conclusion affected by the following example?

4. Which claim that $John_i$ made did he_i deny?

Empty Categories and their Constraints



In Chapter 5, we introduced Principles A, B and C of the Binding Theory and showed how DPs fell into one of three types: anaphors (which include both reflexives and reciprocals), pronouns and R-expressions. We saw that anaphors were subject to Principle A, pronouns to Principle B and R-expressions to Principle C. In this chapter, we're going to re-examine this division based on what we saw about traces in Chapter 6 and conclude that there are actually two different kinds of traces, and that these two traces (together with PRO and another element called *pro*) are distinguished by Binding Theory. Traces of A-movement are anaphors, for example, while traces of A′-movement are R-expressions.

In this way, Binding Theory acts as a kind of indirect constraint on movement. If a moved element needs to bind its trace in its trace's governing category (because that trace is an anaphor) then that's going to constrain the places to which you can move. In the second part of the chapter, we'll see that another constraint on movement is also needed. By the end, we'll have developed a very important principle that constrains movement, called the Empty Category Principle.

8.1 Binding Theory: a recap and (minor) reformulation

To begin with, notice that we can pare our theory down a bit and eliminate a few stipulations with respect to Binding Theory. At the moment, we have three different 'labels' for DPs: anaphors, pronouns and R-expressions. We could reduce those three labels to two, by saying that there aren't three different kinds of DPs, rather DPs possess two features or properties, 'a' and 'p'. These features would indicate whether an element was subject to Principle A (the 'a', or anaphoric, feature) or Principle B (the 'p', or pronominal, feature). If a DP had the +a feature, it would be subject to Principle A. If it had the –a feature, it would not be subject to Principle A. The features +p and –p would work in the same way. Anaphors as we saw them in Chapter 5 would then correspond to [+a, –p] – a DP that was subject to Principle A and not Principle B. Pronouns would be the reverse, and therefore [–a, +p]. What about [–a, –p]? This would correspond to a DP which is subject to neither Principle A nor Principle B. Have we seen such a DP? Yes.

R-expressions, which are subject to Principle C. But what about the fourth possibility [+a, +p]? Well, this would have to be a DP which is subject to *both* Principle A and Principle B simultaneously. That would require a DP to be both bound and free in its governing category at the same time, which seems impossible. It therefore appears as though we don't need to worry about that case. It would seem to be excluded on logical grounds. So, summarizing, we have Table 8.1.

Table 8.1 DPs and their feature combinations.

Feature combination	*Type of DP*
+a, −p	Anaphor
−a, +p	Pronoun
−a, −p	R-expression
+a, +p	Cannot exist

However, we are now going to look at a further kind of DP that doesn't seem to have a place in Table 8.1: traces. Remember we argued in Chapter 6 that traces of passive, raising and wh-movement were DPs in their own right. The question obviously arises: are they anaphors, pronominals, R-expressions or none of the above?

You might be inclined just to place them outside the system by saying that Binding Theory applies only to overt DPs, that is DPs which are pronounced. Since traces aren't pronounced, Binding Theory doesn't apply to them. However, Chomsky (1982) suggested a more interesting way of looking at things. He argued that there are different kinds of unpronounced or 'empty' categories, which correspond to all four feature possibilities, including [+a, +p].

We've obviously got some work to do here. We've gone from just worrying about whether traces fit in to our a/p system to suddenly claiming that, using the terms of that system, there are four different kinds of empty categories! What we'll see in this chapter is that traces actually break down into two different kinds: traces of A-movement (passive and raising), which are anaphors [+a, −p], and traces of A′-movement (wh-movement), which are R-expressions [−a, −p]. Null pronouns, which would have the features [−a, +p] are usually assumed not to exist in English, but if we take a look at other languages, like Italian and Spanish, we can see them pretty clearly. The last one [+a, +p] will turn out to be our old friend PRO, the unexpressed subject of infinitivals, and we'll finally be able to clear up some of the things we had to leave as loose ends back in Chapter 4.

Before we can begin, though, we need to introduce one additional stipulation regarding Binding Theory which we didn't discuss in Chapter 5. Since that chapter came before we introduced movement, we only discussed DPs that were in A-positions like Spec IP or sister to V. (You need movement in order for DPs to get to A′-positions like Spec CP or 'adjoined to IP'.) It turns out that elements in A′-positions are 'invisible' for Binding Theory. When looking to see, for example, whether a given trace is bound in its governing category, only DPs in A-positions are considered.

8.2 The Binding Theory status of DP-trace

In trying to decide the status of different DP-traces, let's take a look at a typical case of A-movement, with a view towards what kind of DP the trace might be:

(1) $Mary_i$ seems t_i to be the best candidate.

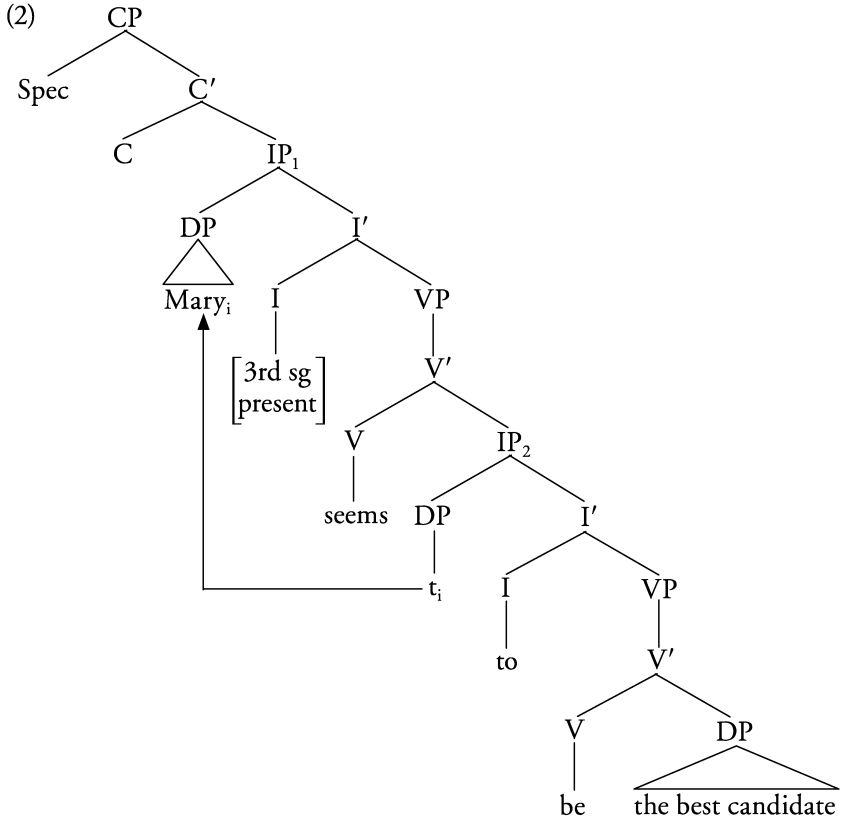
BEFORE CONTINUING

Draw the tree for (1) and determine the governing category for the trace.

One possibility seems to be excluded immediately. It doesn't look as though the trace could be an R-expression. R-expressions, remember, are subject to Principle C, which means that they can't be bound by an element in an A-position. But that's exactly what we have here. *Mary* is in an A-position, subject of the matrix clause. It c-commands the trace, and it's coindexed with the trace. Therefore *Mary* binds the trace and (1) would be predicted to be ungrammatical if the trace were an R-expression.

Let's see what the governing category of the trace is. That way, we can see whether *Mary* is outside it or inside it, and that will tell us whether we should home in on Principle A or Principle B. The first step as always is to draw the tree, and that gets us (2):

(2)



We need to find the minimal XP which contains: (a) the trace; (b) a governor for the trace; and (c) a subject which is accessible to the trace. IP_2 is the minimal maximal projection containing the trace, but it doesn't contain a governor for the trace (remember that infinitival *to* doesn't govern), so we'll have to look higher. V is a governor, and no barrier intervenes between the V *seems* and the subject (remember that IP doesn't count as a barrier). So the governing category could be VP. We're two-thirds of the way there. The problem is that, of course, VP doesn't contain a subject, so we need to go higher still. The matrix IP_1 contains a subject, *Mary*, and that subject is accessible to the trace, since coindexation between them does not create an i-within-i configuration.

So it looks like the matrix IP is the governing category for the trace. We've already established that trace is bound by *Mary*, and *Mary* is within the matrix IP. *Mary* is also in an A-position. Therefore the trace is bound, rather than free, in its governing category – and we have our answer. The trace has to be [+a, –p], subject to Principle A and not Principle B, which makes it an anaphor.

EXERCISE 1

Using the same logic, show that the traces of passive must also be [+a, –p]. (e.g. $John_i$ was killed t_i)

The conclusion that traces of A-movement are anaphors also does some other work for us that you might not have expected. Consider a sentence like (3) (these sentences are often referred to as 'superraising' violations):

(3) *[$_{IP1}$ $John_i$ seems that [$_{IP2}$ it is believed t_i by everyone]].

It's not immediately obvious why (3) isn't grammatical, meaning something like 'It seems that John is believed by everyone'. In particular, notice that Subjacency is *not* violated. The movement crosses only one IP: IP_2. IP_1 is *not* crossed, as it dominates both the departure site and the landing site. However, if traces of A-movement are anaphors, we have the explanation we're looking for. Since *John* is moving to an A-position (subject of the matrix sentence), t_i must be an anaphor. But if that's the case, notice that t_i won't be bound in its governing category. Its governing category is going to end up being the finite IP_2, and *John* is not within IP_2. Therefore the ungrammaticality of sentences involving 'superraising' violations follows from Binding Theory.

8.3 The Binding Theory status of wh-trace

Having looked at traces of A-movement, let's now take a look at traces of A′-movement. In Chapter 6, we didn't differentiate between the two. However, a moment's reflection will make it clear that we now need to. As just discussed,

traces of passive and raising are anaphors, but traces of A′-movement clearly can't be. To see this, consider a simple case of wh-movement, like (4):

(4) $[_{CP}$ Who$_i$ $[_{IP}$ t_i left]]?

If t_i, the trace of wh-movement, is an anaphor, (4) ought to be a Principle A violation. The trace is bound by *who*, but of course *who* is in Spec CP, which is not an A-position. *Who* will therefore be 'invisible' to Binding Theory, and this means that the trace has no binder. So if traces of wh-movement were anaphors, (4) would be predicted to be ungrammatical.

The fact that (4) is perfectly fine means that traces of A′-movement can't be anaphors. They must be something else. What other possibilities are there? The fact that the trace has no binder in an A-position would seem to rule out the trace being any kind of anaphor, either [+a, +p] or [+a, –p]. That leaves us with only two remaining possibilities. The wh-trace could be [–a, +p], a null pronoun or [–a, –p], a null R-expression.

Notice that the grammaticality of (4) is compatible with either possibility. Because the trace has no A-binder whatsoever, it would satisfy either Principle B or Principle C. So (4) by itself doesn't tell us whether the trace is a pronoun or an R-expression. What would distinguish the two possibilities is a sentence where the wh-trace is A-bound by a coindexed pronoun, but where the pronoun falls outside the trace's governing category. That's a case where Principle B would be satisfied, but Principle C would be *violated*. If the trace is a null pronoun, that configuration should be grammatical. If the trace is a null R-expression, the resulting sentence should be ungrammatical, since R-expressions must be A-free, full stop. Here's the relevant test sentence:

(5) *Who$_i$ does he$_i$ think $[_{CP}$ t_i' that Mary likes t_i]?

In (5), the pronoun is attempting to act as a bound variable, creating a 'crossover' configuration. In this case, it is an instance of *strong* crossover because the coindexed pronoun c-commands the wh-trace. If this indexation were possible, then (5) would be synonymous with (6):

(6) Who$_i$ t_i thinks that Mary likes him$_i$?

From a structural point of view, both (5) and (6) should have the interpretation 'which person x is such that x thinks that Mary likes x'. However, (5) cannot have that interpretation. It looks like the pronoun *he* can't have the same index as the wh-trace when the pronoun c-commands the wh-trace, as it does in (5). The pronoun must have a different index from the wh-trace, as in (7):[1]

(7) Who$_i$ does he$_j$ think $[_{CP}$ t_i that Mary likes t_i]?

We can explain this by assuming that traces of wh-movement are null R-expressions; that is, they have the features [–a, –p]. Under this assumption, the wh-trace must be A-free. There can be no element in an A-position which

c-commands the trace and is coindexed with it. But that's precisely the situation which arises in the ungrammatical (5). The pronoun in the matrix subject position c-commands the trace, which is in direct object position of the embedded clause. So if the pronoun and the trace have the same index, as would happen if we were trying to force the bound-variable interpretation, you violate Principle C and the result is ungrammaticality. If the pronoun has a different index than the wh-trace, as in (7), no A-binding of the pronoun takes place (remember *who* is in an A′-position) and there's no violation of Principle C.

Notice that this is now an argument that traces of A′-movement can't be pronouns. The offending pronoun *he* is outside the governing category of the wh-trace. If traces were [–a, +p], we'd predict that both (5) and (7) should be grammatical. Only the assumption that traces of A′-movement are null R-expressions correctly predicts that (5) is ungrammatical, but (7) is fine.

EXERCISE 2

Show that the wh-trace in (5) is not bound in its governing category.

8.4 The search for null pronouns

We now seem to have discovered analogues to two of the three kinds of overt NPs. We've got traces of A-movement, which are anaphors (+a, –p), and traces of A′-movement, which are R-expressions (–a, –p). Let's now see if we can find a null counterpart to overt pronouns. We're looking for some kind of empty category which is subject to Principle B of the Binding Theory.

Unfortunately, English isn't helpful here. The one outstanding empty category in English we haven't talked about is PRO, but we'll have reason to see in Section 8.5 that PRO is [+a, +p].

However, let's look at 'null subject' languages like Spanish and Italian, where the subject of a finite verb is normally unpronounced. So for example, the neutral way of saying *I am going to Madrid* in Spanish is:

(8) Voy a Madrid.
go-1sg to Madrid
'I am going to Madrid.'

In general, unless you were emphasizing that *you* were going to Madrid, as opposed to someone else, you would leave out the first person singular subject pronoun *yo*. Null subject languages contrast (unsurprisingly) with 'non-null subject' languages like English and French, where the subject pronoun must always be expressed.[2]

(9) *am going to Madrid

(10) *vais à Madrid (French)
go-1sg to Madrid
'I am going to Madrid.'

How do we approach sentences like (8) and the contrast between null subject and non-null subject languages? We could just say that a sentence like (8) has no subject, that is, that there's no DP in Spec IP, but it's pretty clear why we wouldn't want to do that: the argument structure of the verb *to go* is the same in Spanish and in English. The semantics are certainly the same. It's not that the Spanish sentence (8) has no concept of 'goer' in it. And if that's the case, then there will be an Agent θ-role which the verb will have to assign at D-Structure. An obvious way out of this problem is to say that (8) *does* have a subject, merely an unpronounced one.

So if Spanish and the other languages have a null subject where English would have a subject pronoun, could this null subject be the null pronoun that we're looking for? Let's take a look at a slightly more complicated sentence:

(11) Juan_i dijo que va a Madrid
Juan said that ø goes to Madrid
'Juan_i said that he_i is going to Madrid.'

For the same reasons as in (8), it must be the case that the embedded clause has a subject in Spec IP even though nothing is pronounced. (We've put 'ø' in (11) as an indication that there's some kind of empty category present.) Therefore, it must be the case that there is an empty category of some kind serving as the subject. Furthermore, that empty category is coindexed with the subject *Juan*, as can be seen from the interpretation.

Could this empty category which we've indicated as ø be either of the two that we've seen before? It seems unlikely. It can't be a trace of A-movement [+a, –p]. Those are anaphors. $ø_i$ does have a binder: *Juan*. But *Juan* is outside $ø_i$'s governing category, so Principle A won't be satisfied. That rules out $ø_i$, being [+a, +p] as well. But the fact that *Juan* is binding $ø_i$ from an A-position (albeit an A-position outside $ø_i$'s governing category) means that $ø_i$ can't be a trace of A′-movement. Those are R-expressions and need to be A-free.

So what we've got is an empty category that can be A-bound, but outside its governing category. Sound familiar? $ø_i$ has to be a null pronoun, [–a, +p]. By a process of elimination, that's the only possibility we're left with. Fortunately, this seems right. The subject position of the embedded clause in (11) is a place where pronouns can appear. It's a θ-position and is Case marked by finite I. Also, most tellingly, an overt pronoun can appear in that position when the embedded clause subject is not co-referential with the matrix clause subject.

(12) Juan dijo que vosotros vais a Madrid.
Juan said that you-pl go-2p1 to Madrid
'Juan said that you (pl) are going to Madrid.'

Furthermore, this null element seems to have all of the interpretive possibilities that overt pronouns have. That is to say, in addition to being bound outside its governing category, this null element can also be free, just like an overt pronoun. So, for example, if we're having a conversation about our friend Juan, I can say:

(13) $ø_i$ Dijiste que $ø_j$ va a Madrid
ø said-2s that ø go-3s to Madrid
'You said that he is going to Madrid.'

Here we've got empty categories both in the subject position of the matrix clause and the subject position of the embedded clause, and both are free. This makes it even more likely that we're dealing with null pronouns here.

We'll refer to this null pronominal element as 'pro', which is usually read/said as 'little pro'. The lowercase letters are important, because we need to keep it separate from PRO, the unexpressed subject of an infinitive, which we saw in Chapter 4 (which is called 'big *pro*'). English does of course have PRO, as we've seen, but it doesn't have *pro*, since subjects of finite verbs can't be null in English.[3]

Null objects in Italian

So far, we've only seen *pro* as a subject. If *pro* really is just an empty pronoun, it should be able to turn up in other positions as well. After all, overt pronouns aren't just restricted to subject position. They can be objects of verbs and prepositions – in fact, anywhere an overt DP can turn up.

Rizzi (1986) argues that *pro* can also appear as the object of certain verbs in Italian, like *condurre* 'to lead'. Rizzi points to contrasts like the following:

(14) a. Questo conduce la gente$_i$ a [PRO$_i$ concludere quanto segue].
This leads people inf. to conclude what follows
'This leads people to conclude what follows.'

b. Questo conduce pro a [PRO$_i$ concludere quanto segue].
This leads inf. to conclude what follows
'This leads one to conclude what follows.'

In one respect, the argumentation follows exactly the same line as we saw earlier. From the appearance of the DP *la gente* in (14a), it seems as though *condurre* has a θ-role to assign to its object. Since no DP is present phonetically in (14b) to receive that θ-role, the θ-criterion requires that there be a phonetically unrealized DP there. You might be inclined to object to this line of reasoning with the observation that perhaps *condurre* in Italian is like *eat* in English, in which the object θ-role is only optionally realized.

(15) a. John ate dinner.
b. John ate.

However, this possibility doesn't look too strong when we take a look at the English equivalent of *condurre*. As the translation indicates, *lead* seems to obligatorily assign a θ-role to its object.

(16) a. Societal problems should lead people to question the political system.
b. *Societal problems should lead – to question the political system.

If *lead* were like *eat* in having an internal θ-role which is optional, then we'd expect that (16a) and (16b) should both be grammatical, on a par with (15a) and (15b). The fact that they aren't suggests that the internal θ-role possessed by *lead* is obligatory. Since optionality/obligatoriness of θ-role assignment with respect to a given predicate is not something that we expect to vary cross-linguistically, given the close ties between θ-role assignment and semantics, the fact that *lead* in English has an obligatory internal θ-role strongly suggests that *condurre* does in Italian as well.

In fact, it seems that *pro* provides exactly the solution that we're looking for. *Lead* and *condurre* each have an internal θ-role which must be assigned. In Italian, that internal θ-role can be assigned to *pro*, which is phonetically null, so (14b) is grammatical. English, on the other hand, doesn't have *pro*, as we've seen, so (16b) constitutes a θ-criterion violation.[4]

8.5 [+ a, +p] revealed!

We've managed to track down three of the four logical feature possibilities in the domain of empty categories: traces of A-movement are null anaphors [+a, –p], traces of A′-movement are null R-expressions [–a, –p] and *pro* is our null pronoun [–a, +p]. That leaves [+a, +p].

Now you might think that [+a, +p] is going to be a lost cause. When we talked about the typology of *overt* elements at the beginning of the chapter, we just wrote off the possibility of [+a, +p] with the simple, if plausible, assertion that such a thing couldn't exist because it would have to be both bound and free in its governing category; which is, of course, impossible. However, I hope that by now you've learned not to jump to conclusions, however apparently plausible (my recommended approach to things in general, by the way, not just syntax). As it turns out, things are a little bit more complicated. What I'm going to argue is that PRO ('big pro'), the phonologically null subject of infinitivals and gerunds, has the features [+a, +p].

We saw in Chapter 4 that PRO seems to split into two types with regard to its interpretation. On the one hand, we have *controlled* PRO, as in (17):

(17) John_i promised PRO, to go home early.

Unlike ordinary pronouns, which can take an antecedent but don't have to, controlled PRO does require an antecedent. There's no possible situation under which (17) could mean 'John promised that somebody or other would go home early'. It can only mean 'John promised that he (John) would go home early'.

On the other hand, we also have *arbitrary* PRO, as in (18):

(18) PRO to go home early would be a terrible mistake.

Here PRO's interpretation has a much more pronominal feel to it. In fact, it's essentially synonymous with something like 'for one to go home early would be a mistake'.

The PRO Theorem

Now it's one thing to say that PRO has both anaphoric and pronominal interpretive tendencies and quite another to say that it literally has the Binding Theory features [+a] and [+p]. We still haven't addressed the fundamental question of how something could simultaneously satisfy both Principle A and Principle B of the Binding Theory.

Let's begin by taking a closer look at Principles A and B, and see if we've got any hitherto unsuspected room to manoeuvre.

(19) a. *Principle A*: An element with the feature [+a] must be bound in its governing category.
b. *Principle B*: An element with the feature [+p] must be free in its governing category.

On the surface, they do look contradictory, but let's dig a little deeper. The contradiction is coming about because of the way *bound* and *free* are defined. Since *free* means 'not bound', it's hard to imagine how something could be both bound and not bound.

However, don't lose sight of that last little prepositional phrase 'in its governing category'. The Binding Theory doesn't say that an anaphor must be bound, full stop, or that a pronoun must be free, full stop. What it says is that they need to be bound or free *in their governing categories*. How does that help us? Well, if you had an element which *didn't have* a governing category for some reason, maybe that would fall outside the system somehow because Principles A and B presuppose that the element *does* have a governing category.

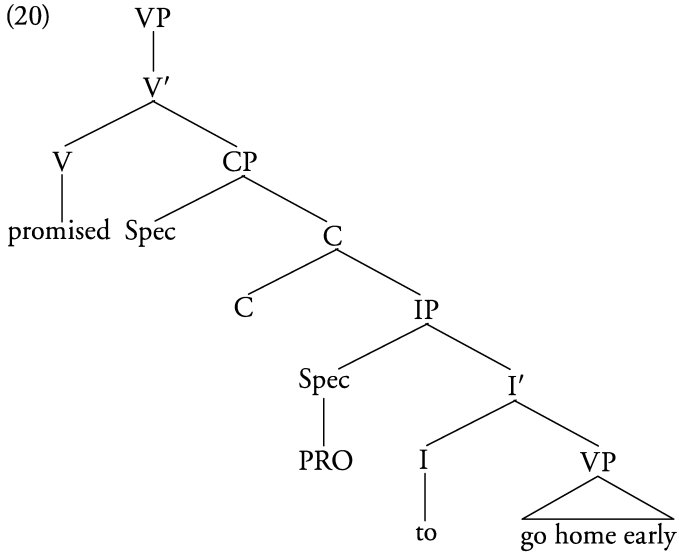
Let's pursue that thought. The first thing we need to think about is how you could ever have an element which didn't have a governing category. From Chapter 5 you'll remember that the governing category for a given element is the minimal maximal projection which contains: (i) the element itself; (ii) a governor for the element; and (iii) a subject which is accessible to the element. If for whatever reason an element doesn't have one of these three things, then it could be reasonably said that it doesn't have a governing category.

Starting with (i), it looks like there's no help there. There will always be some category which contains the element. (iii) doesn't look too promising either. As we saw back in Chapter 5 the subject of the clause that the element is in will always be an accessible subject, unless the element is inside a subject, in which case the subject of the next higher clause will serve.

(ii) looks more promising. If an element has no governor, then it makes sense to say that it has no governing category. And it looks like PRO fits the bill. First,

let's take a look at a partial structure for (17), the case of controlled PRO:

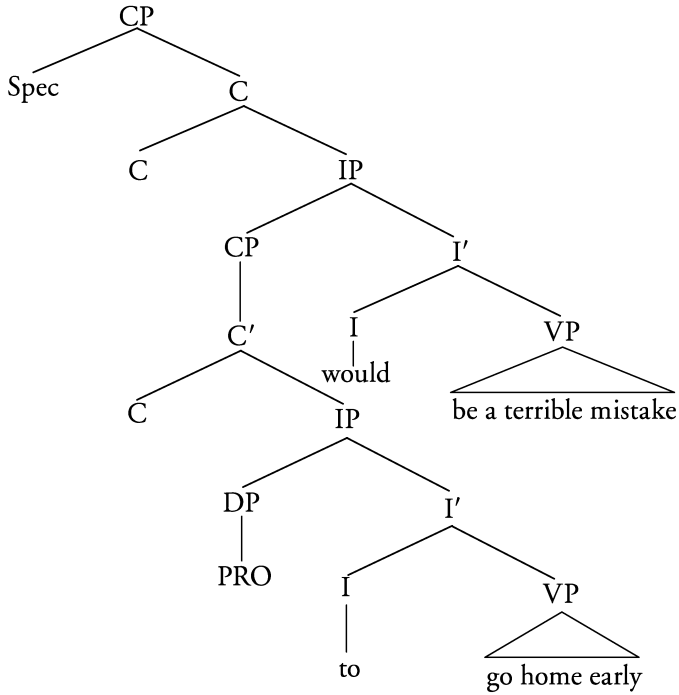
(20)



Here PRO doesn't have a governor. Government from outside the CP, by *promise* for example, is blocked by the barrier CP. And there's no government from I°, the other possible source. Remember that non-finite I is defective: unlike finite I, it doesn't govern its specifier position, Spec IP.

The same situation obtains in the case of arbitrary PRO. If we take a look at a partial structure for (18), we've got (21):

(21)



Again, government from the matrix finite I is blocked by the barrier CP and there's no government from non-finite I because it's defective.

So it looks like PRO, both controlled and arbitrary, has no governor. It therefore seems plausible to say that PRO has no governing category. So what effect does this have with respect to Principles A and B? Are Principles A and B satisfied or violated by an element which has no governing category?

The traditional claim is that elements which have no governing category *satisfy* Principles A and B automatically. This intuition might seem slightly odd. Is (19a) 'satisfied' by an anaphor that has no governing category? But I think the intuition becomes clearer if we reverse things and ask ourselves whether Principle A or B is *violated* by an element which has no governing category. If you can't figure out whether something is bound or free in its governing category because it simply has no governing category, it's hard to say that Principle A or B is violated. And it seems reasonable to define 'satisfies' as equivalent to 'doesn't violate'. This is what we will do.

This is not just a bit of terminological juggling. If we can find it in our hearts to say that Principles A and B are satisfied when an element has no governing category, it turns out that some interesting (and correct) predictions about the distribution of PRO seem to follow automatically from the assumption that it is both [+a] and [+p].

If not having a governing category is the only way to satisfy both Principles A and B at the same time, then it follows automatically that PRO can *never* appear in a governed position. That is, it follows that it can never be the subject of a finite clause, the object of a verb, the subject of an ECM infinitive or the object of a preposition. These predictions all seem right:

(22) a. *PRO goes to the market every Sunday
b. *I hit PRO
c. *I believe PRO to be here
d. *I gave the book to PRO

If we hadn't made the assumption that PRO was [+a, +p], and therefore couldn't appear in a governed position, it's hard to see what other principle rules out (22a) to (22d). Interpretively, PRO can either be controlled or arbitrary, so, on a par with (18), (22a) could mean that some unspecified person or people in general go to the market every Sunday. And (22b) could mean something like *I hit myself*. But the fact is that they don't. They're just ungrammatical. But fortunately, instead of having a mystery, we have an explanation.

Putting all of these ideas and arguments together, we have what's known as the PRO Theorem:

(23) *The PRO Theorem*
PRO must be ungoverned.

This is nice in that it's a genuine *theorem*, as opposed to a description or a rule-of-thumb or something. That is to say, this fact about PRO follows completely from other principles. Once we say that PRO has the both the feature [+a] and the feature [+p], it follows that PRO cannot have a governor.

Notice that at last we've got a genuine explanation for why there's no overt counterpart to PRO; that is, an overt element which has the features [+a] and [+p]. As we've seen, the only way that such an element could satisfy both Principles A and B of the Binding Theory would be to have no governing category, and for that it would have to have no governor. But if an element doesn't have a governor, then it can't be assigned Case. (Remember from Chapter 4 that you have to govern an element in order to Case mark it.) And an overt element which doesn't get assigned Case is always going to violate the Case Filter. Therefore, you could never have an overt element which was [+a, +p] as it would always result in a Case Filter violation. *Voilà!*

8.6 Interlude: a note on Plato's Problem

Let's pull back and reflect on what we've just discovered. When you think about it, all these things that we've just been talking about are vivid examples of Plato's Problem. Remember from the Introduction that Chomsky argued that you seem to know a lot more than just what you could have figured out from the data/acoustic disturbances that you were exposed to in childhood. There must be some guiding hand, in the form of a Language Acquisition Device, which makes possible the acquisition of such a body of knowledge, which is not only very intricate, but also quite independent of anything you can actually see or hear.

First of all, there's the concept of empty categories in general. We've found what I think is pretty solid evidence that part of knowing your language consists of knowing that there are elements in the grammar *which are never pronounced*. That in itself is really surprising. In acquiring English, you could have had no direct evidence that they were there, so how did you figure out what was going on? It would be like trying to learn the rules of chess by watching games played, but where crucial pieces were invisible.

Furthermore, it's not just that you know that there are invisible pieces. You know that there are, and you distinguish between, different *kinds* of invisible pieces – *three* if you're an English or French speaker and *four* if you're a Spanish or Turkish speaker. And they all have different properties! In English, you've got traces of A- and A′-movement, which are anaphors and R-expressions respectively, and you've also got PRO which is *both* an anaphor *and* a pronominal – something which couldn't exist if it were an overt category. And if you're a speaker of a null subject language, you've got all these plus *pro* to deal with.

It's hard to see how analogy, inductive generalization or any other kind of generalized learning mechanism could explain how you know facts like these. Nonetheless, the intuitions are sharp and something that every speaker knows. It's these kinds of issues that lie at the heart of Chomsky's 'Logical Problem of Language Acquisition'. Not only did you manage to figure out all this information from the data you were exposed to, but *everybody* did regardless of intelligence. How did everyone converge on exactly the same thing given the varying primary linguistic data they were exposed to? The only answer seems to be that an innate, inborn, guiding force, Universal Grammar, is at work in the acquisition process.

8.7 The Empty Category Principle

In this section, we're continuing our exploration of conditions and constraints on movement. We've seen a few of these already. In Chapter 6, we discussed the Subjacency Condition, which prevents a single movement operation from crossing more than one bounding node. In the previous sections, we saw how Binding Theory can act as an indirect constraint on movement. Because traces of A-movement are anaphors, for example, the antecedent of the trace must bind the trace within its governing category. In this section, though, we're going to see that more is needed. By the end, we'll have developed a very important principle that constrains movement, called the Empty Category Principle (ECP).

A beginning: the government requirement on traces

As we just mentioned, we've seen that Subjacency and Principle A act as constraints on movement, but it turns out that we need more. Consider the following D-Structure:

(24) e is crucial John to see this
(cf. e seems John to see this)

The first question we need to ask ourselves is 'what's the structure?'. In particular, does *crucial* take an IP or a CP complement? In order to help decide this, let's take a look at the sentences in (25):

(25) a. It is crucial PRO_{arb} to understand that change is needed.
b. It is crucial for Mary to attend the meeting.

Since we appear to get PRO and *for/to* infinitives, it looks as though government by the adjective *crucial* is blocked. Therefore, *crucial* must take a full CP complement as opposed to an IP complement. Otherwise, PRO would be governed by *crucial* in violation of the PRO Theorem (since IP isn't a barrier), and we'd have no explanation for why a Case-assigning complementizer like *for* can appear.

We can see this more clearly if we compare *be crucial* in (25) with *be likely* in (26), which, as a raising predicate, we know takes a 'bare' IP complement:

(26) a. *It is likely PRO_{arb} to understand that change is needed.
b. *It is likely for Mary to attend the meeting.

In (26a), the 'defective' IP allows PRO_{arb} to be governed by the adjective *likely*, in violation of the PRO theorem. In (26b), there's no CP, and therefore no C^{o} position to host *for*. The grammaticality of (25a) and (25b) as compared with (26a) and (26b) would seem to confirm that *crucial* takes CP, rather than an IP complement.

Now as (24) stands, the matrix predicate *is crucial* needs a subject, and *John* needs Case, but we could solve both these problems by treating (24) like a raising structure and doing A-movement from the embedded subject position to the matrix subject position. If we do this, the result is (27):

(27) *John$_i$ is crucial [$_{CP}$ t$_i$ to see this]

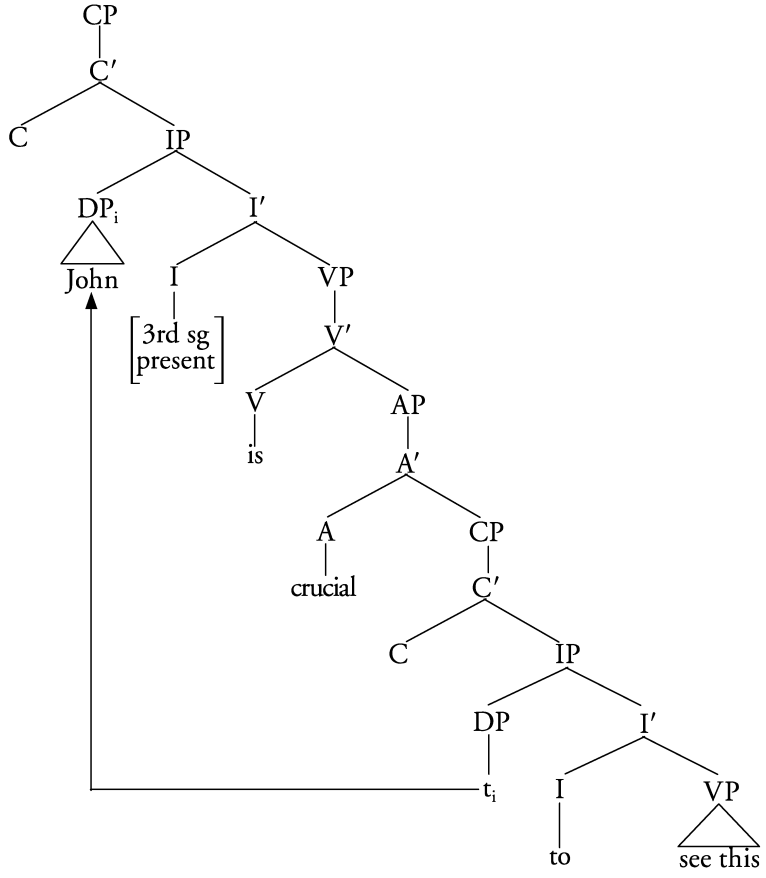
There's nothing incoherent about the thought that (27) expresses. It could just mean something like *it is crucial for John to see this* – by analogy with other raising structures that we saw in Chapter 6. However, instantaneously and without prior instruction, native speakers of English know that (27) is ungrammatical. And that's a problem, because nothing that we've got so far predicts that (27) is ungrammatical.

BEFORE CONTINUING

Draw the tree for (27). Verify that neither Subjacency nor Principle A is violated.

If we were to try to rule (27) out, we'd have to appeal either to Subjacency or to Principle A, but neither one of those is violated here. Consider first Subjacency:

(28)



As you can see, the movement in question crosses only one bounding node (the embedded IP), so there's no Subjacency violation. You might think that maybe we can change Subjacency so that AP counts as a bounding node. DPs are bounding nodes, after all. However, that solution's not going to work. If we included AP among the bounding nodes, we'd end up ruling out standard raising sentences like (29):

(29) John_i is $[_{AP}$ likely $[_{IP}$ t_i to be here]].

The movement in (29) would hypothetically end up crossing two bounding nodes (IP and AP) and we'd predict that (29) has the same ungrammatical status as (27). So it doesn't look like changing Subjacency is the right move to make.

What about Principle A then? That seems at first glance like it might be a better alternative. Surely the movement in (28), which crosses a CP barrier, must result in a trace which isn't bound in its governing category.

In order to see if this is true, we need to find the trace's governing category. Remember that that's the minimal XP containing the trace, a governor for the trace and a subject which is accessible to the trace. It's obvious where the trace is, but notice what happens when we come to look for a governor. I can't serve as a governor because only *finite* I is a member of the class of governors. C is empty and so isn't a possible governor. And government from outside CP is blocked because the maximal projection CP is a barrier. Therefore, the trace has no governor and therefore has no governing category. (This result shouldn't really be too surprising, by the way. We saw earlier that PRO can appear in the position that the trace is in (subject of a CP infinitival). Remember from the last section that PRO cannot have a governing category, and therefore can only appear in positions which are ungoverned.)

But look at the problem that's created by the lack of a governing category for the trace: we had to say in the last section that things which have no governing category don't violate *either* Principle A or Principle B. That's how PRO can be both [+a] and [+p]. Therefore, the trace in Spec IP of the embedded clause in (28) must *satisfy* Principle A! And that means that there's no possible Binding Theory violation.

So what can we do about this? It looks as though there's some condition on movement/traces that we need to add over and above Subjacency and Principle A. Given the data that we've looked at so far, it seems as though the easiest thing to say would be that traces of A-movement have to be governed:

(30) *Government Requirement on A-movement*
Traces of A-movement must be governed.

A government requirement on traces would be sufficient to rule out sentences like (27). Non-finite I doesn't govern, and government from outside is blocked by the CP. Other adjectives which are going to cause a similar problem are *vital*, *probable* and *necessary*.

Let's now take a look at A′-movement to see whether (30) is more general or something specific to A-movement. If we try to do A′-movement from the

subject of the embedded IP in (29), we get (31):

(31) *[$_{CP}$ Who$_i$ is it crucial [$_{CP}$ t$_i'$ [$_{IP}$ t$_i$ to be here]]]?

Just like (27), it's not that the thought expressed by (31) is incoherent. It should mean 'who is the person x such that it is crucial for x to be here?'. However, like (27), (31) is simply ungrammatical. The fact that (31) is just as bad as (27) suggests that the government requirement *should* be extended to traces of A′-movement as well, giving us (32):

(32) *Government Requirement on Traces*
Traces must be governed.

A government requirement on traces means that traces and PRO will be in complementary distribution. Traces must be governed, but PRO must not be governed (the PRO Theorem). That looks like a valid generalization. Examples in the last section, where we said that PRO cannot appear because the positions are governed, include things like the subject of a finite verb, the direct object and the ECM object, as illustrated by (33):

(33) a. *PRO goes to the market every Sunday.
b. *I hit PRO.
c. *I believe PRO to be here.

These are all positions in which traces can appear:

(34) a. Who$_i$ t$_i$ goes to the market every Sunday? (trace of A′-movement)
b. John$_i$ was hit t$_i$. (trace of A-movement)
c. John$_i$ is believed t$_i$ to be in the area. (trace of A-movement)

8.8 From government to proper government

However, it looks as though just adding the government requirement on traces in (32) to the list of constraints on movement won't be sufficient by itself. To see why, consider the following contrast:

(35) a. Which suspect do you think [$_{CP}$ t$_i'$ [$_{IP}$ t$_i$ committed the crime]]?
b. *Which suspect do you think [$_{CP}$ t$_i'$ that [$_{IP}$ t$_i$ committed the crime]]?

This illustrates the so-called 'that-trace' effect. Normally in English, complementizer *that* is completely optional. You can put it in or leave it out, and there's no change in interpretation or grammaticality:

(36) a. I think ø the police arrested Bill.
b. I think that the police arrested Bill.

(37) a. Which suspect do you think [$_{CP}$ t_i' ø [$_{IP}$ the police arrested t_i]]?
b. Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ the police arrested t_i]]?

However, when you do wh-movement from the *subject* position of an embedded clause, complementizer *that* must be left out, as in (35a). There's some problem caused by a sequence of *that* followed by a trace (35b).[5] In this section we will focus on the 'that-trace' paradigm, particularly on the ungrammatical (35b), and the contrast between (37b) and (35a).

BEFORE CONTINUING

Show that Subjacency, Principle C and the Government Requirement on Traces are satisfied in (35b).

To start with, notice that nothing we've said so far explains why (35b) is ungrammatical. Since the movement proceeds in a successive cyclic fashion, Subjacency is never violated. Movement from the embedded subject to Spec CP crosses only one bounding node, the embedded IP, and subsequent movement from the embedded Spec CP to the matrix Spec CP crosses only one bounding node, the matrix IP. As we saw in the last chapter, the trace of wh-movement is an R-expression, but there is no Principle C violation. The only elements in (35b) which are potential Binding Theory antecedents of the trace are the wh-phrase itself and its intermediate trace. However, both of these elements are in Spec CP, which is an A′-position, and all elements in A′-positions are invisible for Binding Theory purposes. Therefore neither the wh-phrase nor the intermediate trace binds the trace as far as Binding Theory is concerned. Finally, the government requirement on traces is also satisfied. The embedded I is tensed (i.e. it's not infinitival) and so assigns Nominative Case to the embedded subject position. Since we have Case assignment, we have government to that position as well. Therefore, nothing we have so far predicts the ungrammaticality of (35b).

Clearly, we need to modify our theory in order to explain why (35b) is ungrammatical. However, in doing so we need to steer a course between the devil and the deep blue sea, as represented by the grammatical (35a) and (37b).

(35) b. *Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ t_i committed the crime]]?

(35) a. Which suspect do you think [$_{CP}$ t_i' [$_{IP}$ t_i committed the crime]]?

(37) b. Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ the police arrested t_i]]?

The only difference between (35a) and (35b) is the presence of complementizer *that*. So that must be relevant for the solution. But comparing (35b) and

(37b) suggests that the difference between subjects and objects is relevant as well. Movement of the embedded subject over the complementizer *that* causes a problem, but movement of the embedded direct object doesn't. We'll see that *proper government*, a stronger variety of government, will allow us to account for 'that-trace' effects. We'll be able to predict that (35b) is ungrammatical, while still allowing (35a) and (37b).

(35b) vs (37b)

Let's begin by comparing the ungrammatical (35b) with the grammatical (37b):

(35) b. *Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ t_i committed the crime]]?

(37) b. Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ the police arrested t_i]]?

Both sentences have an overt complementizer *that*. The difference between the two is that in (35b) you have extraction of a subject and in (37b) you have extraction of a direct object. What we see is a *subject/object asymmetry*. Subjects and objects are not being treated in the same way.

As usual, in order to figure out where we should modify our theory, we should think about what principles are relevant and what the structure of the relevant sentences is. The principles that we have at the moment which govern movement and traces are (1) Subjacency, (2) Binding Theory and (3) the Government Requirement on Traces. It doesn't look like modifying Subjacency will work. There are actually *fewer* nodes in total crossed in the ungrammatical (35b). Modifying Binding Theory looks hopeless too.

Since traces of A′-movement are R-expressions and must be free, there would have to be a c-commanding antecedent somewhere in (35b) that wasn't in (37b), but that looks impossible. All of the elements which c-command the trace in (35b) also c-command the trace in (37b). Again, if anything, (37b) should be the ungrammatical one!

The last principle we have to check is the government requirement on traces. Is there some government-related difference that we can exploit to explain what's going on? If not, we're out of luck. However, there does seem to be a small sliver of light. In (35b), the governor of the subject position is I, a functional category. On the other hand, the governor of the direct object position in (37b) is V, a lexical category.

That seems to be the only difference we've got, so let's go with it and incorporate that difference into our theory. Anticipating the discussion a little bit, and bringing ourselves into line with standard terminology, let's say that the requirement isn't just that traces have to be governed, but rather that they must be *properly* governed, and that being governed by a *lexical* category is one way to be properly governed. This requirement that traces be properly

governed is referred to in the literature as the *Empty Category Principle* (usually abbreviated ECP):

(38) *The Empty Category Principle*
Traces must be properly governed.

(39) α properly governs β if and only if
(i) α governs β, where α is a lexical category (usually referred to as 'lexical proper government').

Using (38) and (39), we can explain the difference between (35b) and (37b) in terms of proper government. The trace in (37b) is properly governed by the verb *arrested*. The trace in (35b) on the other hand is not properly governed and therefore violates the ECP. We therefore predict that (37b) is grammatical but (35b) is not.

(35b) vs (35a): a second kind of proper government

Lexical proper government accounts for the difference between (35b) and (37b), but notice that we've gone too far. We correctly predict that (37b) is grammatical and that (35b) is ungrammatical, but we now predict that (35a) is *also* ungrammatical:

(35) a. Which suspect do you think $[_{CP}\ t_i'\ [_{IP}\ t_i$ committed the crime]]?
b. *Which suspect do you think $[_{CP}\ t_i'$ that $[_{IP}\ t_i$ committed the crime]]?

In both (35a) and (35b), the trace in subject position is not governed by a lexical category. When the complementizer is absent (as in (35a)), the only governor is I, a functional category. When the complementizer is present (as in (35b)), both I and C are potential governors (since IP is not a barrier), but, again, both are functional categories. Therefore, both (35a) and (35b) violate the ECP as we have it in (38) and both are predicted to be ungrammatical.

In order to solve this problem, we've got to pull back somehow. If the grammaticality of (35a) doesn't follow from lexical proper government, perhaps there's some other path to proper government. Is there some government-related reason why (35a) is grammatical, but (35b) is not?

Maybe there is. Notice that in both (35a) and (35b) the intermediate trace m-commands the trace in subject position (the first maximal projection dominating the intermediate trace is CP). The maximal projection IP intervenes, but as we saw before IP is the one maximal projection which doesn't act as a barrier. Therefore, if we expanded the list of governors to include intermediate traces, the intermediate trace in the embedded Spec CP would govern the trace in subject position. (Only heads govern the way we have the rule at the moment.)

Let's assume that this is the second path to proper government that we were looking for. Even if you're not governed by a lexical category, you can still be

properly governed provided that you're governed by an antecedent. This other type of proper government is referred to in the literature, not too surprisingly, as *antecedent government*. So, we need to keep the ECP in (38), but modify the definition of proper government in (39):

(40) α properly governs β if and only if
(i) α governs β, where α is a lexical category (lexical proper government), or
(ii) α governs β, where α is an antecedent of β (antecedent government).

Adding antecedent government to lexical proper government as the two ways traces can be properly governed, we'll correctly predict the grammaticality of (37b) and (35a). In (37b), the trace in direct object position is properly governed because it's lexically properly governed by the verb *arrested*. It couldn't be antecedent governed because, for example, the barrier VP intervenes. In (35a), the trace in subject position is properly governed in virtue of being antecedent governed by the intermediate trace. It can't be lexically properly governed because the trace is only governed by I, which is not a lexical category.

But now the pendulum has swung too far in the *other* direction. We've explained the grammaticality of (35a) by saying that the trace is antecedent governed, but that explanation should carry over to (35b). We've gone from predicting that (35a) and (35b) are both ungrammatical to predicting that they're both grammatical. However, with one more little push, we can get the three-way distinction that we need.

A minimality requirement

As usual, let's take a look at what's different about (35b) and see if there's a government-related difference that we can appeal to. However, we want to make sure that (35a) and (37b) aren't once again (wrongly) predicted to be ungrammatical! Notice that the overt complementizer is a head, and, since again IP is not a barrier, the complementizer will govern the subject position just as much as the intermediate trace antecedent in Spec CP does. This means that, although both the complementizer and the antecedent will govern the trace in subject position, in (35b) the antecedent is not the *closest* governor. Instead, the complementizer is the closest governor. In the grammatical (35a), by contrast, there is no closer governor than the antecedent.

If we impose a *minimality* requirement on government, we can explain the difference between (35a) and (35b). If you're governed by some element, then that element blocks government by any other element which is further away. This idea is already present in our definition of government via the notion 'barrier'. Something inside VP, for example, can only be governed by V, the closest head, and not by a head from outside, because the maximal projection VP blocks government from anything outside. However, we need to make sure that minimality is respected for *all* forms of government, both government by heads and antecedent government.

So, not only do you need to be either lexically properly governed or antecedent governed, but the governor involved must also be the closest governor to you. This has the effect of blocking antecedent government by the intermediate trace in (35b), because the intermediate trace is not the closest governor. The closest governor is C, which is not a lexical head, and we thereby correctly predict that (35b) is ungrammatical. In the grammatical (35a) on the other hand, the C position is not filled. Therefore the closest governor is the intermediate trace in Spec CP, and the trace can be antecedent governed.

(37b) is unaffected by this. We continue to (correctly) predict that it should be grammatical. When extracting from a direct object, the trace will always be lexically properly governed by the verb. This means that the complementizer *that* may be either present or absent because the verb will always be there to provide lexical proper government. We therefore predict that both (37a) and (37b) are grammatical.

So, by way of summary, let's put together all the definitions that we have so far, including the changes that we've needed to make to the definition of *government*:

(38) *The Empty Category Principle*
Traces must be properly governed.

(40) α properly governs β if and only if
(i) α governs β, where α is a lexical category (lexical proper government), or
(ii) α governs β, where α is an antecedent of β (antecedent government).

(41) α governs β if and only if
(a) α is a head or α is coindexed with β (this covers the antecedent government cases), and
(b) α m-commands β, and
(c) no barrier intervenes (where any maximal projection except IP is a barrier), and
(d) minimality is respected.

With these definitions, we can now explain the 'that-trace' paradigm that we started the section with:

(35) a. Which suspect do you think [$_{CP}$ t_i' [$_{IP}$ t_i committed the crime]]?
b. *Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ t_i committed the crime]]?

(37) a. Which suspect do you think [$_{CP}$ t_i' ø [$_{IP}$ the police arrested t_i]]?
b. Which suspect do you think [$_{CP}$ t_i' that [$_{IP}$ the police arrested t_i]]?

When extracting from a subject, you rely on antecedent government to satisfy the ECP, since the subject position is not lexically properly governed. But the presence of an overt complementizer blocks antecedent government

because the antecedent is not then the *closest* governor (a minimality violation). We thus predict that (35a) is grammatical but (35b) is not. On the other hand, since direct objects don't rely on antecedent government, the complementizer has no effect on grammaticality. It can either be present (37b) or absent (37a) because the verb will always be there to provide lexical proper government.

8.9 Subjacency vs the ECP

Having motivated the ECP in the last section, I want to show how it interacts with one of the other principles of movement that we've been talking about: the Subjacency Principle. As we discussed in the Open Issue in Chapter 4, we can use degrees of ungrammaticality to argue for the existence of separate grammatical principles. We can explain why sentence A feels more ungrammatical than sentence B by saying that sentence B violates more principles than sentence A. This provides a way of showing the reality of the principles that we're proposing – by showing that they correspond with our intuitions about degrees of ungrammaticality. In the particular cases under discussion, if a sentence which contains both a Subjacency violation and an ECP violation feels more ungrammatical than a sentence with either one alone, then we'll have good evidence for the existence of these as separate principles. If, on the other hand, the violations don't feel different, then we'll have evidence that Subjacency and the ECP shouldn't exist as separate principles of grammar, and we should look for ways to collapse them or reduce one to the other.

With this in mind, let's remind ourselves of what constitutes a standard Subjacency violation:

(42) *What$_i$ were [$_{IP}$ you wondering [$_{CP}$ who$_j$ [$_{IP}$ t$_j$ read t$_i$]]]?

In (42) we have a typical Subjacency violation. *Who* is filling the embedded Spec CP, forcing *what* to move in one fell swoop to the matrix Spec CP. In doing so it crosses the embedded IP and the matrix IP, resulting in a Subjacency violation.

BEFORE CONTINUING

Does (42) violate the ECP? Explain why or why not.

Let's see what we've got in (42) with respect to the ECP. The trace t$_i$ in direct object position is properly governed, as it is governed by the verb *read*. The trace in subject position t$_j$ is antecedent governed by *who*. Therefore, (42) satisfies

the ECP and only Subjacency is violated. As a way of getting a handle on our intuitions, notice these mere Subjacency violations have a vaguely comprehensible feel to them. On the other hand, if we change the wh-phrases around, we get almost a 'word salad' effect:

(43) **Who$_j$ were [$_{IP}$ you wondering [$_{CP}$ what$_i$ [$_{IP}$ t_j read t_i]]]?

Intuitively, (43) feels much worse than (42). We're in a position to explain this fact under the assumption that the ECP and Subjacency are separate principles.

BEFORE CONTINUING

Are t_i and t_j properly governed in (43)?

First of all, in (43) there's a Subjacency violation, just as there is in (42). *Who* has to be moved in one fell swoop to the matrix Spec CP, crossing two bounding nodes, the embedded IP and the matrix IP (remember that we cannot fill a single Spec CP position twice). However, (43) also contains an ECP violation. The trace of the subject, t_j, fails to be lexically properly governed. The only governor is the functional element I. But, unlike (42) the subject trace in (43) also fails to be antecedent governed. *What* in the embedded Spec CP can't antecedent govern t_j, because it's not coindexed with t_j. However, by minimality, what it does do is block antecedent government by anything further away, including t_i's 'real' antecedent *who*. That leaves the subject trace stranded and unable to satisfy the ECP either by lexical proper government or by antecedent government. Therefore, while (42) violates one principle – Subjacency – (43) violates both that principle and a second principle – the ECP. And the fact that (43) is significantly worse than (42) provides us with an argument for keeping the two principles independent and separate. In the literature, violations like (42) are usually referred to as 'weak' violations (or 'mere Subjacency violations'), whereas (43) is a 'strong' (Subjacency plus ECP) violation.

8.10 Adjuncts and antecedent government

In motivating the ECP so far, we've looked particularly at two phenomena: that-trace effects (Section 8.8), and weak vs strong 'island' violations (Section 8.9). In doing so, we've been focusing on the ways in which subjects are different from direct objects. We've seen that subjects trigger that-trace violations when the complementizer is present and they also cause strong violations when extracted from a wh-island. Direct objects, on the other hand, *can* be extracted over an overt complementizer, and they trigger *weak* violations when extracted out of a wh-island. In this section, we'll take a look at

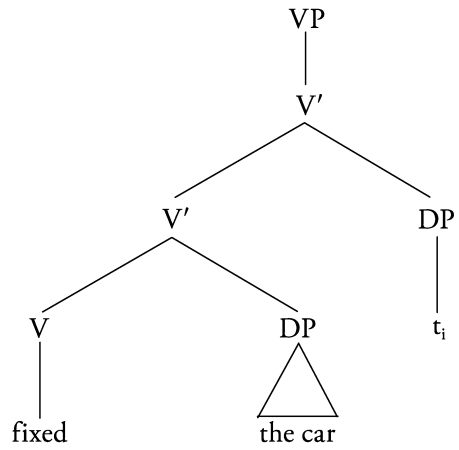
VP-adjuncts and see that they'll require us to modify our assumptions about lexical proper government.

The problem in a nutshell is that, unlike direct objects and subjects, VP-adjuncts don't behave in the same way with respect to these two phenomena. They act like direct objects in that they don't show that-trace effects. However, they act like subjects in that they seem to trigger strong wh-island violations. Consider first (44):

(44) How$_i$ do you think [$_{CP}$ t$_i'$ that [$_{IP}$ John fixed the car t$_i$]]?

The string of words in (44) can be (and probably most naturally is) interpreted as having the structure indicated. That is, it's a question about how John fixed the car, as opposed to a question about how you think. An appropriate answer is 'by replacing the manifold'. So if the structure in (44) is a possible one, then, following the logic that we used in Section 8.8, t$_i$ must be lexically properly governed by *fixed*. Since (44) is grammatical, t$_i$ must satisfy the ECP. Antecedent government is presumably blocked both for minimality reasons (the complementizer *that* intervenes) and because the barrier VP intervenes. That leaves lexical proper government. This doesn't look like a problem, though, since the structure of the VP is (45):

(45)



Fixed m-commands the trace, no barrier intervenes, and there's no closer governor to present a minimality problem.

However, this reasonable-looking conclusion starts to unravel when we consider sentences like (46):

(46) **How$_i$ were you wondering [$_{CP}$ who$_j$ [$_{IP}$ t$_j$ fixed the car t$_i$]]?

(46) seems to represent a strong wh-island violation, on a par with the subject extraction that we saw in (43) above. The only possible interpretation for the string of words in (46) is as a question about how you were wondering. It just can't be interpreted as a question about how someone fixed the car. If you were

to respond to the question in (46) with the answer 'by replacing the manifold', it would sound impossibly odd (because people don't wonder by replacing the manifold). Compare (46) with (44). What in (44) is the most natural interpretation for the string of words is just not possible for (46).

But now we've got a problem. Using the logic that we developed in Section 8.9, the strong ungrammaticality of (46) would lead us to conclude that the ECP (as well as Subjacency of course) is violated. But if VP-adjuncts satisfy the ECP by being lexically properly governed, (46) should be only a weak Subjacency violation like (42). t_i is governed by the verb in (46) just like it is in (42). The relevant structure (see (45)) is identical. (Clearly, there's no problem with t_j. That's antecedent governed by *who* in the expected way.)

So, it seems as though what we *really* need is a three-way distinction between subjects, direct objects and VP-adjuncts rather than just a two-way distinction between subjects and everything else. In particular, there's a difference between direct objects and adjuncts that we're not capturing. As usual, what we need to do is think about what the difference is between the two and see if that can form the basis for a way forward.

θ-marking, Case Marking and lexical proper government

The main difference between direct objects and adjuncts is that the verb doesn't θ-mark adjuncts. In a sentence like *John fixed the car with a wrench*, the verb assigns the Patient θ-role to *the car*, but it's the preposition *with* which assigns the Instrument θ-role to *a wrench*, leading to the interpretation that *a wrench* is the instrument of the fixing. Had the prepositional phrase been *in Valencia* instead of *with a wrench* then *Valencia* would have been assigned the Location θ-role by *in*. If we restrict lexical proper government so that you only lexically properly govern something if you θ-mark it, then we predict that adjuncts must be antecedent governed.

That would explain why (46) is an ECP violation: t_i fails to be antecedent governed. *Who* is in the embedded Spec CP, but it has a different index from t_i and so can't antecedent govern t_i. However, by minimality, it blocks antecedent government from *how*.

So, it looks like introducing a θ-marking requirement into the definition of lexical proper government will explain why (46) is strongly ungrammatical. But before we do anything else we should double-check Exceptional Case Marking contexts. ECM subjects, remember, are Case marked as though they were direct objects of the verb, but are θ-marked as if they were the subjects of the infinitival clause. So tying θ-marking to lexical proper government is going to have the knock-on effect that ECM subjects should *not* be lexically properly governed. Does that look like the right result? Well, no actually. Consider (47):

(47) *Who, were you wondering [$_{CP}$ why$_j$ John believed [$_{IP}$ t_i to be the best candidate] t_j]?

(47) seems to have the vaguely comprehensible feel to it which characterizes violations of Subjacency only, and not the 'word salad' feel of an ECP

violation. t_i must therefore be properly governed. Notice that it couldn't be *antecedent* governed. The nearest antecedent governor for t_i is *why*, but that has the wrong index, and also by minimality blocks antecedent government from anything farther away. Therefore, the subject position of an ECM infinitive must be *lexically* properly governed, which contradicts what we would have predicted.

Fortunately, the solution to this problem isn't too difficult. What we need to do is have lexical proper government depend not just on θ-marking, but on *either* θ-marking *or* Case assignment. In the standard case of direct objects, it won't make any difference because they're both assigned a θ-role and assigned Case. However, in order to correctly account for the fact that ECM subjects also seem to be properly governed, we need to allow for just Case assignment being sufficient. Just to be concrete about everything, let's work this discussion that we've had into a revised definition of *lexical proper government* (I've put the new material in bold):

(38) *The Empty Category Principle*
Traces must be properly governed.

(48) α properly governs β if and only if
(i) α lexically properly governs β, or
(ii) α antecedent governs β.

(49) α lexically properly governs β if and only if
α governs β, where a is a lexical category **and α assigns either Case or a θ-role to β.**

Let's summarize where we are at this point: contrary to what we thought initially, adjuncts satisfy the ECP by being antecedent governed, rather than by being lexically properly governed. We wanted to say that in order to account for the strong ECP ungrammaticality of (46), where an adjunct has been extracted from a wh-island. (Otherwise we would have predicted just a weak Subjacency violation.) But that meant that we needed to modify our definition of lexical proper government. We made sure that adjuncts wouldn't be lexically properly governed by tying lexical proper government to either Case or θ-role assignment. Direct objects and ECM subjects will therefore be lexically properly governed, while adjuncts won't be. However, we've got one loose end that needs to be tied up.

Levels of application and a solution to the problem of adjuncts

We've explained why adjuncts pattern with subjects in triggering strong violations when extracted from a wh-island – both subjects and adjuncts are antecedent governed. And the wh-word in the embedded Spec CP, which creates the wh-island, will always present a minimality barrier to antecedent government. However, what we *haven't* explained is why adjuncts don't trigger that-trace

violations. In other words, in correctly predicting that (46) is ungrammatical, we predict that (50) is also ungrammatical, which it isn't:

(50) How$_i$ do you think [$_{CP}$ t$'_i$ that [$_{IP}$ John fixed the car t$_i$]]?

If adjuncts are antecedent governed like subjects, then the presence of the complementizer *that* should, by minimality, prevent antecedent government of the trace just like it does for subject extraction. In other words, (50) should be just like (51):

(51) *Who$_i$ do you think [$_{CP}$ t$'_i$ that [$_{IP}$ t$_i$ fixed the car]]?

Lasnik and Saito (1984) suggested that the difference between (50) and (51) lies in the fact that subjects are arguments, whereas adjuncts are just that – adjuncts, not arguments. They are not required by the syntax or semantics of the verb. What Lasnik and Saito propose is that the ECP applies at different levels for arguments as opposed to adjuncts. For arguments, the ECP has to be satisfied both at S-Structure and at LF. For adjuncts, on the other hand, the ECP only needs to be satisfied at LF.[6]

How does this help us with the (50)/(51) contrast? Well, as Lasnik and Saito note, one operation that you can carry out between S-Structure and LF is the deletion of the complementizer *that*. Normally, of course, you can't delete things between S-Structure and LF. Otherwise you'd end up predicting that 'John left yesterday at 3 p.m.' can mean just 'John left', with deletion of the rest of the material between S-Structure and LF.

However, complementizer *that* differs from *yesterday at 3 p.m.* in that *that* has no semantic content. It contributes nothing to the interpretation of the sentence, and LF is the level at which those aspects of the syntax which are relevant for interpretation are represented. Since it contributes nothing at LF, there's no information lost if you delete it. (Arguably, one is even *forced* to delete it, if things which aren't relevant for the interpretation *must not* be present.) In this respect, complementizer *that* is similar to the expletive *there* which we discussed in Chapter 7 which also gets deleted between S-Structure and LF.

But if we're allowed (or forced) to delete complementizer *that* between S-Structure and LF, then we get rid of the minimality problem that it presents. If the ECP then checks adjuncts only at LF, we predict that there should be no that-trace problems.

Let's step through it, sticking with (50) as our example. At S-Structure, we have (52):

(52) How$_i$ do you think [$_{CP}$ t$'_i$ that [$_{IP}$ John fixed the car t$_i$]]?

Both t$_i$ and t$'_i$ need to be properly governed in order to satisfy the ECP. Complementizer *that* is blocking t$'_i$ from antecedent governing t$_i$, but at this stage that doesn't matter. Since we're talking about an adjunct chain, the ECP doesn't care about what's going on at S-Structure.

At LF, what we have is (53):

(53) How$_i$ do you think [$_{CP}$ t$'_i$ ø [$_{IP}$ John fixed the car t$_i$]]?

Now we need to check what the proper government situation is. t$'_i$ presents no problem. It's antecedent governed in the usual way by *how* in the matrix Spec CP. t$_i$ also now presents no problem. It is antecedent governed by t$'_i$. Deletion of complementizer *that* has removed it as a potential closer governor and the intermediate trace t$'_i$ is thus able to antecedent govern the original trace t$_i$. Therefore, both traces are properly governed and (52) is therefore correctly predicted to be grammatical.

8.11 Summary of the ECP

So, at last, it looks as though we have the three-way distinction that we needed. Direct objects do not trigger that-trace violations and they result in only a weak Subjacency violation when extracted from inside a wh-island. Subjects, on the other hand, behave in exactly the opposite fashion. They trigger both that-trace violations and strong ECP violations, when extracted from a wh-island. Adjuncts behave in a third, different way. They do not trigger that-trace violations, but they do trigger strong ECP violations when extracted from a wh-island.

We achieved this by assuming that direct objects satisfy the ECP via lexical proper government by the verb. Being arguments, direct objects must satisfy the ECP both at S-Structure and LF, but this raises no problem. Therefore any trace in direct object position will always satisfy the ECP. The presence or absence of a complementizer will not therefore trigger any ECP violations and, in the case of extraction from a wh-island, the violation will always be weak.

Subjects, on the other hand, are not lexically properly governed and are therefore dependent on antecedent government to satisfy the ECP. The presence or absence of the complementizer is therefore crucial, because the complementizer represents a closer governor than the intermediate trace. If the complementizer is present, it will block antecedent government, generating the that-trace effect. The possibility of escaping a that-trace violation through complementizer deletion isn't an option for subjects. Since subjects are arguments, they must satisfy the ECP both at S-Structure and LF, and at S-Structure the complementizer is present. When antecedent government fails for some other reason, for example because the antecedent governor is too far away as a result of one-fell-swoop long-distance movement, then a strong ECP and Subjacency violation is triggered.

The fact that adjuncts behave in certain respects like direct objects and in other respects like subjects is explained by the fact that, like subjects, they satisfy the ECP by being antecedent governed rather than by being lexically properly governed. This means that they too trigger a strong ECP and Subjacency violation when antecedent government fails (as in the case of extraction from a wh-island). However, because they are adjuncts and not arguments, they are able to escape the that-trace violations which subjects trigger. At S-Structure, the complementizer blocks antecedent government from the intermediate trace,

but the ECP does not check adjuncts at S-Structure. By the time LF is reached, complementizer *that* has been deleted and therefore no longer presents a minimality barrier to antecedent government. The trace of the adjunct is therefore properly governed at LF and the ECP is satisfied.

8.12 Relativized minimality: an introduction

In this section we're going to look at an important modification that has been made to the ECP, particularly in order to address the conceptual problem with antecedent government that came up in Section 8.10. The essential problem is that the way the ECP is formulated looks to be a bit of a trick. It says that traces need to be 'properly governed', with the two paths to proper government being lexical proper government or antecedent government. Although we've called these two kinds of government 'proper government' – implying that they're really the same thing – in reality they look quite different. Lexical proper government, on the one hand, does look plausibly like a variety of the kind of government that we saw in Chapter 4. It involves heads and seems sensitive to the same kinds of barriers (especially CP) which block 'regular' government. Antecedent government, on the other hand, really doesn't look like a variety of 'regular' government. The elements that antecedent govern are not heads, but instead maximal projections, and, crucially, you can antecedent govern through maximal projections like CP and VP. This makes it look like we're talking about two different things, and that's being obscured because we're calling them both 'government'.

There's also a question about the nature of antecedent government when we consider in more detail the question of elements which block it. Rizzi (1990) raises an important question which we've swept under the carpet until now: why doesn't the subject block antecedent government of adjunct traces? Consider (54):

(54) *[$_{CP}$ how$_i$ did [$_{IP}$ John say [$_{CP}$ what$_j$ [$_{IP}$ Bill fixed t$_j$ t$_i$]]]]?

Just like in the other cases that we discussed in this chapter, the problematic trace is t$_i$, the trace of *how*. t$_i$ being an adjunct, needs to be antecedent governed. However, the closest governor is *what*, which doesn't have the same index as *how*. Therefore, it both fails to antecedent govern t$_i$ itself and also blocks antecedent government from *how*, which is further away.

But consider the implications of this explanation for a simple sentence like (55):

(55) [$_{CP}$ Why$_i$ did [$_{IP}$ John leave t$_i$]]?

(55) is grammatical, and with respect to the trace t$_i$ we would say that it satisfies the ECP in virtue of being antecedent governed by *why*. (In this case, there's no difference between the S-Structure and the LF representations, although technically, being an adjunct trace, it wouldn't be checked until LF.)

But Rizzi points out that we need to answer a crucial question about (55). Why doesn't the subject *John* block antecedent government by *why*? *John* is a DP which c-commands the trace t_i, and *John* doesn't have the same index as the trace. (It's not clear what it would mean for *John* and an adjunct *why*-phrase to be coindexed anyway.) Therefore, *John* (or any other subject) should, by minimality, always block antecedent government, and we therefore predict that any adjunct wh-movement should be ungrammatical.

It's clear that we don't want this result. But now what's the difference between the ungrammatical (54) and the grammatical (55)? There's an obvious possibility – in (54), the element which blocks antecedent government is a wh-word in Spec CP. In (55), on the other hand, the subject (which we want not to be a blocker) isn't in Spec CP. Subject position isn't even an A′-position, it's an A-position.

Rizzi (1990) sees this as the key difference. He proposes changing the conception of minimality. As we've discussed it thus far, the concept of minimality in the ECP is a *rigid* one. That is to say, if a given element has a closer governor *of any kind*, it blocks government from *any* other governors which are further away. What Rizzi proposes we do is change to a *relative*, rather than a rigid, conception of minimality. We've got different kinds of government – government by heads (lexical proper government) and government by phrases (antecedent government) – and the idea behind relativized minimality is that different kinds of governors are only blocked by closer governors *of the same type*.

According to Rizzi, we actually need to differentiate between not just *two*, but *three* different types of government. First, in order to explain the difference between (54) and (55), we need to break antecedent government down into two separate subcases, depending on whether the potential governor is in an A-position or an A′-position. The basic idea is that we've got two kinds of chains: A-chains and A′-chains (that is, chains formed by movement to an A-position and to A′-positions), and, under the *relativized* minimality approach, antecedent government from one type of position won't be blocked by an element in the other type of position.

In (54), we need to see if the trace t_i is antecedent governed. *How* is the trace's antecedent, but *what* is a closer governor. However, does *what* count as a blocker under relativized minimality? Yes, it does. It is, in Rizzi's terminology, a 'typical potential governor' for the trace. It's in an A′-position – the embedded Spec CP – and it's closer to t_i than *how* is. It therefore blocks antecedent government, and we still correctly predict that (54) violates the ECP.

But look how the situation changes now with respect to (55) under relativized minimality. We again are looking at t_i and we need to see whether antecedent government by *why* is blocked. In this case, the potential intervenor is *John*. However, *John* is in an A-position, not an A′-position. It therefore isn't a 'typical potential governor' for a wh-trace and therefore doesn't block antecedent government from Spec CP. We therefore avoid the prediction that we seemed to be stuck with under rigid minimality, in which *John* should count as a blocker and (55) should be ungrammatical.

Rizzi's relativized approach to minimality can also explain the other question which we raised initially in Section 8.10 and came back to at the beginning of

this section: why does antecedent government seem able to pass right through supposed barriers like CP and VP when lexical proper government can't? Under relativized minimality, it's not that VP and CP are barriers in themselves. Rather it's the fact that CP or VP indicate the presence of a closer head, namely C or V, and those C or V heads will block government by any head which is further away (for example, if we're talking about Case assignment or lexical proper government). On the other hand, what's relevant for antecedent government is XPs (full phrases) in A- or A′-positions, not heads. So antecedent government will only be blocked by XPs in the relevant positions. Therefore, while a C head might block a verb from lexically properly governing a trace in Spec CP (as it would constitute a closer governor), that same trace could be successfully antecedent governed by something in the higher Spec CP (provided of course that there are no other intervening XPs in A′-positions) because antecedent government doesn't care about heads.

So, as you can see just from this introduction, it looks like Rizzi's concept of *relativized*, as opposed to *rigid*, minimality has a lot going for it. However, rather like the ending of *Star Trek II: The Wrath of Khan*, the price of success in the mission is the life of an old friend. In moving from rigid to relativized minimality, we lose the ECP analysis of that-trace effects, which crucially depends on rigid minimality. Consider the familiar contrast in (56):

(56) a. *Which suspect do you think $[_{CP}$ t'_i that $[_{IP}$ t_i committed the murder]]?
b. Which suspect do you think $[_{CP}$ t'_i $[_{IP}$ t_i committed the murder]]?

Since t_i is the trace of a subject, and therefore the trace of an argument, its status with respect to the ECP is checked at S-Structure. At this level, though, the intervention of the complementizer blocks antecedent government and (56a) is ungrammatical. Under relativized minimality, however, it's not clear why complementizer *that*, which is a head, should be a blocker of antecedent government by a wh-trace in the embedded Spec CP. Being a head, we expect it to block lexical proper government, but not antecedent government. Rizzi is not unaware of this problem, however, and devotes a significant portion of his (1990) book to coming up with a new analysis of that-trace effects compatible with relativized minimality.

Summary

In this chapter, we've reopened the books on traces and found that traces aren't a separate kind of DP from the anaphors, pronouns and R-expressions that we saw in Chapter 5. Instead, following Chomsky (1982), what we've argued is that traces are just a part of a wider group of empty categories, which include our old friend PRO from Chapter 4, and a new kind of empty category, *pro*, which doesn't exist in English, but which is found in null subject languages like Spanish and Italian.

We then discussed an important principle which constrains movement and therefore the distribution of empty categories, which we called the *Empty Category Principle*.

(38) *The Empty Category Principle*
Traces must be *properly* governed.

Consideration of that-trace violations and the existence of weak vs strong wh-islands led us to the conclusion that there were two paths to proper government: lexical proper government (for direct objects and ECM objects) and antecedent government (for subjects):

(48) α properly governs β if and only if
(i) α lexically properly governs β
(ii) α antecedent governs β.

(49) α lexically properly governs β if and only if
(i) α governs β, where α is a lexical category and α assigns either Case or
(ii) a θ-role to β.

(46) α antecedent governs β if and only if
(i) α c-commands β,
(ii) α and β are coindexed, and
(iii) minimality is respected.

The ECP predicts a difference between subjects, which are antecedent governed, and objects which are lexically properly governed. However, in order to explain the behaviour of adjuncts, which are antecedent governed, but pattern with direct objects in not showing that-trace effects, we needed to assume, following Lasnik and Saito (1984), that the ECP applied at different levels, depending on whether the trace in question is in an A-position or an A′-position:

(52) *Levels of Application of the ECP*
The ECP applies:
(a) at S-Structure to traces in A-positions;
(b) at LF to traces in A′-positions.

Open issue: does *pro* exist in english?

Throughout this chapter, we've been assuming that the feature combination [–a, +p], that is, *pro*, was automatically excluded in a non-null subject language like English. It seemed pretty straightforward. We assumed that *pro* wasn't possible in English because of the ungrammaticality of sentences like (57), which are perfect in languages like Italian, Spanish or Turkish.

(57) *is going to the store.

However, Epstein (1984) presents some very interesting evidence to suggest that English *does* have *pro*, just not in subject position.

Consider (58), a sentence which up until now we might have analysed as containing arbitrary PRO:

(58) It is fun [$_{CP}$ [$_{IP}$ PRO$_{arb}$ to play baseball]].

PRO, the subject of *play* is interpreted in (58) as 'anyone' or 'people in general', which is exactly the interpretation that we've seen for other cases of arbitrary PRO, like (59):

(59) PRO$_{arb}$ to return early from vacation is unpleasant.

Given that interpretation, it would seem clear that PRO$_{arb}$ should be identified as some kind of universal quantifier and should therefore undergo Quantifier Raising just like other quantifiers that we saw in Chapter 7.

However, there's a problem. We saw in Chapter 7 that QR was strictly clause bounded. That is, the quantifier can be adjoined at LF only to the 'nearest' IP node. That would seem to conflict with the interpretational facts in (58), in which it appears that QR is not clause bounded.

To see this, we need to take a closer look at what (58) means and the implications that the meaning has for what the structure is at Logical Form. All other things being equal, we'd expect (58) to have the (partial) LF in (60), in which the subject, PRO$_{arb}$, has been QRed to the closest IP that dominates it:

(60) [$_{CP}$ [$_{IP}$ it is fun [$_{CP}$ [$_{IP}$ ($\forall$x) [$_{IP}$ x to play baseball]]]]]

Unfortunately, this LF is of a sentence which asserts (61):

(61) If everyone plays baseball, it is fun.

This is not what (58) means. What (58) means is that, for any person, it is fun for that person to play baseball. *That* meaning would be associated with the LF in (62), in which the universal quantifier has scope over the matrix clause:[7]

(62) [$_{CP}$ ($\forall$x) [$_{IP}$ it is fun [$_{CP}$ [$_{IP}$ x to play baseball]]]]

This is a problem. The only way to derive this LF at the moment is to say that QR is always clause-bounded, except when we're dealing with PRO$_{arb}$.

And things just get worse, as Epstein notes. It's not just that QR is not clause bounded when dealing with PRO$_{arb}$. It looks like PRO$_{arb}$ can be QRed exactly one clause higher and no further. Consider (63):

(63) Josh said it is fun PRO to play baseball.

PRO can have a controlled interpretation here (where Josh asserted that it is fun *for him* to play baseball), but let's focus on the PRO_{arb} interpretation, where Josh is trying to assert something general. What we see is that the LF associated with (63) has to be (64), and not (65) or (66):

(64) $[_{CP}\ [_{IP}$ Josh said $[_{CP}\ (\forall x)\ [_{IP}$ it is fun $[_{CP}\ [_{IP}$ x to play baseball]]]]]].

(65) $[_{CP}\ [_{IP}$ Josh said $[_{CP}\ [_{IP}$ it is fun $[_{CP}\ [_{IP}\ (\forall x)\ [_{IP}$ x to play baseball]]]]]]].

(66) $[_{CP}\ (\forall x)\ [_{IP}$ Josh said $[_{CP}\ [_{IP}$ it is fun $[_{CP}\ [_{IP}$ x to play baseball]]]]]].

Just as with (60) above, (65) is the LF of a statement which asserts that Josh said 'if everyone plays baseball, it is fun', which is clearly not what (63) means. (66) on the other hand asserts not that Josh made some universal claim, but rather that Josh made a claim or claims which just happen to be true of everyone. For example, suppose that the only people in the universe are Josh and Erich. If Josh at one point says 'It is fun for Erich to play baseball' and then later says 'It is fun for me to play baseball', then (66) would be true. But that doesn't seem like a circumstance under which we would want to say that the sentence in (63) was true. Our intuition is that (63) is only true if Josh said 'It is fun [for everyone] to play baseball'. *That's* the interpretation in (64), in which PRO_{arb} is QRed *exactly one clause higher* than that one it's in, and no more. This is a very, very, strange state of affairs indeed.

However, Epstein notices that there's another aspect to the interpretation of (58) which we haven't considered yet, and therein lies a possible way out. Notice that you can have an expressed DP in the place of PRO_{arb}:

(67) It is fun $[_{CP}$ for $[_{IP}$ Babe Ruth to play baseball]].

This is a little hard to see, for reasons which will become clear momentarily, but (67) can mean something like 'it is fun for everyone for Babe Ruth to play baseball'. It gets even better if you make the context a little more plausible, something like *it is good business sense for Babe Ruth to play baseball.* However, notice that we seem to have an extra argument there – a complement of the adjective *fun* – as indicated by the translation of (67). We can even put that second argument in by adding another prepositional phrase:

(68) It is fun for the spectators $[_{CP}$ for $[_{IP}$ Babe Ruth to play baseball]].

Therefore, going back to (58), it looks like there's not just PRO_{arb}, but also a *second* unexpressed argument which is the complement of *fun* and which is coindexed with PRO. What (58) really means is 'for any x: if x plays baseball, it is fun *for x*'. (58) doesn't mean that 'for any x: if x plays baseball, then it is fun for y' (where y is maybe some sadist who likes watching other people run around). This suggests that the logical form that we gave for (58) was wrong. It's not (62), but rather (69), with a variable in the matrix clause and a *controlled* PRO:

(69) $[_{CP}\ [_{IP}\ (\forall x)\ [_{IP}$ it is fun (for) $x_i\ [_{CP}\ [_{IP}\ PRO_i$ to play baseball]]]]]

This suggests that the S-Structure of (58) is really (70), where e_i is a null universal quantifier which undergoes QR:

(70) $[_{CP}\ [_{IP}$ it is fun $e_i\ [_{CP}\ [_{IP}\ PRO_i$ to play baseball]]]]

This would seem to solve all of our problems about scope. If it's not PRO_{arb}, but rather some null complement of *fun* which is the universal quantifier which undergoes QR, we explain immediately the interpretational facts of (63) above. No modifications are necessary at all. The universal quantifier has exactly the scope it has because it's in the same clause as *fun* and is moving to take scope over the minimal IP in which it appears.

The question now is what kind of empty category e_i can be. Since e_i itself has no antecedent, [+a, –p] is ruled out. Assuming that e_i is in the same position that the benefactive phrase would be located (as in (68)), it would be governed by *fun*, which rules out [+a, +p]. [–a, –p] is a possibility, but that would result in a variable that has no c-commanding operator (e_i does need to *create* a variable when it undergoes QR, but remember we're looking into what kind of empty category e itself is). That just about leaves [–a, +p] as the only choice.

However, what are we now going to do about sentences like (57), which seem to show that null pronouns aren't allowed in English? Epstein offers a solution, which, conveniently enough, also accounts for the ungrammaticality of (71):

(71) *It is fun for *pro* to play baseball.

He proposes that a crucial feature of *pro* in English is that it cannot be Case marked. This way, *pro* in English will be correctly barred from subject position of a tensed clause (like in (57)), as that will always receive Nominative Case from I. In (71), *pro* will be inappropriately Case marked by the preposition *for*.

Epstein's proposal is extremely interesting, although it does come slightly at a cost, not the least of which is the weakening of the link between *pro* and overt pronominals. If the four featural possibilities for empty categories are truly nothing more than unpronounced versions of their overt counterparts, then we would expect not only that *pro can* bear Case, but that it *must*, in line with other overt pronominals. This seems to be true for *pro* in true null subject languages, but would need to be modified for English.

Also, as Epstein himself notes, there are an additional class of cases which seem to be interpreted in the same way as (58) and (63), but where there is no syntactically expressible argument position:

(72) John knows $[_{CP}$ how $[_{IP}$ PRO to solve the problem]].

(72) seems parallel to the cases we've seen above. The scope of the universal quantifier seems to be the matrix IP, rather than just the embedded IP.

(73) $[_{CP}\ [_{IP}\ (\forall x_i)\ [_{IP}\ John_j$ knows $[_{CP}$ how $[_{IP}\ PRO_i$ to solve the problem]]]]].

Unlike *fun*, however, *know* does not have a position in its argument structure where a null *pro* might be lurking:

(74) *John knows for/to Bill how PRO to solve the problem.

For these reasons, we'll have to take Epstein's arguments as tentative, rather than conclusive. However, in Chapter 10 we'll see another context where it's been claimed that *pro* exists in English, even in subject position.

FURTHER READING

The typology of empty categories as I've laid it out in this chapter is originally due to Chomsky's (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, but there are antecedents in Chomsky (1981). It's in the (1981) book, for example, that Chomsky sets out the PRO Theorem, although *pro* doesn't get invented until *Concepts and Consequences*. (Chomsky (1981) says that PRO is the unexpressed subject of null subject languages.) Section 5 of Harbert (1995) also has a discussion of these issues at a slightly more advanced level but with lots of interesting additional data. The ECP is probably the topic with more written about it than any other in Chomskyian generative grammar. There are all manner of permutations and formulations, and what we've presented here is only a basic, vanilla version. Like a lot of things, the original proposal can be found in Chomsky (1981). One of the first big revisions/expansions of the idea was mentioned back in Section 8.10: Lasnik and Saito (1984). It has a (deserved) reputation for being extremely difficult to read, though, and I definitely wouldn't recommend that you start there. I'd start with Lasnik and Uriagereka (1988), which is an intermediate/advanced textbook which covers the ECP. There's also some excellent discussion in Webelhuth (1995a), which you should be able to follow at this stage. Both Lasnik and Uriagereka and Webelhuth have a lot of references to the primary literature which you can follow up on.

ANSWERS TO IN-TEXT EXERCISES

Exercise 1

What we need to do is find the governing category for the trace and then see where *John* falls inside it or outside it. The element is the trace, its governor is the past participle *killed*, but, going up the tree, the first category which has a subject is IP. The subject is accessible and therefore the governing category for the trace is IP. *John* falls within IP and therefore the trace of Passive must be anaphoric and not pronominal [+a, –p]. Any other combination would result in either Principle A or Principle B being violated.

Exercise 2

This is very similar to the previous question, actually. The governing category for t_i is the IP *Mary likes* t_i. The element is t_i, its governor is *likes*, and the first category which has a subject is the IP *Mary likes* t_i, Coindexation between *Mary*; and t_i does not create an i-within-i violation. Therefore *Mary* is an accessible subject for the trace and *Mary likes* t_i is the governing category for the trace. *He* is outside of that IP and therefore the trace is not bound in its governing category.

FURTHER EXERCISES

Exercise 1

Give the S-Structure and LF representations for the following sentences, being sure to insert all appropriate traces. (You may use labelled brackets rather than trees.) Show how the ECP accounts for the indicated judgement.

1. What did John buy?
2. I know who John likes.
3. Everyone believes John to be a fool.
4. Who does John believe to be a fool?
5. John seems to be here.
6. How did John see the man who bought the book? (Answer: with a telescope.)
7. *How did John see the man who bought the book? (Answer: with a credit card.)
8. Who seems to be seen every day?

Exercise 2

Do the extracted objects of prepositions, as in 1, satisfy the ECP by being lexically properly governed or antecedent governed?

1. Who_i did John give the book to t_i?

What does that predict about the following sentences? Is that prediction confirmed?

2. Who did John say that Bill gave the book to?
3. Who did John wonder whether Bill gave the book to?

What's the problem posed by the following sentence?

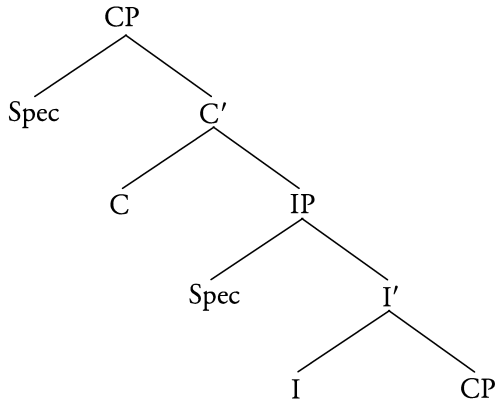
4. *Who did you get angry after Bill showed the photograph to?

Towards the Minimalist Program: Clause Structure Revisited

9

Since introducing phrase structure and X′-Theory in Chapter 2, we haven't seen any reason to modify the basic structure of the clause. As it stands, in addition to whatever *lexical* categories are needed (VP, DP, PP, etc.), the clause contains the *functional* categories IP and CP, giving us the basic structure of the clause in (1):

(1)

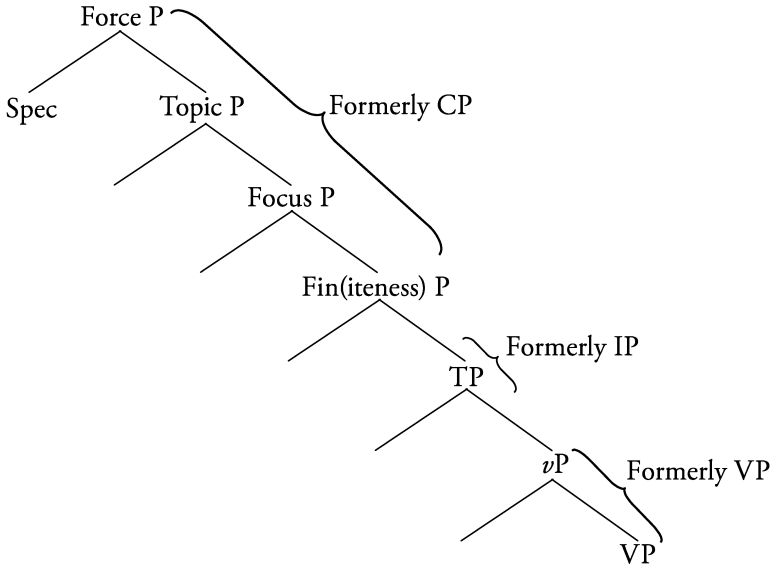


What we'll do in this chapter is revise the picture in (1), working our way upward through the clause, concentrating on three main areas: the VP area, the IP area and the CP area, with a view towards moving to a system which more closely approximates that of Chomsky's more recent framework, the Minimalist Program (Chomsky 1995, 2001; Hornstein et al. 2005; Bošković and Lasnik 2007; among many others). The Minimalist Program in reality consists of a very radical revision of quite a number of basic assumptions about structure building, and it would require an entirely separate book to explain those properly. However, we can introduce at least *some* of these ideas in a basic

way with the syntax that we have so far. That way, for those of you who continue on studying syntax, at least some of the concepts and phenomena won't be totally unfamiliar.

In reviewing the VP area, the first thing we'll do is argue for what's called the VP-internal Subject Hypothesis. As the name suggests, the hypothesis claims that subjects are base-generated within VP at D-structure and then *move* to the specifier of IP (rather than just being in the specifier of IP all along). We'll then go on to argue that we need an additional functional category within VP, called *v*P (with an italicized lower case v and read as 'little v P'). We'll reanalyse Accusative Case assignment as movement of the direct object to a specifier of *v*P, and in doing so introduce the idea that categories can have more than one specifier. IP will remain more or less as it is, but will change its name from IP to TP, to signal more clearly that its main role in the Minimalist Program is with respect to tense, rather than agreement. Finally we'll take a look at CP, and discuss a suggestion that goes back to Rizzi (1997) that CP is not a single functional projection, but rather should be expanded into a series of them. So, by the time we finish, the clause will look like (2) rather than (1):

(2)



9.1 The VP-internal subject hypothesis

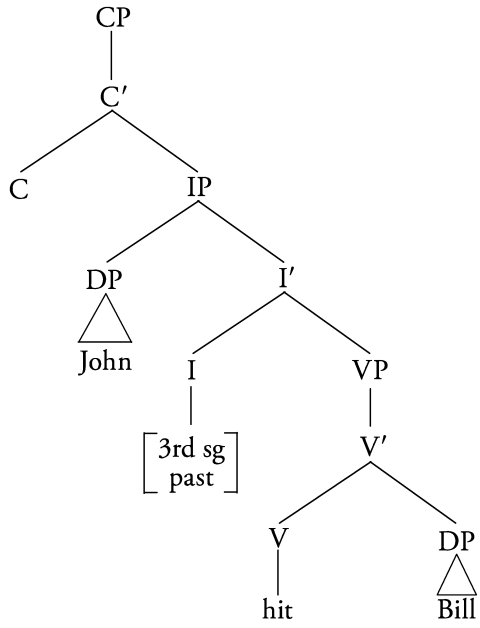
In this section, we'll look at some arguments for the idea that subjects are base-generated within the VP at D-structure. We'll start by looking at some problems that our current view has regarding θ-role assignment to subjects, and then move on to some empirical considerations to do with the position of quantifiers and Binding Theory.

θ-role assignment to subjects

Ever since our first discussions about the structure of the clause back in Chapter 3, we've assumed that the subject was base-generated in the specifier of IP, so a sentence like (3) has the D-structure tree in (4):

(3) John hit Bill

(4)



But this view actually raises a number of problems. One of them has to do with θ-role assignment to the subject. After the initial discussion in Chapter 4, we've been a little vague about the configuration under which θ-role assignment takes place, and that's been sort of deliberate. The subject is an argument of *hit* and is assigned the Agent θ-role by *hit*. But how exactly does that happen? The subject is neither governed (m-commanded) nor c-commanded by *hit*, nor do the subject DP and V seem to be in any other syntactic relation that we've seen.

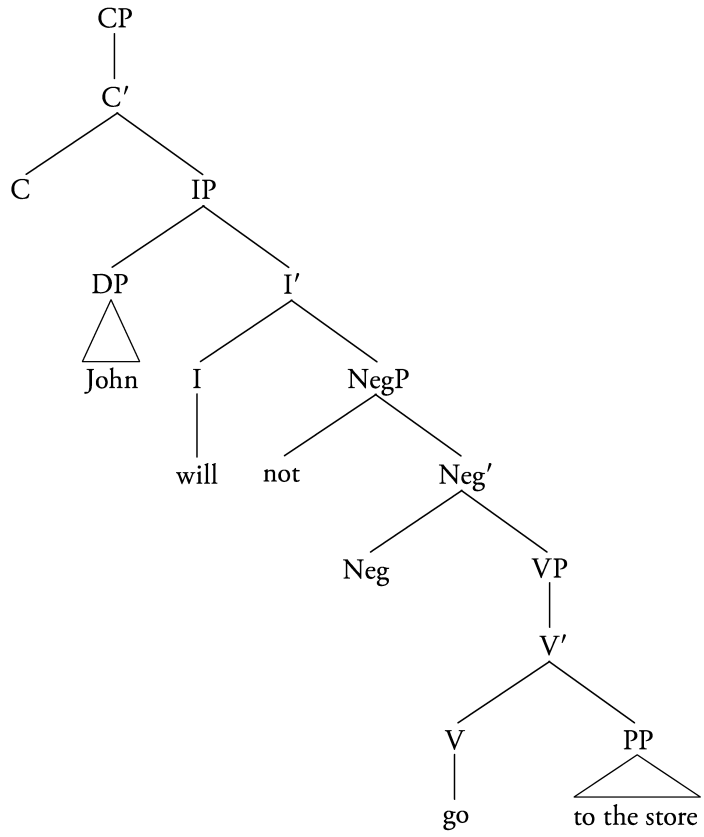
One approach to this problem might be to claim that it isn't V that assigns the θ-role to the subject, but rather the whole VP. V might not m-command the subject, but *VP* does, and that might be a way to bring θ-role assignment under government. However, the idea that it's the VP which assigns the subject θ-role to the subject is complicated by simple negated sentences such as (5):

(5) John will not go to the store

The usual assumption for English is that sentential negation has its own functional projection between IP and VP, and that *not* is the specifier of NegP

(though for the present discussion it doesn't matter whether it's the head or the specifier):

(6)



This seems like a reasonable thing to say based on the word order in (5). And if we look at how *not* patterns with multiple auxiliaries, it appears to follow consistently the first one:

(7) a. John will not be going to the store.
b. John has not been going to the store for years.
c. John would not have gone to the store unless...

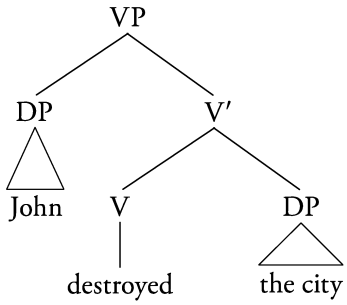
However, if NegP is between VP and IP, then the VP (a projection of the verb) no longer governs the subject (since it no longer m-commands it) and the connection between the verb and the subject (the subject is supposedly an argument of the verb after all) becomes more tenuous.

There are, of course, ways that you could imagine getting around this problem. Because NegP is a sister to VP, you might claim that θ-role assignment somehow 'percolates' up through NegP, and then it's NegP which assigns the

θ-role to the subject. But that doesn't look right from a semantic point of view. There doesn't seem to be any effect that sentential negation has on the θ-role of *John* per se.

But there's a way we can come at the whole problem from another direction. If the subject were actually *inside* the VP at D-structure, say, in the specifier of VP, that would solve the problem. The verb now governs the subject, and in fact governs all of its arguments.

(8)



The subject would be m-commanded by V, just as we would expect. Fortunately, placing the subject within the VP at D-structure also looks like empirically the right thing to do.

Quantifier float

One of the oldest and most straightforward empirical arguments in favour of the VP-internal subject hypothesis has to do with 'floating' quantifiers. These are quantifiers such as 'each', 'both' or 'all', as illustrated in (9)–(11):

(9) Each boy/each of the boys left a book on the table.

(10) Both girls/Both of the girls bought a car.

(11) All (of) the stores closed early.

What's interesting about these quantifiers is that the DPs that the quantifiers are modifying can 'float' away, leaving the quantifier 'stranded' in a lower position. (Don't worry about the slight meaning difference in some cases. We're just interested in the position.)

(12) *The boys each* left a book on the table.

(13) *The girls both* bought a car.

(14) *The stores all* closed early.

It's not simply that the quantifier and the DP have been inverted. If we add in an auxiliary verb, we can see that the quantifier when it's in the 'lower' position looks to be below IP. It follows both auxiliaries and sentential negation, but precedes the verb:

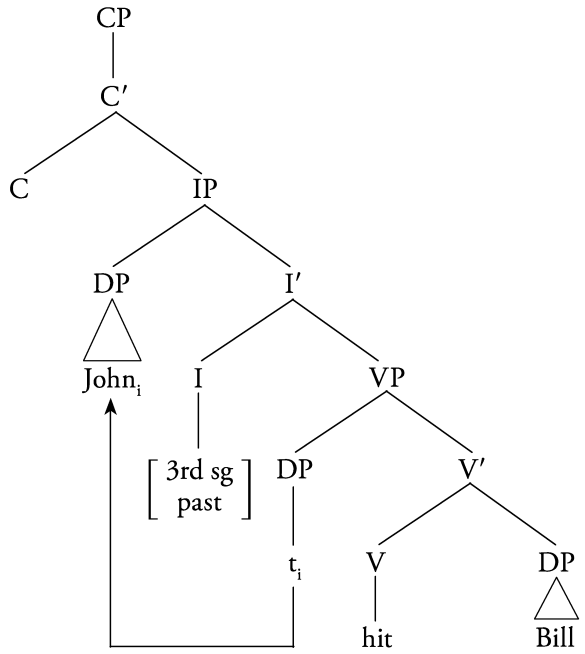
(15) *The children* will *all/both/each* buy a book.

(16) *The children* will not *all/both/each* go to the store.

This really looks like the 'stranded' quantifier is in the specifier of VP. It must precede the verb, but follow auxiliaries and sentential negation.

So, what are the implications for the syntactic derivation? Since the Proper Binding Condition rules out downward movement, it must be the case that subject DPs are base-generated in Spec VP at D-Structure and then move upward to Spec IP to satisfy the Extended Projection Principle.

(17)



So *John* is governed by the verb at D-structure and is a sister to (and therefore semantically modified by) the V′ containing the verb and the direct object. What's unusual about *each*, *both* and *all* is that they allow part of the DP to remain behind in the D-Structure position, which is normally impossible. Determiner plus noun combinations can't 'float away' from a PP or a CP, for example.

(18) a. *The book could [t about linguistics] arrive tomorrow.
b. *The government will [t of France] withdraw its ambassador.
c. *The man has [t that I met yesterday] seen the film.

In addition to the facts in (15) and (16), there's some potentially interesting confirming data from the behaviour of stranded quantifiers in ECM constructions. Consider (19):

(19) *The children* are believed to have *all/both/each* gone to the store.

Here the stranded quantifier can appear in what looks like the specifier of the lower VP, which provides another argument for the idea that ECM objects actually do originate in the lower clause and then undergo A-movement to the higher clause (and, in the case of (19), then move to subject position of the higher clause because *believe* is passive).[1]

Another advantage to the VP-internal Subject Hypothesis is that it explains agreement facts which otherwise would be strange. Romance languages like Spanish show person and number agreement between the stranded quantifier and the floated subject:

(20) *Los niños/Las niñas* están *todos/todas* en el jardín
the boys/the girls are all.MP/all.FP in the garden
'The boys/The girls are all in the garden.'

The floated DP and the stranded quantifier agree in number and gender, even though they're separated by the finite verb. If the quantifier and the rest of the DP are base-generated in different places, then some story about the 'long-distance' agreement needs to be provided. On the other hand, if the quantifier and the rest of the subject are base-generated together in Spec VP, then the agreement is no different from cases like (21):

(21) *Todos los niños/Todas las niñas* están en el jardín.

9.2 More on the VP: double object verbs, VP-shells and light verbs

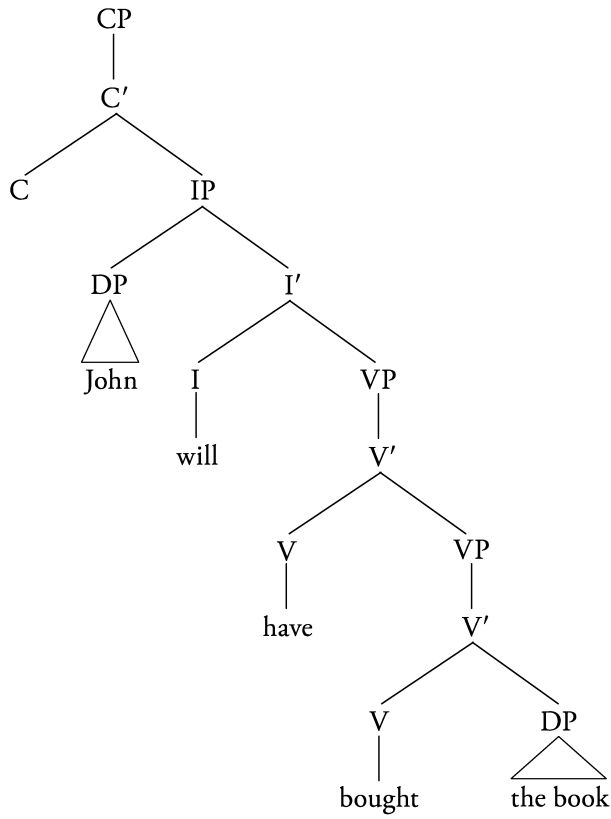
In the previous section, we argued that subjects are actually base-generated at D-structure within the VP. Now what I want to do is (finally!) redeem the promissory note from the Open Issue of Chapter 2 about double object verbs, like *give*.

(22) a. I gave John the book
b. I gave the book to John

What we'll argue is that when a verb has more than one internal argument, we have a VP-shell structure; that is to say, the structure that we introduced back in Chapter 3 to account for sentences with multiple auxiliaries, in which a V takes a VP as its complement (ignoring the VP-internal Subject Hypothesis for a moment).

(23) John will have bought the book

(24)

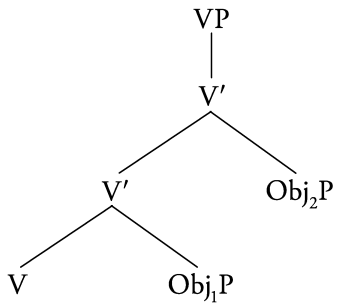


However, giving a VP-shell analysis of double-object verbs isn't quite as straightforward as giving one for multiple auxiliary sentences because there's a key difference which I'm sure you've spotted: (23) has two lexical V elements in the VP (*have* and *bought*), so it's not hard to imagine that there might be two VPs. (22) has only *give* as the lexical verb. What we'll argue is that there's actually an 'extra' phonologically null verb, called a 'light' verb, in the VP in (22). This might seem really ad hoc: we want to have a second verb and so an invisible one magically appears out of nothing. But not only are there some good arguments for it in the case of double object verbs, but we'll also see that having it around for transitive clauses in general does a lot of work for us.

The problem of double object verbs

As we noted back in the Open Issue for Chapter 2 (without really doing anything about it), double object constructions are a bit of a problem for us at the moment. If we're restricting ourselves to binary branching (for acquisition reasons), then a double-object VP in a sentence like (22) pretty much has to have the structure in (25):

(25)



Unfortunately, though, this structure doesn't look right for at least a couple of reasons.

The first one has to do with idioms. The standard assumption is that idioms need to be a constituent. With respect to constituency in the VP, the tree in (25) clearly makes three claims. First, the verb and the *first* object together form a constituent. (They are exhaustively dominated by V′.) Second, the verb and *both* objects form a constituent. (They are also exhaustively dominated by a V′.) Third, and crucially, the verb and the second object *don't* form a constituent to the exclusion of the first object. There is no node in the tree which dominates only the verb and the second object.

So what we predict is that there should be no double-object verb idioms where the idiom consists of the verb plus the second object, without including the first object. But that prediction looks wrong, as Larson (1988) observes:

(26) a. John *carries* such behaviour *to extremes*
b. Bill *threw* John *to the wolves*
c. Albert *took* Scott *to the cleaners*

The sentences in (26), containing idioms made up of the verb plus the second object, suggest that the verb and the second object *are* a constituent. So that's problem number one with the tree in (25).

The second problem is not about constituency, but about c-command relations. The structure in (25) makes a very clear prediction: the second object asymmetrically c-commands the first object. The first branching node dominating Obj_2P is VP, which dominates Obj_1P. But the first branching node dominating Obj_1P is V′, which does not dominate Obj_2P.

But this prediction seems not to be borne out by the facts. As various people started noticing in the mid-1980s (e.g. Barss and Lasnik 1986), a number of phenomena which refer to c-command seem to indicate that the c-command relations are the *reverse* of what we predict; that is, that it is the *first* object that asymmetrically c-commands the *second* object.

Consider first Binding Theory and some data that we mentioned back in the Open Issue for Chapter 2 – but before we'd talked about Binding Theory per se. Principle A of the Binding Theory says that anaphors must be bound by their antecedent (in their governing category) and that binding requires c-command. So if we look at various possible configurations for anaphors and antecedents, we should get some indication as to what the c-command relations are.

(27) a. I showed Mary herself (in the mirror).
b *I showed herself Mary (in the mirror).

(28) a. I showed Mary to herself (in the mirror).
b. *I showed herself to Mary (in the mirror).

Regardless of whether the double object is expressed as *verb X Y* or as *verb Y to X*, it looks as though the first object has to be the antecedent and the second object has to be the reflexive. This makes it look like it's the *first* object that c-commands the *second* one, not the other way around. This is the reverse of what we'd expect, given the tree in (25).

Negative polarity is another phenomenon for which c-command is relevant. Recall that negative polarity items are items like *any*, which are licensed when they're c-commanded by another negative expression or by sentential negation. (So negative polarity *any* is crucially different from the so-called 'free choice' *any*, as in 'Any book will do.') Negative polarity *any* is illustrated by the sentences in (29):

(29) a. I did not buy any books.
b. No tourists went to any museums.
c. *Any book didn't remain on the shelf.

Any books is c-commanded by sentential negation in (29a) and *any museums* is c-commanded by the negative phrase *no tourists* in subject position in (29b). However, since sentential negation does not c-command the subject position, you can't have a negative polarity item there in (29c).

So, again, if negative polarity items require a c-command relationship with the negative element which licenses them, then we should be able to use it to get a handle on the c-command relationships between the two objects in a double object construction. Unfortunately, though, we get exactly the same result as we did with Binding Theory:

(30) a. I gave no gift to any politician
b. *I gave any gift to no politician

(31) a. I gave no politician any gift
b. *I gave any politician no gift

The negative element has to be the first object and the negative polarity item has to be the second object. That suggests that the first object c-commands the second object, but not vice versa, which, again, is the reverse of what we predict.

Finally, let's look at how bound variables behave in the context of double objects. Remember from Chapter 7 that for a pronoun to act as a bound variable it needs to be c-commanded by the relevant operator:

(32) a. Every mad scientist$_i$ thinks that he$_i$ is smart.
b. *He$_i$ thinks that every mad scientist$_i$ is smart.

In (32a), the pronoun *he* is c-commanded by the quantifier, and so can take on a bound variable interpretation. In (32b) on the other hand, the pronoun is not, and it can't.

No prizes for guessing what happens when we look at double object constructions. (Don't forget that we're focusing on the *bound variable* interpretation. (33b) and (34b) are both fine on a non-bound variable reading for the pronoun, where *his* and *its* refers to some single, previously mentioned person or thing.)

(33) a. I sent [every prize]$_i$ to its$_i$ recipient.
b. *I sent his$_i$ prize to [every recipient]$_i$.

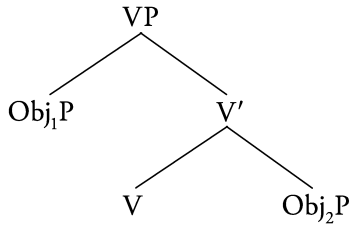
(34) a. I sent [every recipient] his prize.
b. *I sent its$_i$ recipient [every prize]$_i$.

It's the same result again. The pronoun is fine as a bound variable when it's in the second object, but can't be one when it's in the first object. So that suggests again that the first object asymmetrically c-commands the second object.

VP-shells

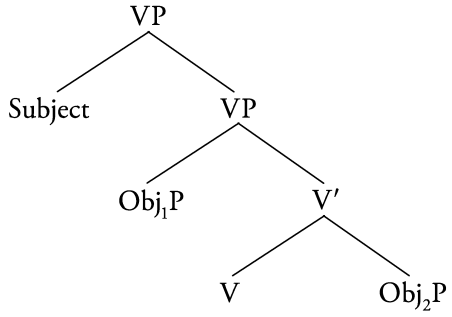
So, what we need is a new structure for double-object VPs which captures the two facts we've discussed above: the fact that the verb and the second object need to be a constituent at some stage, and the fact that the first object needs to c-command the second object. There's an 'obvious' way to do it:

(35)



This gets the facts right, but seems to conflict with the VP-internal Subject Hypothesis. It can't just be that the subject is above the two objects, as some kind of second specifier:

(36)



This would predict that, after the subject raises from Spec VP to Spec IP, you should get the following order:

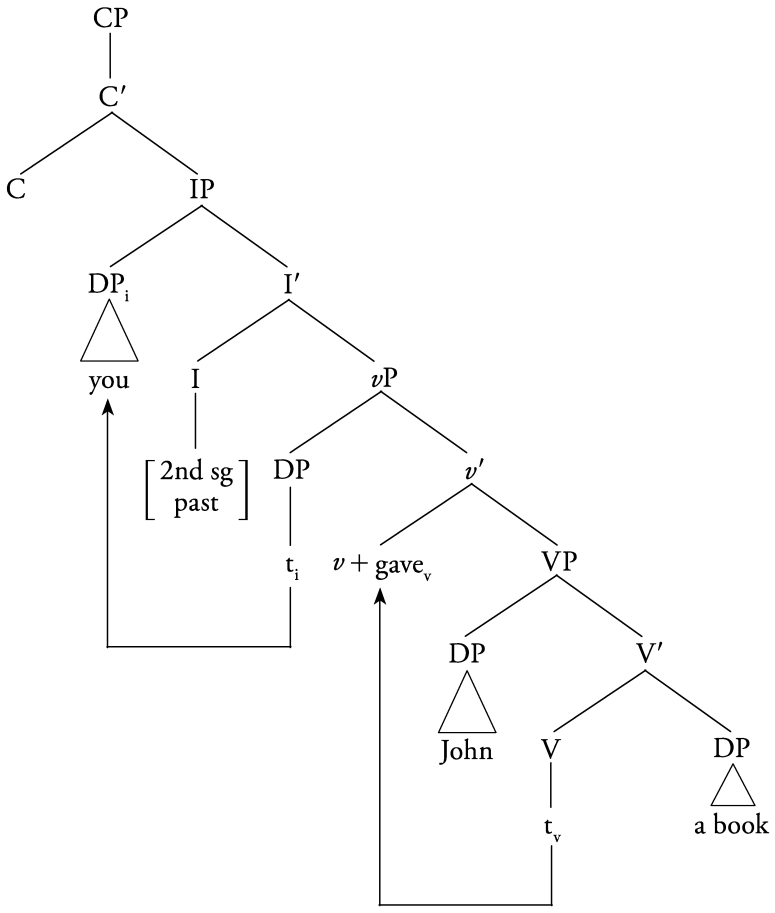
(37) *I every prize sent to its recipient

However, if we *did* have a structure somehow like (36) above, and the verb raised to some higher position, that would get the words back in the right order.

Going back to Larson (1988), the assumption is that this higher position is another VP, which dominates the VP that we've given in (36). We thus have what's called a 'VP-shell' structure, where one VP is layered inside another, just as we did when we talked about multiple auxiliaries in Chapter 3. So a sentence like (38) has, at D-structure, the VP in (39), with a lower VP headed by the lexical verb *give* and the higher VP headed by a 'light' verb, which Chomsky (1995) calls *v* (an italicized lowercase 'v' and read as 'little v'):

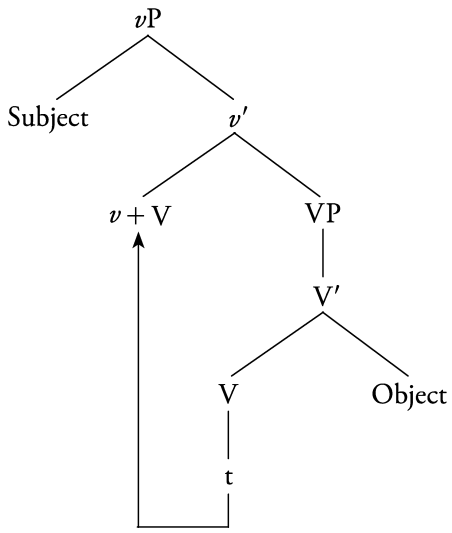
(38) You gave John a book

(39)



At D-structure, *gave* is a sister to its direct object *a book*, but because *John* is in the specifier of the lower VP, it asymmetrically c-commands *a book*, giving us the right c-command facts. Then *gave* raises from V to the light verb *v* (and the subject raises to the specifier of IP). If the verb has only one argument, it is base generated as a sister to the verb, as illustrated by (40).

(40)



Let's now explore just a little bit more with respect to this 'light' verb *v*. Traditionally, the term 'light verb' is used to refer to verbs which have very little semantic content. Often they are used together with nouns to create predicates whose meaning differs little from the noun itself. The semantic contribution of the verb *do* in (41a), for example, is minimal at best.

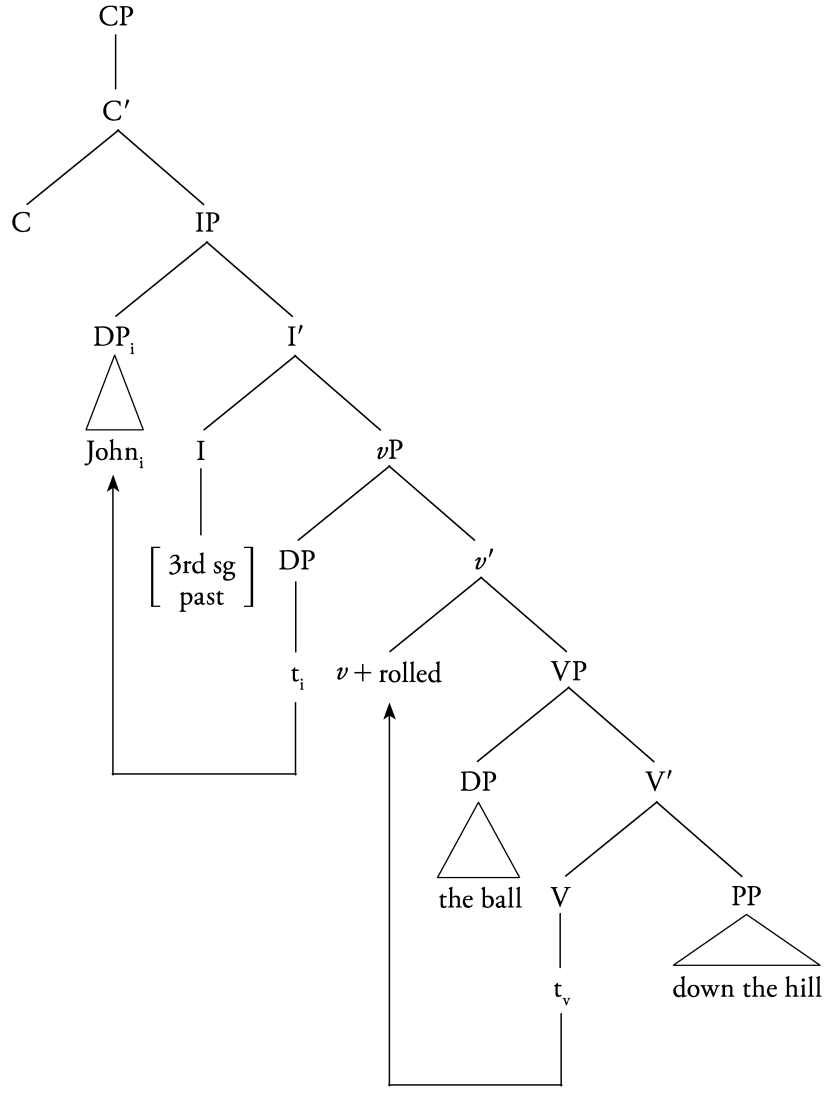
(41) a. Television does harm to children
b. Television harms children

Following Hale and Keyser (1993, 2002) and Chomsky (1995), among many others, the assumption is that the light verb *v* contributes the meaning of 'cause' or 'make' and, as such, is crucially involved in assigning the Agent θ-role to the subject. The VP complement of *v* then expresses the result of the action of causation. You can see this clearly with a lot of verbs in English that undergo the so-called *causative/inchoative* alternation:

(42) a. John rolled the ball down the hill
b. The ball rolled down the hill

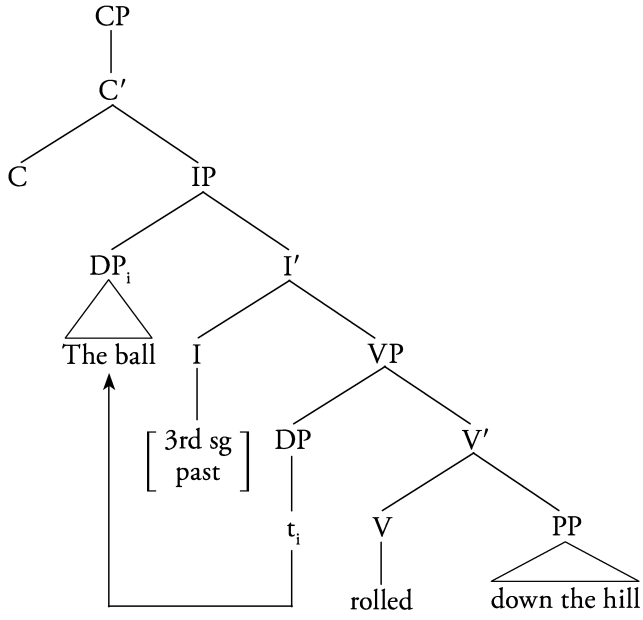
(42a) contains a VP shell, in which *v* takes a VP complement. *John* is base-generated in the specifier of *v*P and is assigned the Agent θ-role, *the ball* is the specifier of the complement VP, and the PP *down the hill* is the complement of V:

(43)



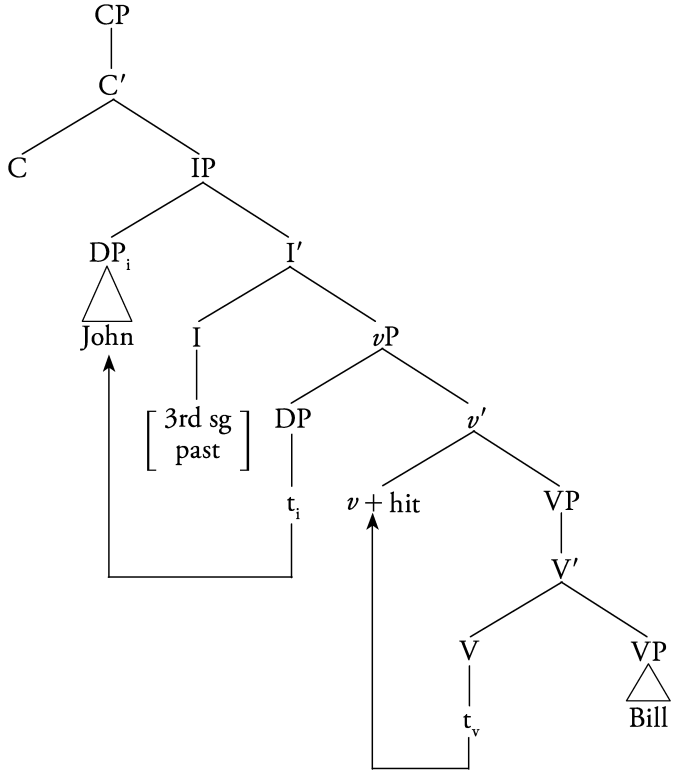
In (42b) on the other hand, *v*P is not present, and there is no agentivity. The surface subject, *the ball*, is base-generated as the specifier of VP and is assigned the Theme θ-role.

(44)



Where the verb does have a causative agent, but only one object, the agent is the specifier of *v*P and the object is the complement of the verb, as you'd expect:

(45)



9.3 *v*P and Accusative Case assignment

Having introduced the functional category *v* and talked a little bit about its role in assigning the θ-role to the subject, what I want to do in this section is give some more arguments for *v* and, in particular, argue that not only does *v* assign a θ-role to the subject, it also assigns Accusative Case to the object.

One of the initial motivations for changing the way we think about Case assignment is conceptual, as first discussed by Chomsky in joint work with Howard Lasnik (published as Chapter 1 of Chomsky 1995). In Chapter 4 we argued that government provided the unifying concept for Case assignment. I^o and V^o are both governors. With respect to nominative Case, I^o m-commands the subject in Spec IP, with no barrier intervening. It is therefore able to assign Nominative Case to the subject. For Accusative Case, the verb m-commands the direct object and again no barrier intervenes.

However, Chomsky and Lasnik note that, despite the fact that government allows us to provide a unified analysis of Nominative and Accusative Case assignment, it glosses over the fact that, in reality, Case assignment takes place in two different structural configurations. Nominative Case is assigned in a specifier–head relationship (between I and Spec IP), but Accusative Case assignment takes place in a head–complement relationship (between V and its sister DP). Chomsky and Lasnik believe that we would have a better theory if we could make Nominative and Accusative Case assignment more unified, specifying a single configuration in which Case assignment takes place.

Of course, it is possible that the single unified configuration we're looking for is neither of the ones we have at the moment, neither specifier–head nor head–complement. However, the easier approach would be to try to collapse one into the other. Perhaps Nominative Case assignment 'really' happens in a head–complement configuration. Or perhaps it's the other way around: Accusative Case assignment 'really' happens in a specifier–head configuration.

Chomsky and Lasnik suggest that it's Accusative Case assignment which is the one that is 'out of step', so to speak, and that Case assignment should be unified as a specifier–head relationship with a functional category. For Nominative Case assignment, we already have the functional category IP, with the subject in the specifier of IP. For Accusative Case assignment, Chomsky and Lasnik originally introduced a functional category AgrO, but, as the theory developed, its functions were taken over by little *v*.

So, it's because of general considerations of theory neatness that we would prefer to have one configuration for Case assignment rather than two. These are the kinds of considerations that I've been calling 'conceptual'. However, as mentioned in the Introduction, we can be more sure of our conclusions when we find *empirical* motivation. In other words, conceptual considerations are most useful when they serve to predict concrete facts, especially ones that we had not noticed before. In the remaining parts of this section, we'll examine some empirical arguments for the correctness of Chomsky and Lasnik's hypothesis, and specifically for the idea that little *v* is involved both in Agent

θ-role assignment and Accusative case assignment, starting with an observation which is known as 'Burzio's Generalization'.

Burzio's Generalization

In his 1986 book *Italian Syntax*, Luigi Burzio noticed a correlation between θ-role assignment to an external argument and Accusative Case assignment, presenting an overwhelming amount of evidence for what came to be known as Burzio's Generalization (1986: 178):

(46) All and only the verbs that can assign a θ-role to the subject can assign Accusative Case to an object.

In essence, verbs which don't assign a θ-role to a subject DP also don't assign Accusative Case.

One of the most obvious illustrations of Burzio's Generalization is the process of passivization.

(47) John was arrested t

As we discussed in Chapter 6, passivization does two things: it takes away (1) the verb's ability to assign a θ-role to its subject and (2) its ability to assign Accusative Case to its object. As a result, *John*, which is the complement of *arrested* and gets the Patient θ-role associated with it, has to move to subject position in order to get Nominative Case from I. There's no problem with the θ-criterion, because the subject position now doesn't have a θ-role associated with it. (Or, if we're thinking in terms of the VP-internal Subject Hypothesis, the empty Spec VP at D-Structure doesn't cause a θ-criterion violation because the verb's ability to assign its Agent θ-role to an element in this position has been taken away.) As things stand, though, there's nothing obvious that connects these two phenomena.

However, now that we have the light verb *v*, it looks like things fall together really well if we assume that *v*, in addition to assigning the subject θ-role, also is what assigns Accusative Case to the object. First, we get an explanation for Burzio's Generalization, which is now just about the presence or absence of *v*. If it's present, you get a θ-role for the subject and Accusative Case for the direct object. If *v* is not there, you don't get either of these things. But we also solve the conceptual Case assignment problem. If *v* assigns Accusative Case, then both Nominative and Accusative Case involve a specifier–head relationship with a functional category.

Antecedent contained deletion in English

Hornstein (1995) provides further interesting empirical evidence from English which, although couched in an earlier framework, can be 'reinterpreted' into an argument that Accusative Case is assigned by moving to the specifier of *v*P.

In this case, the argument comes from a phenomenon referred to as *antecedent contained deletion*, which is a subpart of a more general process called *ellipsis*, which we've seen before. (48) illustrates ellipsis:

(48) John has gone to the store and Bill has ... too.

Clearly, the second conjunct is missing some information, indicated by the ellipsis. Specifically, it seems to be missing its VP. However, (48) is not ungrammatical. Instead, the VP for the second conjunct is interpreted as identical to the VP in the first conjunct. So, as far as the interpretation is concerned, we have (49):

(49) John has gone to the store and Bill has [gone to the store] too.

There are various ways in which such an ellipsis is accounted for, but one standard assumption is that the whole VP from the first conjunct is copied into the VP of the second conjunct at LF. That will ensure that *Bill has too* gets the correct interpretation. The predicate of which Bill is the subject isn't just *gone*. It's *gone to the store*. (48) doesn't mean that John has gone to the store and Bill has just gone somewhere. It seems that it's the whole VP *gone to the store* which gets copied.

There are cases, however, where the VP gap that we're trying to fill is *inside* the VP that we're trying to copy. These are instances of *antecedent contained deletion*. It turns out that the VP-copying explanation for ellipsis runs into some potentially serious problems when we look at sentences with antecedent contained deletion. Take a look at (50):

(50) John kissed everyone that Sally did [e].

Just as we did above, in order to recover the 'missing' material, and figure out what the meaning of (50) as a whole is, we need to copy a VP into the place of the gap (indicated by '[e]'). However, watch what happens when we try to apply our VP-copying rule. The only VP around is the matrix VP *kissed everyone that Sally did [e]*. And what happens when we copy that VP? We get (51):

(51) John kissed everyone that Sally did kiss everyone that Sally did [e].

And this doesn't solve our problem. We still don't know what the sentence as a whole means, because we've still got elided material. In other words, we've still got a gap to fill. Just as before, there's no VP to copy into the gap other than *kissed everyone that Sally did [e]*, but that contains another gap. As you can see, we can continue this procedure for as long as we want to, but we'll never get to the end. It seems that no matter what we do, we'll never be able to figure out what (50) means, because we'll always have missing material that needs filling in.

May (1985) realized that this problem of the interpretational infinite regress with antecedent contained deletion could be solved by QR (and in fact gave this as an argument for QR). Let's look at (50), not at S-Structure, but at LF:

(52) $[_{IP}$ everyone that Sally did $[e]_i$ $[_{IP}$ John $[_{VP}$ kissed $t_i]]]$

At LF, the entire quantified phrase *everyone that Sally did [e]* undergoes QR and adjoins to IP. Notice how this changes the picture. At LF, the matrix VP no longer contains a gap. It contains a trace, which is crucially different. If we take the VP *kissed* t_i and paste it into the gap, we get (53):

(53) $[_{IP}$ everyone that Sally did $[_{VP}$ kissed $t_i]_i$ $[_{IP}$ John $[_{VP}$ kissed $t_i]]]$

There's no elided material anywhere in (53), so we've managed to avoid the infinite regress problem. And we've also managed to get exactly the interpretation that (50) has. The LF in (53) is interpreted as 'for every x such that Sally kissed x, John kissed x'. And that seems to be exactly what (50) means.

It turns out that we can use antecedent contained deletion in ECM contexts to construct an argument for *v*P in English. Consider the contrast in (54):

(54) a. ?I expect everyone you do to visit Mary.
b. *I expect everyone you do will visit Mary.

Clearly, what (54a) and (54b) are trying to convey is mutual expectation. You have some number of people who you expect will visit Mary, and I want to assert that I have the same people in mind.

Let's take these back to front and start with (54b). What we find in (54b) is that we run into another infinite regress problem, even with QR. The LF of (54b) is (55), with the quantified phrase *everyone you do [e]* adjoined to the embedded IP:

(55) $[_{CP}$ $[_{IP}$ I $[_{VP}$ expect $[_{CP}$ $[_{IP}$ everyone you do $[e]_i$ $[_{IP}$ t_i will visit Mary]]]]]].

We need to find a VP to plug into the gap $[e]_i$, but there isn't one that works. We could try the VP *visit Mary*, but that's not what (54b) means. The gapped VP has to do with expectation, not with visiting Mary. But the VP which is headed by *expect* is *expect everyone you do[e]* t_i *will visit Mary*, and this VP has another [e] gap. If we plug this VP into the gap in the LF in (55), we'll get another gap and so on and so on.

Look how this changes with (54a), which feels a little clumsy, but which is much better than (54b). In (54a), the gap is in the ECM object. Under the hypothesis that we're advancing here, even though it's the subject of the lower clause, it must move at LF to the specifier of *v*P in order to be assigned Accusative Case:

(56) $[_{CP}$ $[_{IP}$ I $[_{vP}$ everyone you do $[e]_i$ $[_{VP}$ expect $[_{IP}$ t_i to visit Mary]]]]]

It will then undergo QR and adjoin to the matrix IP:

(57) $[_{CP}$ $[_{IP}$ everyone you do $[e]_i$ $[_{IP}$ I $[_{vP}$ t'_i $[_{VP}$ expect $[_{IP}$ t_i to visit Mary]]]]]]

With the LF in (57), we've solved our infinite regress problem. We've got a VP *expect t_i to visit Mary* which we can copy into the position of the gap:

(58) $[_{CP}\ [_{IP}$ everyone you do $[_{VP}$ expect $[_{IP}\ t_i$ to visit Mary$]]_i\ [_{IP}$ I $[_{vP}\ t'_i\ [_{VP}$ expect $[_{IP}\ t_i$ to visit Mary$]]]]]]$

This also results in the correct interpretation: for every person x such that you expect x to visit Mary, I expect x to visit Mary.

This is a good argument for movement to *v*P at LF, because it's the move to Spec *v*P which allows the quantifier to escape from the embedded IP. If we didn't have movement to the specifier of *v*P, then when the quantified phrase underwent QR, it would adjoin to the embedded IP and we'd have (59) at LF:

(59) $[_{CP}\ [_{IP}$ I $[_{VP}$ expect $[_{IP}$ everyone you do $[e]_i\ [_{IP}\ t_i$ to visit Mary$]]]]]$

When we then tried to look around for a VP to fill the gap in [e], we'd have the same problem we had when we were looking at (54b). There are only two VPs around: *visit Mary* and *expect everyone you do [e]$_i$ t_i to visit Mary*. We don't want the first one, because that doesn't get the right meaning, and we don't want the second one because that has elided material in it, and we'll get into an infinite regress.

The fact that we get the right meaning without an infinite regress suggests that the ECM subject is raising to a position outside the embedded IP, something which is already predicted under the assumption that it must move, for Case reasons, to the specifier of *v*P. Notice that this has the effect of making ECM less 'exceptional'. Both direct objects and ECM objects receive Accusative Case because both move to the specifier of *v*P. That's another nice conceptual bonus.

*v*P and 'multiple' specifiers

So this all sounds great, but what are the mechanics of Accusative Case assignment by *v*? Clearly, it must be the case that the specifier–head relationship happens at LF rather than at S-structure. Just as we mentioned when discussing VP-shells, if the direct object moves to the specifier of little *v* in the syntax, then we're going to end up with the wrong word order (ignoring the VP-internal Subject Hypothesis for just a second):

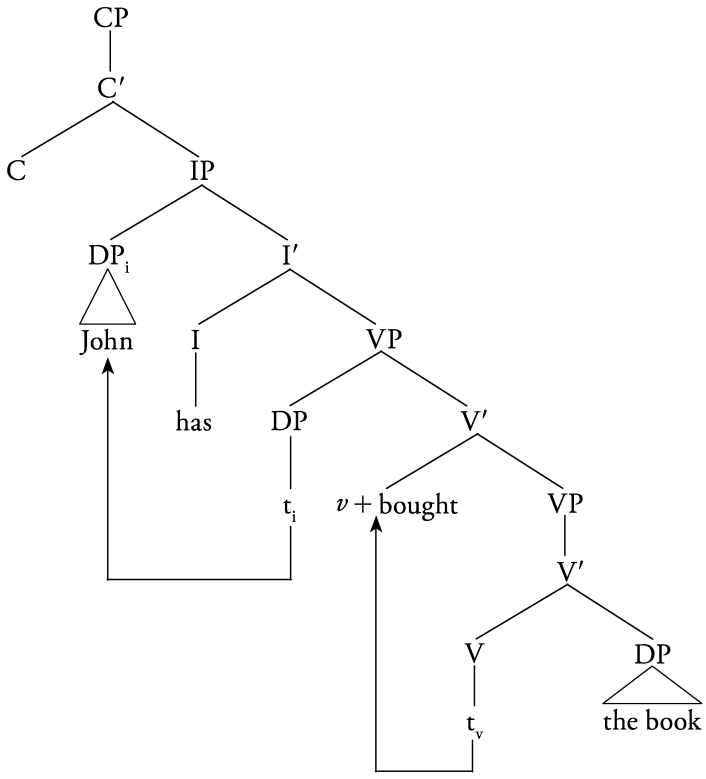
(60) $[_{IP}$ John $[_{I}$ has $[_{vP}$ the book$_i$ v+bought$_v\ [_{VP}\ t_v\ t_i\ ]]]]$

It looks more at S-structure like the direct object is in its original D-structure position as complement to the verb:

(61) $[_{IP}$ John $[_{I}$ has $[_{vP}$ v+bought$_v\ [_{VP}\ t_v$ the book $]]]]$

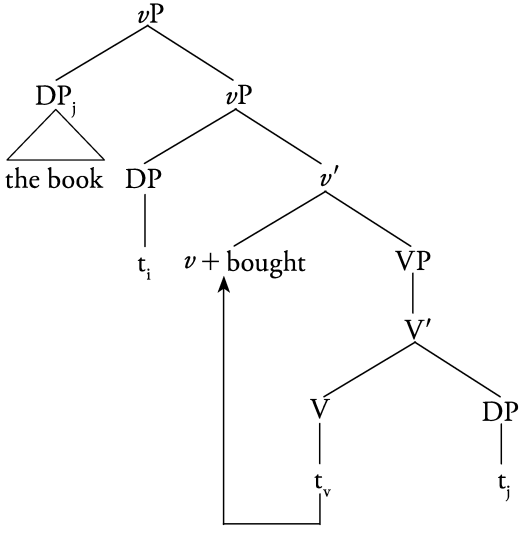
However, as I mentioned a moment ago, we've also got to at least say something about how this integrates with the VP-internal Subject Hypothesis. (61), of course, is really (62):

(62)



So when the object raises at LF, we're going to have (63):

(63)



But this makes the direct object look like it's adjoined to *v*P rather than a specifier of *v*P. And if specifier–head relations are really the key to Case assignment, we're going to have to change the way specifiers are defined, in order to allow for the possibility for an XP to have more than one specifier.

Up to this point, we've defined specifiers in a purely configurational way. Specifiers are immediately dominated by XP and a sister to X′. This in turn makes a clear distinction between specifiers and adjoined elements (like quantifiers that are adjoined to IP). As we discussed in Chapter 7, adjoined categories are immediately dominated by a *segment* of XP and a sister to another *segment* of XP. The problem is, of course, under a configurational definition of 'specifier', you can't have more than one. There just isn't space.

However, if we move away from a purely configurational definition of 'specifier' and think about them more along 'functional' lines, there's no conceptual reason why there couldn't be multiple specifiers for a given XP. What specifiers do is establish relations between a particular element (usually a DP) and a head. So to say that a given element is in the specifier of IP, for example, is just to say that it has established a particular relation with the head of IP (Infl) (in this case, so that Nominative Case can be assigned). If two different DPs establish two different relations with a particular head for two different reasons, then there doesn't seem to be any reason not to say that both of these elements are specifiers. That seems to be precisely what we have in the case of *v*. One DP, the subject, needs to get into a relation with *v* in order to be assigned the Agent θ-role that *v* has. Then another DP needs to get into a separate relation with *v* in order to be assigned Accusative Case.[2]

9.4 From IP to TP

Having spent a great deal of time in the previous sections revising the structure of VP in various respects, in this section we'll move up to consider IP. As it turns out, there is very little that we need to say about IP. As you'll recall from earlier chapters, IP contained two features: (1) subject-verb agreement features and (2) tense features. It's a slight over-simplification, but, for our purposes, the main thing that changes is that IP loses its agreement feature, leaving only the Tense feature. (This is in the context of a fairly radical change to the way agreement is conceptualized, which it would take us too far afield to discuss properly, so we're not adopting it here.) As a result of this loss of agreement features, IP changes name. Instead of IP (Inflection Phrase), the category is known in the Minimalist Program as TP (Tense Phrase).

Again, for our (slightly over-simplified) purposes, nothing else really changes. Rather than subjects moving to the specifier of IP at S-structure (assuming the VP-internal Subject Hypothesis from Section 9.1 above), they now move to the specifier of TP. Instead of Nominative Case being assigned by I°, we'll now say that it's assigned by T°. (This is not to say that subject-verb agreement no longer exists. What happens is that, instead of, for example, third person singular being a feature that both Infl has and the DP *the man* has, it's just *a man* that has the feature, which then 'values' the Tense head. Again, we won't

be pursuing this idea.) And we'll also say that QR adjoins a DP to TP rather than to IP.

9.5 Rizzi (1997) and the 'cartographic' approach to CP

Now that we've had a look at the VP and IP areas, I want to move on to the last part of the clause that we'll be talking about: the CP area. In particular, we'll be talking about an approach pioneered by Rizzi (1997) called the 'cartographic' or 'split CP' approach. He suggests that CP isn't a single unified functional projection. Rather, what we've been calling CP is actually a collection of several functional projections, including ForceP, TopicP, FocusP and Fin(iteness)P.

To begin with, let's consider the role that CP plays in the grammar as we've seen it so far. As Rizzi (1997) notes, CP is the functional category that links IPs with other categories in the syntax. CPs are selected by verbs (and some other categories, like prepositions) and their heads in turn select IPs of particular types. As such, they have two syntactic aspects that we need to account for. One is 'upward-looking', relating to the category which selects CP. The other is 'downward-looking', relating to the IP that C selects.

Let's begin with upward-looking aspect. What I mean by this is that it is properties of C that determine, for example, what verb the CP can be a complement of. As we discussed back in Chapter 7, certain verbs can only take a declarative CP (these are the CPs that we called '-wh'), other verbs can only take an interrogative CP (the CPs we called '+wh'), and some verbs can take either kind. So we saw paradigms like those in (64)–(66):

(64) a. I hope that John fixed the car.
b. *I hope how John fixed the car.
c. *I hope whether/if John fixed the car.

(65) a. *I wonder that John fixed the car.
b. I wonder how John fixed the car.
c. I wonder whether/if John fixed the car.

(66) a. I know that John fixed the car.
b. I know how John fixed the car.
c. I know whether/if John fixed the car.

Saying the same thing with slightly different terminology, some CPs have declarative 'force' and other CPs have interrogative 'force', and verbs select for CPs with different forces. So one thing that CPs need to have is a specification of their force.

The other aspect that CP has is downward-looking, in the sense that CPs themselves select different kinds of IPs. CPs select for IPs which are *finite* (that is, IPs which contain a finite verb) or non-finite (IPs which are infinitival or gerundive).

(67) a. I hope [$_{CP}$ that [$_{IP}$ John fixed the car]] (finite IP)
b. I hope [$_{CP}$ [$_{IP}$ to leave tomorrow]] (non-finite IP)

(68) a. I think that John fixed the car. (finite IP)
b. *I think John to have fixed the car. (non-finite IP)

As (67) and (68) illustrate, some verbs are compatible with CPs that select both finite and non-finite IPs, but not necessarily all are.

There are essentially two ways that we could encode these two syntactic characteristics of C. The first way would be the way that we implicitly took in Chapter 6, which is in terms of features on a single functional head. In the same way that the functional category I contains both tense and agreement features, we could say that C° contains a 'force' feature – which determines whether the CP is declarative or interrogative (or imperative, etc.) – and a 'finiteness' feature, which determines whether the CP selects for a finite or non-finite IP. The other way that we could encode these characteristics is by decomposing C into two different functional categories: ForceP and Fin(iteness)P. Force°, the head of ForceP, would be specified as declarative, interrogative, etc., and Fin°, the head of FinP, would be specified as +finite or –finite.

If that's all there is to say about it, then there's really nothing to decide between the two options. They're just different ways of saying the same thing. But what Rizzi argues is that, in Italian, topics and focalized elements can actually come *between* the Force part and the Finiteness part of CP, and *that* gives us a reason to prefer one approach over the other. If we have a single, unified CP with features, it's not clear how we could fit this fact in. So Rizzi suggests that the second option is preferable and that CP should be decomposed into a series of functional categories: ForceP, TopicP, FocusP and FinP.[3]

The key to Rizzi's analysis is a difference between *che* (literally 'that'), the complementizer that introduces finite clauses, and prepositional complementizers such as *di* (literally 'of') which introduce non-finite clauses. It's the difference between these two complementizers, and the way they interact with other elements around the left edge of the sentence, that make us want to adopt the 'split CP' approach, as opposed to a single CP with different features.

Let's start by just illustrating the two complementizers in (69a) and (69b):[4]

(69) a. Credo che loro apprezzerebbero molto il tuo libro.
believe.1sg that they would appreciate very much the your book
'I think that they would appreciate your book very much.'

b. Credo di apprezzare molto il tuo libro.
believe.1sg of to.appreciate very much the your book
'I think that I appreciate your book very much.'

In either case, the direct object *il tuo libro* can be moved to the left periphery of the embedded clause and doubled with a clitic pronoun. It's a process which is equivalent to the phenomenon of left dislocation in English:

(70) (As for) John, I like him.

Left-dislocated elements are typically set off from the rest of the sentence with a pause (represented orthographically by commas) and represent old information; that is, information which is previously available from the discourse.

What's interesting about the so-called 'clitic left dislocation' in Italian is that you get a different order with *che* and *di* relative to the dislocated element. When you have *che* introducing a finite clause, the dislocated element obligatory *follows che*.

(71) a. Credo *che il tuo libro,* loro lo apprezzerebbero molto
believe.1sg that the your book they it would appreciate very much

b. *Credo, *il tuo libro, che* loro lo apprezzerebbero molto
believe.1sg the your book that they it would appreciate very much
'I think that, as for your book, they would appreciate it very much.'

However, when you have *di* introducing a non-finite clause, the left-dislocated element doesn't follow *di*; it obligatorily *precedes di*:

(72) a. Credo, *il tuo libro, di* apprezzarlo molto.
believe.1sg the your book of to.appreciate.it very much

b. *Credo *di il tuo libro,* apprezzarlo molto
believe.1sg of the your book to.appreciate.it very much
'I think that, as for your book, I appreciate it very much.'

Under a traditional analysis of (69), *che* and *di* would presumably both occupy C^{o}, but that looks like that can't possibly be right given (71) and (72). It looks like we need at least *three* positions. One position for the left-dislocated element, one which *precedes* it, which is presumably where *che* is, and one position which *follows* it, which is presumably where *di* is.

Following Rizzi (1997), let's call the position of the left-dislocated element TopicP. Let's assume that the position which precedes TopicP is ForceP. This will make ForceP the highest of the three functional categories. That sort of makes sense because it will then be ForceP which is the sister to V, and it's the force of the clause (declarative, interrogative, etc.) that verbs and other elements which take CP as their complement care about. So, for example, *wonder* takes as its sister a ForceP which has a [+interrogative] feature. We'll assume that the position which follows TopicP is FinP, and that'll be the functional category that selects IP, finite or non-finite, depending on its feature. *Di*, for example, heads a FinP which selects a non-finite IP.

However, it looks like we need more than just a TopicP between ForceP and FinP. We'll also need another functional category, which we'll call FocusP. To see this, we need to take a look at the way in which focalized elements interact with clitic left dislocation. It turns out that we can have both a clitic-left-dislocated element and a focalized element and, when we do, the focalized element follows the clitic-left-dislocated element.

Focalized elements differ from left-dislocated elements in a number of ways. They appear at or near the left edge of the sentence, but receive focal stress (illustrated by capitalization) and, unlike left-dislocated elements, represent *new*, rather than old, information.

(73) JOHN, I like (BILL, I don't like).

In Romance languages, like Italian, focalized elements are also *not* doubled by a clitic.

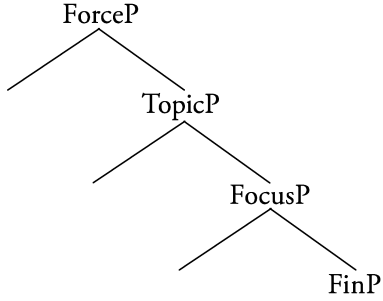
(74) a. IL TUO LIBRO ho letto (, non il suo)
the your book have.1sg read not the his
'Your book, I read (not his).'
b. *IL TUO LIBRO lo ho letto (, non il suo)

What's interesting is that when a sentence contains both a left-dislocated element and a focalized element, the left-dislocated element precedes the focalized element.[5]

(75) Credo che a Gianni, QUESTO, gli dovremmo dire
believe.1sg that to G this to.him should.1pl say.inf
'I think that, to Gianni, THIS, we should say.'

Thus, we have the following structure for the expanded CP, under Rizzi's approach:

(76)



It just remains to finish up by talking a little bit about how various elements that we've seen in previous chapters would fit into the expanded CP. First, as the Italian examples in (69)–(75) suggest, complementizers occupy the highest head, ForceP, within the expanded CP. They precede both topics and focalized elements.

Wh-elements, on the other hand, appear to reside in two different places. In previous chapters, we didn't distinguish between wh-words, as in (77a), and relative clause operators, as in (77b).

(77) a. Who$_i$ t$_i$ left?
b. The man who$_i$ t$_i$ left.

We just had one position, the specifier of CP, and we said that both kinds of elements moved there. However, there's good evidence from Italian that these two kinds of elements move to *different* functional categories now that we've split CP up.

Relative clause wh-operators move to the specifier of ForceP, as they obligatorily precede topics like *il premio Nobel* in Italian:

(78) a. Un uomo a cui, il premio Nobel, lo daranno senz'altro
a man to whom the prize Nobel it will give.1pl undoubtedly
'A man to whom, the Nobel prize, they will undoubtedly give...'
b. *Un uomo, il premio Nobel, a cui lo daranno senz'altro

Il premio Nobel must be a left-dislocated topic rather than a focalized element, as it is doubled by the clitic *lo*, and the relative clause wh-phrase must precede it. Relative clause operators do also precede focalized elements, as you'd expect (Rizzi 1997: 298):

(79) a. Ecco un uomo a cui IL PREMIO NOBEL dovrebbero dare (non il premio X)
here a man to whom the prize Nobel should ought.3pl give non the prize X
'Here is a man to whom THE NOBEL PRIZE they should give (, not Prize X).'
b. *Ecco un uomo IL PREMIO NOBEL a cui fovrebbero dare (non il premio X)

However, question operators do not occupy the specifier of ForceP. Instead, they occupy a lower position. Unlike the relative clause operator in (78), question operators must *follow* topics, rather than precede them:

(80) a. *A chi, il premio Nobel, lo daranno?
to whom the prize Nobel it will give.3pl
'To whom will they give the Nobel Prize?'
b. Il premio Nobel, a chi lo daranno?

Crucially though, question operators are flatly incompatible with a focalized element. They are ungrammatical whether they precede or follow.

(81) a. *A chi IL PREMIO NOBEL dovrebbero dare?
to whom the prize Nobel ought.3pl to give
'To whom THE NOBEL PRIZE should they give?'
b. *IL PREMIO NOBEL a chi dorebbero dare?

Since you can't have a question word and a focalized element in the same clause, this suggests that the two are competing for the same position, the specifier of FocusP. We thus have the following breakdown of positions within

the split CP:

(82)
ForceP:	complementizers, relative clause operators
TopicP:	topics
FocusP:	focalized elements, question operators
FinP:	other prepositional complementizers (e.g. *di*)

However, when the details of the 'split CP' are not relevant to the point at hand, it's very common to see these projections abbreviated together as 'CP'. Like drawing triangles in trees, this is just a space-saving shorthand.

Summary

In this chapter, we've been looking at various ways in which the system of phrase structure containing VP, IP and CP evolved during the late Government and Binding Theory period into the Minimalist Program. In looking first at VP, we saw that we needed to extend the 'VP-shells' idea that we first talked about in Chapter 2. We introduced a 'light' verb *v*, which encodes causation and which takes a VP as its complement. Subjects of a sentence are base-generated in a specifier of *v*P, where they are assigned the Agent θ-role under government from *v*. We also argued that *v* is in addition the functional category which is responsible for assigning Accusative Case, and that objects raise to a second specifier of *v*P in order to receive it. In this way, both Nominative and Accusative Case are assigned in a specifier–head relationship with a functional category. Moving upward, we saw that IP merely underwent a change of name from IP to TP, in order to reflect the fact that the main feature that it bears is a Tense feature, rather than an Agreement feature. Finally, we introduced Rizzi's (1997) 'cartographic' or 'split CP' approach to the highest functional category of the sentence. We suggested that CP needed to be broken down into a series of functional categories, with ForceP as the highest category which Vs take as their complement, down to lowest head Fin(iteness)P, which takes TP as its complement.

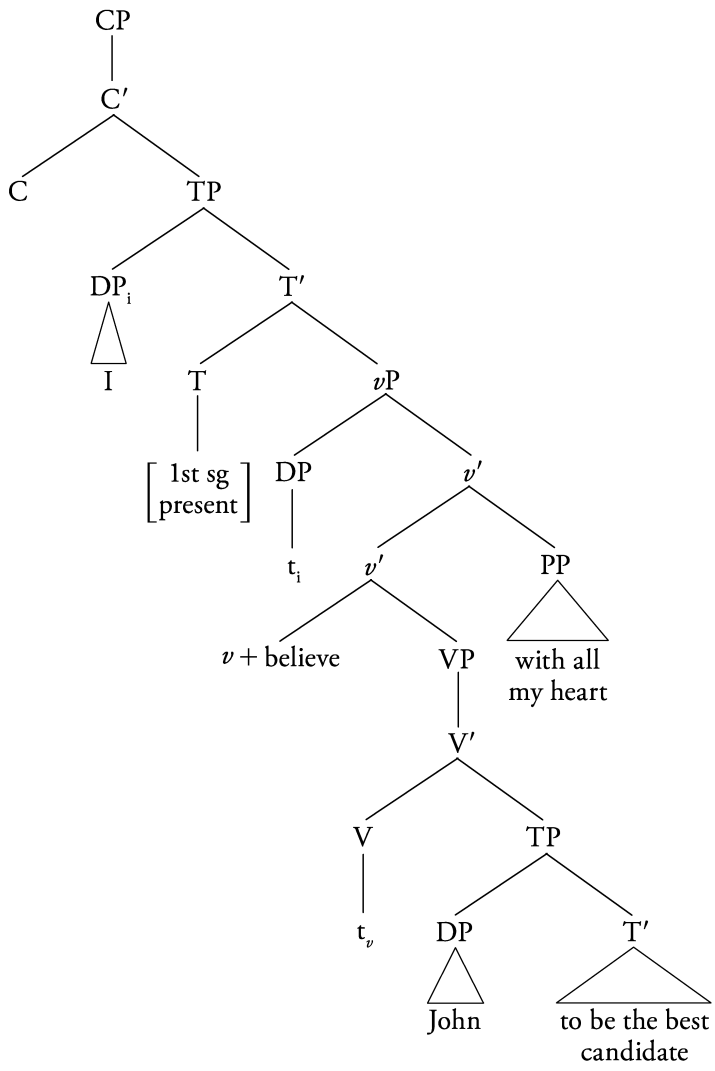
Open issue: more evidence for *v*P in english

I want to return to an argument for *v*P from English. I discuss it here in the Open Issue section rather than in the main text because, although the data and the argument are really interesting, the data are, to my ear, a little unclear and there are some potentially problematic implications.

Lasnik and Saito (1991) suggest that there is further evidence from c-command that subjects in ECM constructions move to a position outside the VP. In a nutshell, ECM subjects, which are in the specifier position of the embedded TP, appear to c-command elements in the *higher* *v*P. This is quite unexpected under the standard analysis of ECM, in which a sentence like (83) has the tree in (84):

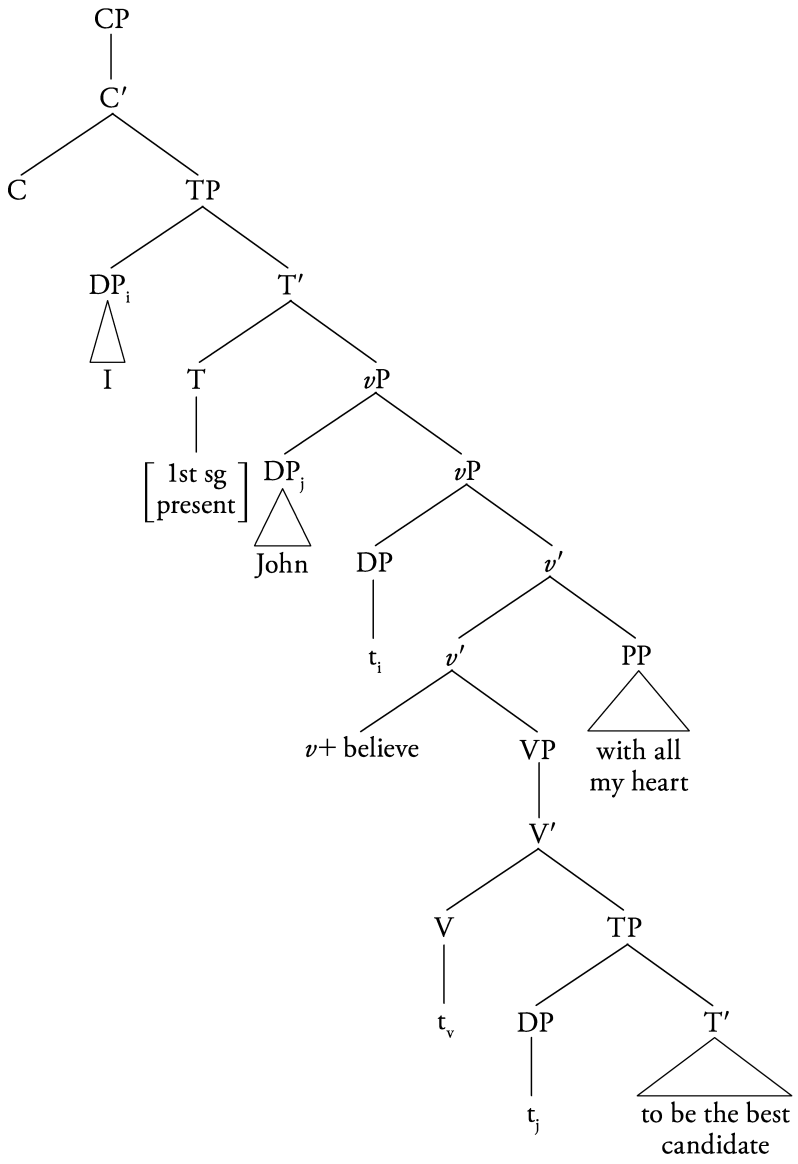
(83) I believe John to be the best candidate with all my heart.

(84)



If the subject of the ECM infinitive *John* is assigned Accusative Case by *believe* without moving from its D-Structure position, it should not be able to c-command a prepositional phrase like *with all my heart*, which in (84) is a sister to the matrix *v*′. If, on the other hand, it could be shown that the ECM subject *is* able to c-command elements in the matrix *v*P, that would be quite strong evidence that ECM subjects are moving from their D-Structure position to somewhere outside the VP. The outer specifier of a *v*P which dominates VP would seem to provide exactly what's needed.

(85)



As you can see from the tree in (85), if the ECM subject were to move to the outer specifier of *v*P, it would be in a position to c-command other elements in the VP.

To take a first example from Lasnik and Saito's data, let's consider their evidence from Binding Theory. As we've discussed in Chapter 5 and Chapter 8, Principles A, B and C all use c-command to define the range of possible antecedent relationships. If anything can be said to be the key unifying force in Binding Theory, it's c-command. Principle A, for example, specifically requires,

among other things, that the anaphor be c-commanded by its antecedent. This is in order to ensure that sentences like those in (86) are correctly ruled out:

(86) a. *Himself$_i$ likes John$_i$.
b. *The woman who talked to John$_i$ yesterday hit himself$_i$.

Observations like those in (86) form some of the core facts about Binding Theory.

However, Lasnik and Saito observe contrasts like the following:

(87) a. The District Attorney proved [two men$_i$ to have been at the scene of the crime] during each other$_i$'s trials.
b. *The District Attorney proved [that two men$_i$ were at the scene of the crime] during each other$_i$'s trials.

Although the contrast in (87) is not as strong for me personally as I'd prefer, Lasnik (1999) reports that (87a) contrasts significantly with (87b) for most of the speakers that he's interviewed. If ECM subjects did not move to the outer specifier of *v*P, we'd predict that (87a) and (87b) should be identical. The first branching node dominating *two men* in both cases is the embedded TP. Therefore both sentences violate Principle A, since the anaphor *each other* is not c-commanded by its antecedent. On the other hand, if ECM subjects move to the outer specifier of *v*P at LF, then in (87a), but not (87b), *two men* will c-command *each other* – leading to the prediction that only (87b) is ungrammatical.

A related kind of example which Lasnik and Saito discuss concerns Negative Polarity Items (NPIs). Remember from the discussion in earlier chapters that NPIs must be c-commanded by a negative DP or by sentential negation:

(88) a. No witness saw any suspect.
b. *Any witness didn't see the car.

What Lasnik and Saito observe is that ECM subjects appear to license negative polarity expressions in the higher VP, which they shouldn't be able to do under the standard assumptions about ECM and Case.

(89) a. The District Attorney proved [*no suspect* to be guilty] during *any* of the trials.
b. *The District Attorney proved [that *no suspect* was guilty] during *any* of the trials.

(You might be able to get a handle on the intuition by thinking about it this way: (89a) is about how the District Attorney *failed* to prove something. It's a 'negative' sentence. (89b) on the other hand is a 'positive' sentence. The District Attorney proved something, namely that no suspect was guilty.) The negative element has to c-command 'any' in order to get this 'not any/ failure' reading, but if the ECM subject remains in the subject position of the embedded infinitive, it's not clear how that happens or why there should be a contrast. On the

other hand, if ECM subjects raise out of the infinitive VP to the outer specifier of *v*P, then they will be in a position to c-command other elements in the VP, and we explain the contrast in (89). In (89a), the ECM subject undergoes raising, but in (89b) the subject of the *finite* clause cannot. Therefore in (89a) we get the 'failure' reading and in (89b) we don't.

Unfortunately, though, this argument for object movement to the outer specifier of *v*P raises an issue which is going to be a little problematic for us. On the basis of Hornstein's (1995) antecedent contained deletion argument in Section 9.3, we concluded that ECM objects moved to the outer specifier *v*P *at LF*. This was consistent with the general view about English and with everything we've seen so far in this chapter. In the next chapter, we'll also see that main verbs in English don't move out of the V *until LF*. At S-Structure, we expect the verb in English to be in its base-generated position.

The problem is that, in order to account for the contrasts which Lasnik and Saito reveal, you need to assume that the ECM subject is moving to the outer specifier of *v*P at *S-Structure*, not at LF. This is because S-Structure A-movement changes binding relations and negative polarity licensing, but LF A-movement doesn't. Consider (90) and (91). In the (a) sentences, we have S-Structure movement. In the (b) sentences, we have a parallel LF movement as part of expletive replacement:

(90) a. Two men$_i$ seem to each other, t$_i$ to be in the room.
b. *There seem to each other$_i$ to be two men$_i$ in the room.

(91) a. No trainee$_i$ seems to any manager t$_i$ to be suitable for the job.
b. *There seems to any manager to be no trainee suitable for the job.

The ungrammaticality of the (b) sentences cannot be attributed to any incompatibility between expletive *there* and the intended associate, or some problem with expletive replacement, as (92) shows:

(92) a. There seem to be two men in the room.
b. There seems to be no trainee suitable for the job.

But this leaves us with a problem. If ECM subjects move *in the overt syntax* to the outer specifier of *v*P, but verbs don't leave the *v*P until LF, then you'd expect that the word order that you get in ECM constructions would have the ECM object in the outer specifier of *v*P *preceding* the verb:

(93) *I Mary$_i$ believe t$_i$ to be the best candidate.

This is clearly not a grammatical sentence of English. And in fact we gave examples like (93) as an argument that the movement had to take place at LF.

Lasnik and Saito's evidence looks really interesting. If we accept it, we've got some nice evidence for the existence of *v*P in English. The downside is that

it creates some problems for us in terms of standard assumptions about verb movement. However, we'll return to an interesting way of resolving this problem in Chapter 10's Open Issue.

FURTHER READING

The bibliography in this chapter and the next chapter are a little bit different from the bibliographies in the other chapters. As I've mentioned a couple of times, some of the discussion in this chapter is adapted from the original so as to fit in with what we've done previously and to filter out assumptions from the Minimalist Program that we haven't discussed. Because of that, what I would recommend people move on to are textbooks that discuss the Minimalist Program. My favourite, and the one that's probably most similar in spirit to this book (in terms of looking at not only the hows but also the whys), is *Understanding Minimalism* (2005) by Norbert Hornstein, Jairo Nunes and Kleanthes K. Grohmann. Andrew Radford has also published a number of textbooks on Minimalist syntax, and his 2009 *Analyzing English Sentences: A Minimalist Approach* I think is the best of the bunch. It's got a lot of discussion of technology – how to *do* a Minimalist analysis of X, Y or Z. What it doesn't do so much is explain *why* a Minimalist analysis might be a good thing or better than the alternatives. David Adger's (2003) book *Core Syntax* also has a lot of good discussion of introductory Minimalist topics. When you're ready to move on to primary Minimalist Program literature, I definitely recommend getting hold of a copy of Bošković and Lasnik's (2007) anthology *Minimalist Syntax: The Essential Readings*.

Having said all that, there are some issues that we've been discussing which don't require a detailed knowledge of the Minimalist Program (often because the ideas pre-date it). The VP-internal Subject Hypothesis, for example, originated in the 1980s and is discussed in a number of PhD dissertations of the time, such as Zagona (1982) and Kitagawa (1986). Also, Bobaljik (2003) provides an excellent overview article on the stranded/floating quantifier issue.

EXERCISES

Exercise 1

Draw the S-structure trees for the following sentences using the revised clause structure of this chapter:

1. John can put the book on the table.
2. Alice slept in the house.
3. Who did Bill like?
4. The man who Bill liked left yesterday.
5. The students can both read the book.
6. She is not a girl who misses much.

Exercise 2

Draw the LF trees for the following sentences using the revised clause structure of this chapter:

1. Mary read every book on the table.
2. Mary read every book on the train.
3. Desmond has a barrow in the marketplace.
4. Martha believed Paul to have spent his days in conversation.
5. Every package was received.
6. Every teacher who had bought the DVD returned it.

Exercise 3

How might the following contrast be used to argue for the existence of a 'split CP' vs an 'unsplit CP' in English?

1a. I know that under no circumstances would I have done such a thing.
1b. *I know under no circumstances that would I have done such a thing.

What is the relevance of the following contrast?

2a. The man who, under no circumstances, would have done such a thing ...
2b. *The man, under no circumstances, who would have done such a thing ...

If you're a native speaker of English, what's your judgement about (3)?

3. I remember who under no circumstances would have done such a thing.

Does your judgement present a problem for the theory? If so, can you think of any way that the data might be explained? (If you're not a native speaker, what would you expect the judgement to be, and how might you explain an 'unexpected' result?)

10 Towards the Minimalist Program: Movement Revisited

In the previous chapter, we revised some of our assumptions about phrase-structure, with a view towards introducing some more advanced concepts and structures. In this chapter, I want to do the same thing with respect to movement. Just as in the previous chapter, what we're doing is making a transition from Government and Binding Theory to the Minimalist Program. Where arguments have been formulated within the Minimalist Program, we'll adapt them into the framework that we've been assuming throughout the book, rather than presenting the arguments exactly as you'd find them in the literature of the Minimalist Program. This again is because the Minimalist Program makes some extremely radical modifications to basic aspects of the way that movement and structure-building work, which it would take us too far afield to explain properly (but again see the references in Chapter 9). However, this will hopefully at least give a flavour of the sorts of suggestions that people have made.

10.1 The copy theory of movement

In this section, I want to introduce one of the most striking changes to the theory of movement that comes out of the transition from Government and Binding Theory to the Minimalist Program. When we introduced the idea of movement back in Chapter 6, we also introduced the idea of traces, a kind of placeholder element which marks the D-structure position of the moved item. Then, in Chapter 8, we argued that traces actually have different Binding Theory properties, depending on whether it's a trace of A-movement (e.g. passive or raising) or a trace of A′-movement (e.g. wh-movement).

However, more recently there have been a number of different arguments to the effect that moved items don't leave traces behind. Instead what you have are actual *copies* of the moved elements which are marked with a feature that says to PF 'Don't pronounce me'. So, instead of representations like (1), you have representations like (2):

(1) a. John_i was hit t_i.

b. What$_i$ did John buy t$_i$?
c. [The man I saw yesterday]$_i$ seems t$_i$ to be in the garden.

(2) a. John was hit [John].
b. What did John buy [what]?
c. The man I saw yesterday seems [the man I saw yesterday] to be in the garden.

In Chapter 6, we alluded to one possible hint that movement leaves copies behind rather than traces: the fact that Binding Theory treats anaphors as if they were in the position of the trace:

(3) [Which picture of himself$_i$]$_j$ does John$_i$ like t$_j$?

John doesn't c-command *himself* in (3). However, *John* would bind *himself* if the anaphor were actually back in the position of the trace. We explained that example by saying that Binding Theory applied at D-structure rather than S-structure, but we'll see in this section some examples that can't be explained in the same way.

But before we get to the empirical arguments though, there has also been a conceptual argument advanced in favour of the copy theory of movement. As always, we're on the lookout for ways in which our theory can be simplified, ways in which we can reduce parts of the theory to other parts. What Chomsky argued was that traces were something 'extra' that our theory was adding. We know lexical items exist and that we need them for our theory. So if we eliminate traces and assume that chains are just composed of copies of lexical items, we've made our theory simpler. The two copies still form a chain, so they only need one θ-role, one Case, etc.

Reconstruction and quantifier raising

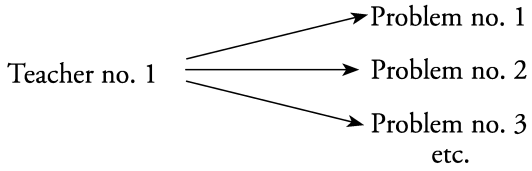
In terms of empirical arguments, let's start by looking at a phenomenon called 'reconstruction'. Like the Binding Theory example in (3) above, reconstruction is a phenomenon in which it looks like we need to say that *the whole, original element itself* is back in the trace position. There are a lot of different contexts in which you get reconstruction effects, but some of the more straightforward are ones involving Quantifier Raising.[1]

To begin, it's long been noted that movement can affect scope, as Fox (1999: 159) observes. Consider first (4):

(4) John seems to a teacher to be likely to solve every one of these problems.

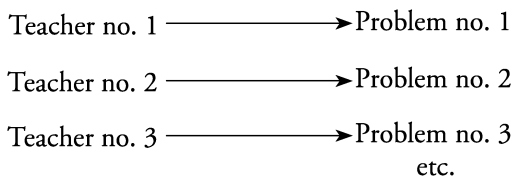
The crucial thing to notice about (4) is that the existential quantifier *a teacher* must have wide scope relative to the universal quantifier *every one of these problems*. So, in (4), there can only be a single teacher (say, John's maths teacher) to whom it seems that John is likely to have solved Problem 1, Problem 2, Problem 3, etc.

(5)



It can't have the reading where there are a multiplicity of teachers, each one of which thinks that John is likely to solve a different one of the problems.

(6)



And if we try to throw in some words to suggest or force the impossible reading, the result is really strange.

(7) #John seems to a different teacher to be likely to solve every one of the problems.

(7) is fine if 'different' refers to a different, single teacher (e.g. not John's *maths* teacher, but his *physics* teacher), but is completely impossible as an attempt to force the reading illustrated in (6), where there is a different teacher for each of the problems.

This is all pretty straightforwardly accounted for under the assumption, which we talked about in Chapter 7, that QR can only adjoin a quantifier to the nearest IP (which is now TP). The existential quantifier *a teacher* is going to adjoin to the matrix TP, and so will have a wider c-command domain than *every one of these problems*, which will only be able to adjoin to the embedded TP.

(8) $[_{CP}\ [_{TP}$ a student$_i$ $[_{TP}$ John seems to t_i $[_{TP}$ every one of these problems $[_{TP}$ to have solved $t_j]]]]]$

But consider now what happens if the lowest clause is passivized and it's the universal quantifier *every one of these problems* that undergoes raising to the matrix subject position instead of *John*.

(9) Every one of those problems seems to a (different) teacher to be likely to be solved by John.

Unlike (4), the universal quantifier now *can* have wide scope relative to the existential quantifier. There can be a multiplicity of teachers in (9), where each one thinks that John solved a different problem (the reading illustrated in (6) above). And unlike (7), changing *a teacher* to *a different teacher* in order to force the wide-scope reading for *every one of those problems* is perfectly fine.

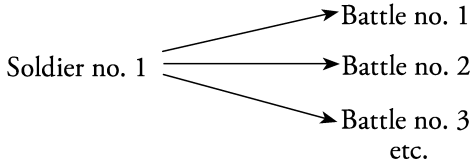
So what this shows is that movement can affect quantifier scope. After the universal quantifier undergoes raising at S-Structure to the matrix subject position, the closest TP is now the *matrix* TP when it undergoes QR at LF, and so it can take wide scope relative to the existential quantifier, which will also be adjoining to the matrix TP. (Of course, the interpretation where the existential quantifier takes wide scope is also still possible, as you'd expect when two quantifiers are in the same clause.)

What's interesting, though, is that movement does not *obligatorily* affect quantifier scope. That is to say, even after something has undergone movement, it can still take scope based on its D-Structure position. Consider (10), again from Fox (1999):

(10) One soldier$_i$ is expected by Napoleon t_i to die in every battle.

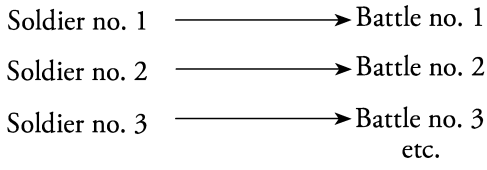
As you'd expect based on the above discussion, the reading where *one soldier* has wide scope relative to *every battle* is available, if rather odd pragmatically. This is the sci-fi/zombie reading in which there is one specific soldier who continually dies, is brought back to life and then dies again in the next battle. (It's weird, but it's there.)

(11)



However, the thing to notice about (10) is that the reading where *every* has wide scope relative to *one* is also available. This is the more pragmatically accessible reading (since it doesn't involve reanimation of the dead), where the one soldier who dies is different in every battle.

(12)



How do we explain this? Well, since the two quantifiers must be adjoined to the same TP in order to interact, there are basically two possibilities. The first possibility is that the two quantifiers are interacting because they're both adjoined to the *matrix* TP at LF. Since *one soldier* is already in the matrix clause at S-Structure, what we'd need to say is that *every battle* can optionally undergo some kind of 'long-distance' QR which is able to move it out of the embedded clause and adjoin it to the matrix TP. *Every battle* then could have wider scope than *one soldier*, assuming that *one soldier* was adjoined to the matrix TP first.

The second possibility is that the two quantifiers are interacting because they're both adjoined to the *embedded* TP at LF. *Every battle* is already in the

embedded clause. So, for this to work, what we need to say is that *one soldier* can take scope either via its S-structure position in the matrix clause or via its D-structure position in the embedded clause. When it takes its scope in the embedded clause, it's then possible for *every battle* to have wider scope than it.

However, as Fox points out (1999: 160–1), it looks like the first possibility, where *every battle* is able to undergo 'long-distance QR', isn't going to work. It predicts that you should *always* be able to get ambiguity in configurations like (10). If 'long-distance QR' out of an embedded clause is always possible, then there should always be the possibility for items in embedded clauses to interact scopally with items in the matrix clause, leading to scope ambiguities. However, this doesn't seem to be the case. Consider (13):

(13) One soldier$_i$ is expected by his$_i$ commander to die in every battle.

Under the interpretation where *his* is a bound variable, (13) is not ambiguous. *One soldier* must take wide scope relative to *every battle*. (In other words, (13) has only the pragmatically odd 'reanimation of the dead' reading.) It's not obvious why this should be. If the plan is to explain the ambiguity of (10) by claiming that *every battle* can undergo long-distance QR (so as to be able to interact with *one soldier*), why is long distance QR impossible in (13)? The only difference is that *Napoleon* has been replaced by *his commander*.

The other possibility is that the ambiguity in (10) arises because the quantifier in the matrix clause, *one soldier*, can, via whatever mechanism, take scope in the embedded clause in (10). On this way of looking at it, there must be some reason why *one soldier* taking lower scope is possible in (10) but blocked in (13). And it looks like there is one. In order for *his* to function as a bound variable, it needs to be c-commanded by the operator portion of the quantifier. If the scope of *one soldier* is fixed in the lower clause, then the requisite c-command relationship won't exist, since the quantifier will be adjoined to the *embedded* TP. It's only if *one soldier* is adjoined to the *matrix* TP that it will c-command *his*. So that explains why the change from *Napoleon* to *his commander* should make a difference. The need to bind the variable *his* in the matrix clause means that *one soldier* must take scope in the matrix clause and cannot take scope in the embedded clause, where it could then interact with *every battle*. So we predict that *one soldier* can only have wide scope in (13) relative to *every battle*.

The next question is what the mechanics behind the different scope possibilities are. Given the theory that we've got at the moment, (9) has a trace in the subject position of the embedded clause.

(14) [$_{CP}$ One soldier$_i$ is expected by Napoleon [$_{CP}$ [$_{TP}$ t$_i$ to die in every battle.]]]

By hypothesis, the trace is just a placeholder, a different element, syntactically speaking, from the quantified DP *one soldier*. So, in order to obtain the reading where *one soldier* takes its scope in the embedded clause only, we would need to postulate some kind of 'quantifier lowering' operation which moved *one soldier* at LF from the matrix Spec TP and adjoined it to the embedded TP.

However, there are a couple of reasons why literal 'quantifier lowering' doesn't look like the right thing to do. First, in Chapter 7 we discussed evidence for the Proper Binding Condition, which requires that traces be c-commanded by their antecedents. That principle is going to rule out the kind of quantifier-lowering at LF that we'd need. Perhaps more importantly though, it looks like quantifier lowering is also empirically the wrong way to go. If the ambiguity in (9) is the result of quantifier lowering, then the position of the trace is irrelevant. Quantifier lowering is just an operation that we have which allows a quantifier to take scope in a lower clause. It's just an accident that we happen to have a trace in exactly the clause we want to move to. We should be able to use quantifier lowering to put a quantifier wherever we like, whether there's a trace there originally or not, since the lowering movement is completely independent from the existence of the trace. But consider (15):

(15) One salesman$_i$ was expected [$_{CP1}$ t$_i$ to persuade Bill$_j$ [$_{CP2}$ PRO$_j$ to buy every product]]

Here, we've embedded an object control verb under passive *expect*. *One salesman* starts as the subject of *persuade* and raises to the specifier of the matrix TP, but there's no trace of *one salesman* in CP$_2$ [PRO to buy every product], just a PRO controlled by *Bill*. But if quantifier lowering is a possible operation, we might expect that (15) can be ambiguous, with *one salesman* undergoing hypothetical quantifier lowering to adjoin to the TP in CP$_2$, where it could then interact with *every product*. We then might expect that *every product* could have wide scope relative to *one salesman* – the interpretation where there are a multiplicity of salesmen who each persuade Bill to buy a different product.

(16)

Salesman no. 1 ⟶ Product no. 1
Salesman no. 2 ⟶ Product no. 2
Salesman no. 3 ⟶ Product no. 3
etc.

But this reading doesn't seem possible in (15). It seems as though *one salesman* must have wider scope than *every product*: there can be only a specific salesman who is expected to persuade Bill to buy all the products.

(17)

Salesman no. 1 ⟶ Product no. 1
Salesman no. 1 ⟶ Product no. 2
Salesman no. 1 ⟶ Product no. 3
etc.

Contrast (15) with the following sentence, where the quantifier *a billionaire* originates in TP$_2$ and then raises to the matrix specifier of TP:

(18) A billionaire$_i$ was believed [$_{TP1}$ t$_i$ to be about [$_{TP2}$ t$_i$ to buy every company]]

Certainly the interpretation in which *a billionaire* takes wide scope is the stronger one, but the interpretation in which *every company* has wider scope (so that the billionaires vary with the companies) is possible.

By now you can probably see where we're going with this. The difference between (15) and (18) is that, in (18), the existential quantifier has a trace position in the lowest clause, the one containing the universal quantifier. In (15), it doesn't. So this makes it looks like it's not a quantifier lowering operation which moves quantified DPs down to some lower position. Instead, the generalization seems to be that, essentially, quantifiers can undergo QR from any of the positions in their chain. As we saw in (13) above though, there might be other constraints, like the need to c-command a bound variable, which will constrain what positions in the chain you can do QR from.

There's a straightforward explanation for these effects if we assume that traces are, in fact, just copies of the moved element, so that the representation for (10) is really (19):

(19) [$_{\text{ForceP}}$ [$_{\text{TP}}$ one soldier was expected by Napoleon [$_{\text{TP}}$ [one soldier] to die in every battle]]]

Again, the change from traces to copies doesn't change any of the things we said in Chapter 6 about the nature of chains. So just as traces don't need an 'extra' θ-role, or an 'extra' Case, the copies don't either. It's the chain as a whole that is assigned Case and a θ-role.

Pronunciation of lower copies of chains

Another argument in favour of the copy theory of movement comes from cases where it appears as though a lower position in the chain, rather than the head of the chain, can be spelled out at PF. If lower positions in the chain are literally unpronounced copies (rather than 'traces', a totally different syntactic object), it becomes easier to imagine how this might be possible. Bošković and Nunes (2007: 17), for example, discuss multiple wh-fronting in Romanian. Romanian is like Polish (which we mentioned briefly in Chapter 7), and unlike English or Chinese, in that *all* wh-words move overtly to the specifier of CP (as we discovered in the last chapter, really FocusP), so you have contrasts like the one between (20a) and (20b):

(20) a. Cine ce precede?
who what precedes
'Who precedes what?'
b. *Cine precede ce.

The one exception is the case where having both wh-phrases initially would result in two consecutive homophonous wh-phrases, as in (21):

(21) *Ce ce precede?
what what precedes
'What precedes what?'

The intended meaning of (21) can only be expressed in Romanian by (22):

(22) Ce precede ce?
what precedes what
'What precedes what?'

Following Bošković (2002), Bošković and Nunes (2007) assume that there is no syntactic violation in (21). Rather, the ungrammaticality of (21) is due to a low-level PF constraint which simply bars consecutive homophonous wh-phrases. Now the question is: how do we analyse (22)? The 'obvious' possibility is that in (22) the second instance *ce* has never moved, and is still in its original D-structure position. But taking this approach to (22) runs into a number of conceptual problems, which are all sort of related to each other.

First, if there's no movement in (22), then our generalization about wh-movement in Romanian becomes less elegant. Instead of 'all wh-phrases in Romanian obligatorily move in the syntax to the specifier of FocusP', we have to say 'all wh-phrases in Romanian obligatorily move in the syntax to the specifier of FocusP, except where this would create two consecutive homophonous wh-phrases, in which case the second one doesn't move'. In effect, we end up with a situation where a syntactic operation is blocked by phonological considerations, which is kind of strange. Subjacency, for example, doesn't say that DP and IP are bounding nodes, except when the moved element begins with [p].

However, if you assume that 'traces' are really just unpronounced copies, then you can construct an appealing-looking alternative. You can explain (22) by saying that both instances of *ce* actually *do* move, but that, because of the PF prohibition against consecutive homophonous wh-phrases, you pronounce the tail of the direct object chain instead of the head of the chain. This is illustrated schematically in (23) (ignoring the irrelevant subject trace/copy):

(23) [ce ~~ce~~$_t$ precede ce$_i$]

This approach seems to solve all our problems. We can maintain the generalization that all wh-words in Romanian obligatorily move to the specifier of FocusP, but if you were to pronounce the heads of both of the chains, as in (21), you'd end up with an ill-formed PF. So, at PF, instead of pronouncing the highest copy/the head of the chain, which you would normally do, you pronounce the lower copy. Now we don't have a phonological problem 'interfering' with a syntactic operation. We have a PF solution to a PF problem. But we can only do that if what we've been calling 'traces' are actually complete copies of the moved element, ready to be pronounced if the need arises.

Pronunciation of intermediate copies

In addition to cases like (22), where it's a question of spelling out the tail of the chain instead of the head, you also have cases where *intermediate* copies can be spelled out. These are particularly interesting because, in addition to providing evidence for the copy theory of movement, they also provide evidence for the concept of successive cyclic movement and intermediate traces in general.

The 'wh-copying' construction in German provides one such example (see Bošković and Lasnik 2007: 356) and Bošković and Nunes (2007: 48ff.) for

discussion and references). There's a fair amount of idiolectal and dialectal variation with this construction, but in those cases where it's allowed, copies of some wh-phrases can appear in what look like intermediate positions, as in (24):

(24) *Wen* glaubt Hans *wen* Jakob gesehen hat?
whom thinks Hans whom Jakob seen has
'Who does Hans think Jakob saw?'

As expected, the head of the chain in the specifier of FocusP is pronounced, but there's also another pronounced copy at the beginning of the embedded clause. As (25) shows, the extra copy is in the normal position for embedded wh-words in German:

(25) Ich weiss wen Jakob gesehen hat
I know whom Jakob seen has
'I know who Jakob has seen'.

So what it looks like is happening is that, in addition to the head of the chain being pronounced, the intermediate trace in the intermediate FocusP is also being pronounced. So not only are the head of the chain and the tail of the chain copies, it looks like intermediate traces are also copies as well.

10.2 A movement analysis of control

Having argued in the previous section that traces don't exist and that they're actually copies, I want to move on to a similar topic in this section – the idea that PRO doesn't actually exist, and that many places where we postulated PRO can be better explained by assuming movement and copies.

Recall from Chapter 4 that we introduced PRO to account for sentences like (26), in the context of the θ-criterion:

(26) John promised to leave.

The θ-criterion says that there needs to be a one-to-one correspondence between the θ-roles that a predicate assigns and the recipients of those θ-roles.

(27) *The θ-criterion*
All syntactic arguments of a predicate must be assigned a θ-role, and all θ-roles of a predicate must be assigned.

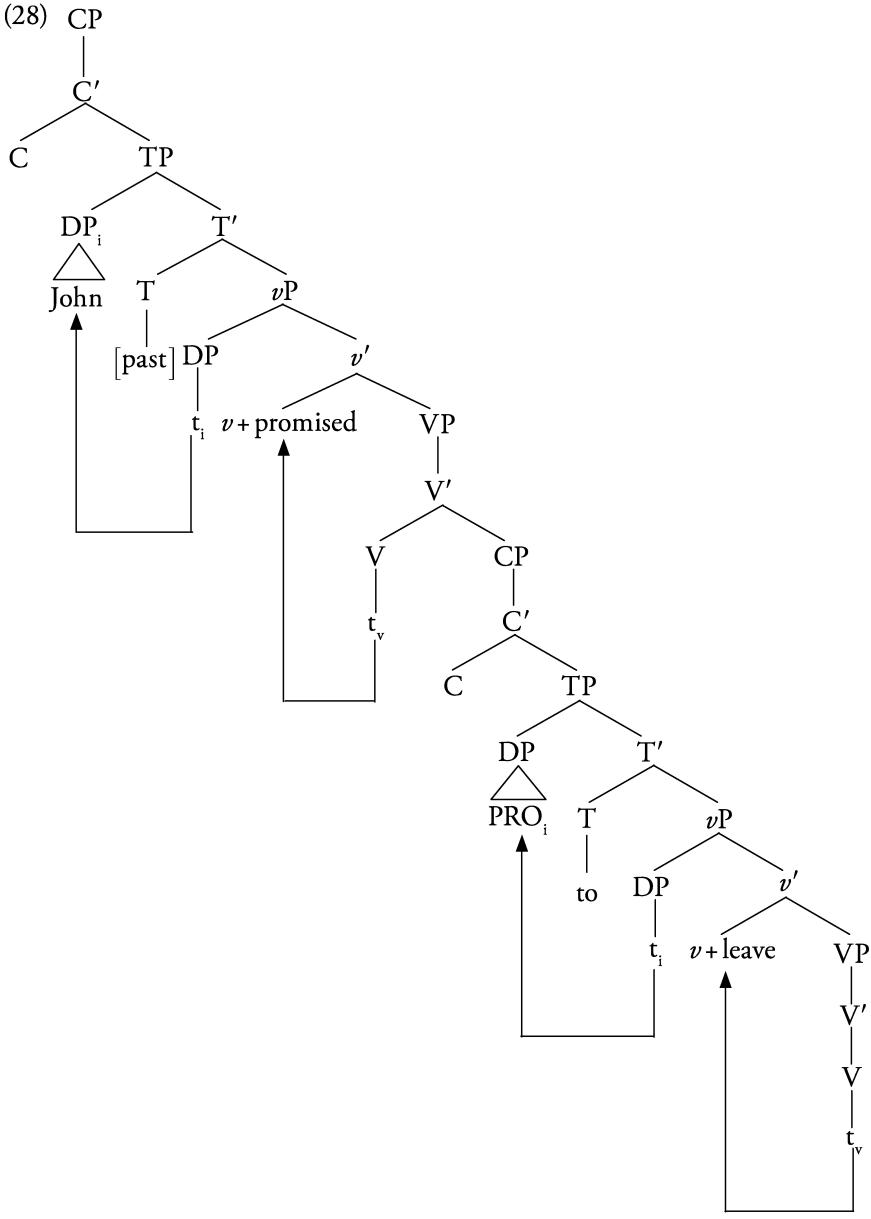
The problem was that the verb of the tenseless embedded clause, *leave*, had an Agent θ-role that, at least on the surface, it looked like it wasn't assigning. We solved that problem by assuming that there was a null DP, which we called PRO, which was the subject of the embedded clause and which was assigned the Agent θ-role of *leave*. This element PRO was 'controlled' by John, since (26) can only mean that John promised that John would leave, not that John promised that some unspecified person would leave or that people in general would leave.

What I want to do in this section is discuss some arguments given by various people starting in the mid-1990s, particularly O'Neil (1995, 1997) and

Hornstein (1999), to the effect that PRO can be eliminated and replaced by movement. Since we hadn't discussed movement yet when we introduced PRO back in Chapter 4, we didn't consider this option, but what this does essentially is make (at least some instances of) control look more like raising.

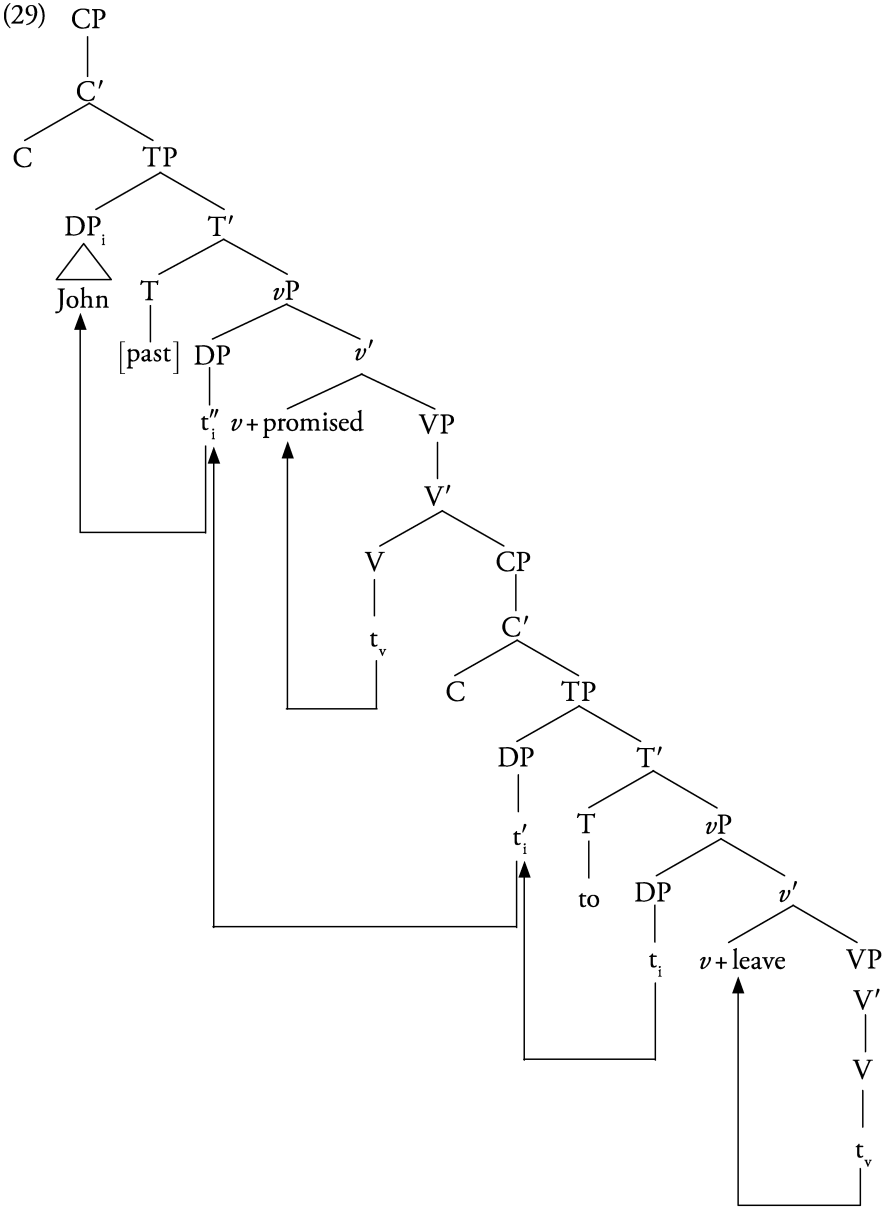
To begin, let's look at a subject-control verb like *promise*. Without assuming control-as-movement, we would have expected something like (26) to have the derivation in (28).

(28)



However, under the movement analysis of control, there is no PRO. *John* is base-generated as the subject of the TP *to leave*, and then *raises* to the specifier of the matrix clause via the matrix VP, as in (29).

(29)



This will require revising a few things, most obviously the θ-criterion, since *John* now gets two θ-roles, but it looks like there are some reasons to think it might be a good idea, which we'll now discuss.

Obligatory vs non-obligatory control and the interpretation of PRO

You may have noticed back in Chapter 8 that we spent a fair amount of time talking about the distribution of PRO, where we argued that it followed from Binding Theory. If PRO is simultaneously both an anaphor and a pronoun, then it can't have a governor. And if it can't have a governor, then it can only appear as the subject of infinitival phrases with CP. This explains the *distribution* of PRO really nicely, but what we didn't talk so much about is PRO's *interpretation*, as it's here that the theory is a little less successful, as Hornstein (1999: 72) notes. All we really did was stipulate that there were subject control verbs, object control verbs and arbitrary PRO and said something kind of vague about how this followed from the fact that PRO had both anaphoric and pronominal tendencies. What I want to do now is look at the interpretation of PRO in a little more detail, because it's in trying to clarify things in this area that we see some arguments for a movement analysis of control.

Since the mid-1980s, there's been a generally accepted distinction between obligatory control and non-obligatory control. It's in obligatory control contexts where PRO looks *interpretively* like an anaphor, whereas in non-obligatory control contexts it looks interpretively more like a pronoun.

Let's take a look at some of the ways in which obligatory control contexts and non-obligatory control contexts are different. First, in an obligatory control context, like subject-control, PRO must have an antecedent. It cannot have 'arbitrary' reference:

(30) John_{i} promised $\text{PRO}_{i/*arb}$ to go home.

In (30), PRO *must* have an antecedent, and that antecedent is *John*. (30) cannot mean 'John promised that people in general would go home', for example, and we can see that clearly if we put a reflexive in the infinitival clause:

(31) John_{i} promised $\text{PRO}_{i/*j}$ to behave himself_{i}/*oneself_{j}/*themselves_{j}.

This isn't required in non-obligatory control contexts such as (32), where PRO is in an extraposed subject. It is possible here to get arbitrary PRO:[2]

(32) It was believed (in the 1950s) that PRO_{i} shaving oneself_{i} was important.

Second, not only does PRO have to have an antecedent with obligatory control, but the antecedent also must c-command PRO, as (33) illustrates:

(33) *[John_{i}'s mother]$_{j}$ promised PRO_{i} to behave himself_{i}.

PRO can only take the c-commanding DP *John's mother* as its antecedent. It can't take just *John*. This is not the case with non-obligatory control. PRO can take either DP as its antecedent:

(34) [John$_i$'s mother]$_j$ said that PRO$_{i/j}$ behaving himself$_i$/herself$_j$ was important.

Furthermore, it looks like there's a kind of locality effect with obligatory control, even where you have c-command.

(35) *John$_i$ thought that it had been promised PRO$_i$ to pay himself$_i$ a lot of money.

Even though *John* c-commands PRO, it seems to be 'too far away' for PRO to take it as its antecedent. In non-obligatory control contexts, distance is not a problem:

(36) John$_i$ knew that it was time PRO$_i$ to pay himself$_i$ a lot of money.

So, in obligatory control contexts, the fact that PRO needs a c-commanding antecedent which is 'not too far away' makes PRO sound very much like some kind of anaphor. But this isn't the case in non-obligatory control contexts.

Obligatory control contexts also have a couple of other characteristics that are worth mentioning, although they're not directly related to any anaphoric qualities of PRO, because they make a movement analysis of 'obligatory control' look appealing, as we'll see.

First, obligatory control contexts do not allow PRO to have a so-called 'split' antecedent, where PRO collectively refers back to several different DPs at the same time. (37) illustrates an object-control context:

(37) *John$_i$ told Mary$_j$ PRO$_{i+j}$ to wash themselves$_{i+j}$/each other$_{i+j}$.

In (37), PRO can't take both *John* and *Mary* as a sort of collective antecedent, but, again, this changes in a non-obligatory control context:

(38) John$_i$ told Mary$_j$ that PRO$_{i+j}$ washing themselves$_{i+j}$/each other$_{i+j}$ would be fun.

Second, obligatory control contexts allow only a 'sloppy' identity interpretation for PRO under ellipsis, whereas non-obligatory control contexts allow a 'strict' identity interpretation.

To see what's meant by the 'sloppy' identity vs 'strict' identity interpretation in the context of ellipsis, consider (39):

(39) John likes his mother and Bill does too.

The second conjunct contains an elided VP *like his mother*, and the presence of a pronoun in the elided VP results in the sentence as a whole being ambiguous.

One interpretation (39) can have is 'John likes his own mother, and Bill likes his own mother too'.

(40) John$_i$ likes his$_i$ mother and Bill$_j$ does [like his$_j$ mother] too.

This is called the 'sloppy' identity interpretation because the pronoun *his* refers to two different people in the two cases. In the first (overt) conjunct, *his* refers to *John*, but in the second (elided) conjunct, *his* refers to *Bill*.

However, (39) can also mean 'John likes his own mother, and Bill likes John's mother too', so that the indexation is not the one in (40), but rather the one in (41):

(41) John$_i$ likes his$_i$ mother and Bill$_j$ does [like his$_i$ mother] too.

This is the 'strict' identity interpretation, so-called because the pronoun *his* refers to the same person, *John*, in both the overt conjunct and the elided conjunct. With that in mind, let's take a look at the possible interpretations when you have an elided VP containing PRO. In obligatory control contexts, only the 'sloppy' identity reading is possible:

(42) John promised PRO to win and Bill did too.

(42) can only mean that John promised that *John* would win and Bill promised that *Bill* would win. It can't mean that *Bill* promised that *John* would win. In other words, the elided PRO can only be controlled by *Bill*.

(43) John$_i$ promised PRO$_i$ to win and Bill$_j$ did [promise PRO$_j$ to win] too.

However, in non-obligatory control contexts, the 'strict' identity interpretation is possible:

(44) John$_i$ thinks that PRO$_i$ behaving himself is crucial and Bill does too.

Like (43), (44) *can* have a 'sloppy' identity interpretation, but it also allows the 'strict' identity interpretation, meaning that Bill thinks that for John to behave himself is crucial.

(45) John$_i$ thinks that PRO$_i$ behaving himself is crucial and Bill$_j$ does [think that PRO$_i$ behaving himself$_i$ is crucial] too.

A movement account of obligatory control

Hornstein (1999) notes that all of these facts about obligatory control that we've been talking about – (1) the fact that PRO needs a c-commanding antecedent which can't be 'too far away', (2) the inability of PRO to take a 'split' antecedent and (3) the fact that PRO allows only the sloppy identity reading

under ellipsis – would fall out completely naturally if obligatory control were in fact the result of movement, just like raising:

(46) a. John seems [John] to be here.
b. John promised [John] to be here.

First, with respect to the 'distance' issue, traces/copies need to have a c-commanding antecedent and are constrained by locality. Sentences such as (47), which we've seen a couple of times in the book already, illustrate so-called 'superraising', and are ungrammatical:

(47) *John seems that it is likely [John] to leave tomorrow.

But this is precisely the configuration from (35) above, where we saw that *John* couldn't control PRO.

Reinterpreting obligatory control as movement would also explain why it doesn't allow 'split' antecedents. It's just literally not possible. You couldn't have some kind of conjoined element *John/Mary* which then moved and split in half, with the *Mary* part moving to object position and the *John* part moving to subject position.

(48) *John_i told Mary_j [John_i/Mary_j] to wash each other$_{i+j}$.

It just can't be done.

Finally, observations about 'sloppy' and 'strict' identity fall together if obligatory control is an instance of movement. Recall from (42) above that in obligatory control contexts, only the 'sloppy' identity interpretation was possible, and that seems to parallel what happens when you have ellipsis with raising:

(49) John seemed to be happy and Mary did too.

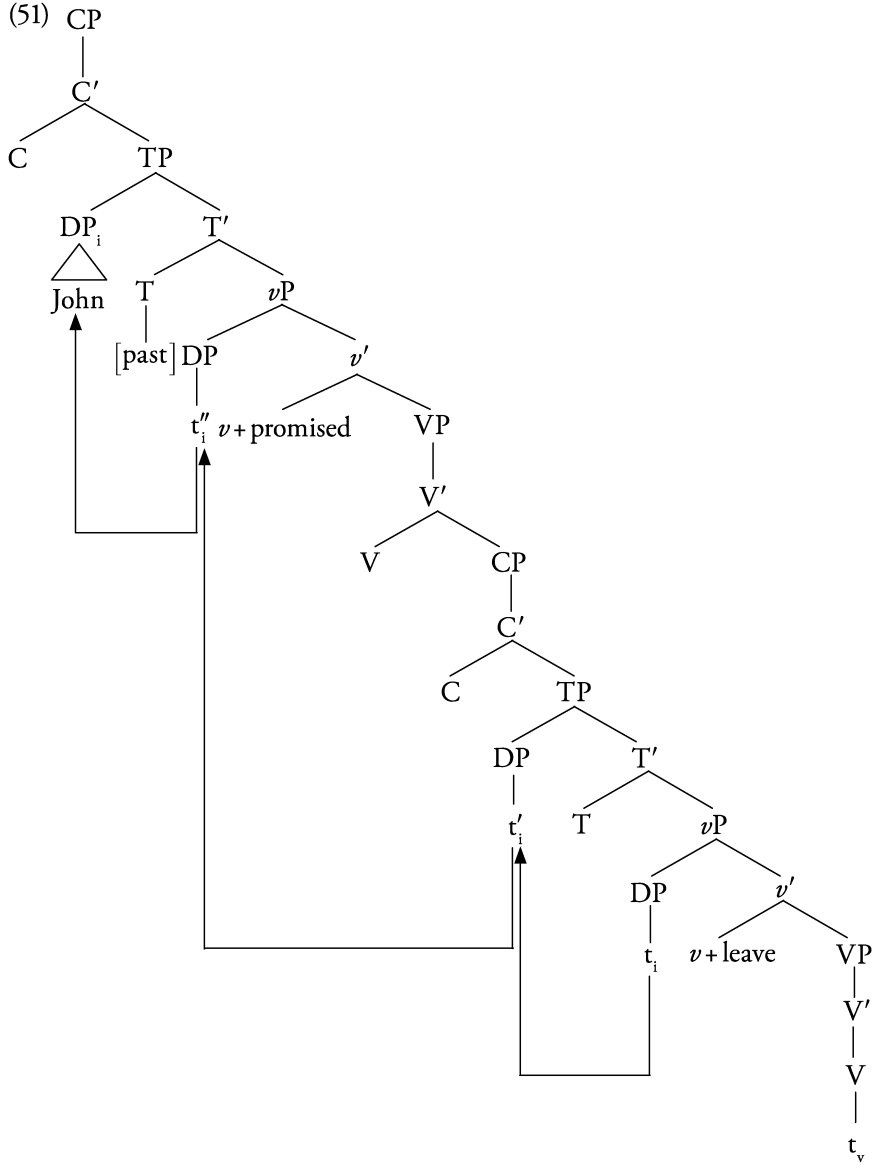
(49) can only mean that John seemed that John was happy and Mary seemed that Mary was happy, which is the 'sloppy' identity interpretation for the trace/copy:

(50) John seemed [John] to be happy and Mary did [seem [Mary] to be happy] too.

The 'strict' identity interpretation is just impossible. (49) can't mean that Mary seemed that *John* was happy (in whatever way Mary might display that). That interpretation is, admittedly, a little strange, but I don't think it's incoherent, so you can't explain the impossibility of the strict identity reading on semantic/pragmatic grounds.[3]

So, in terms of spelling this out with actual derivations, how does control-as-movement work? Pretty much as you'd expect:

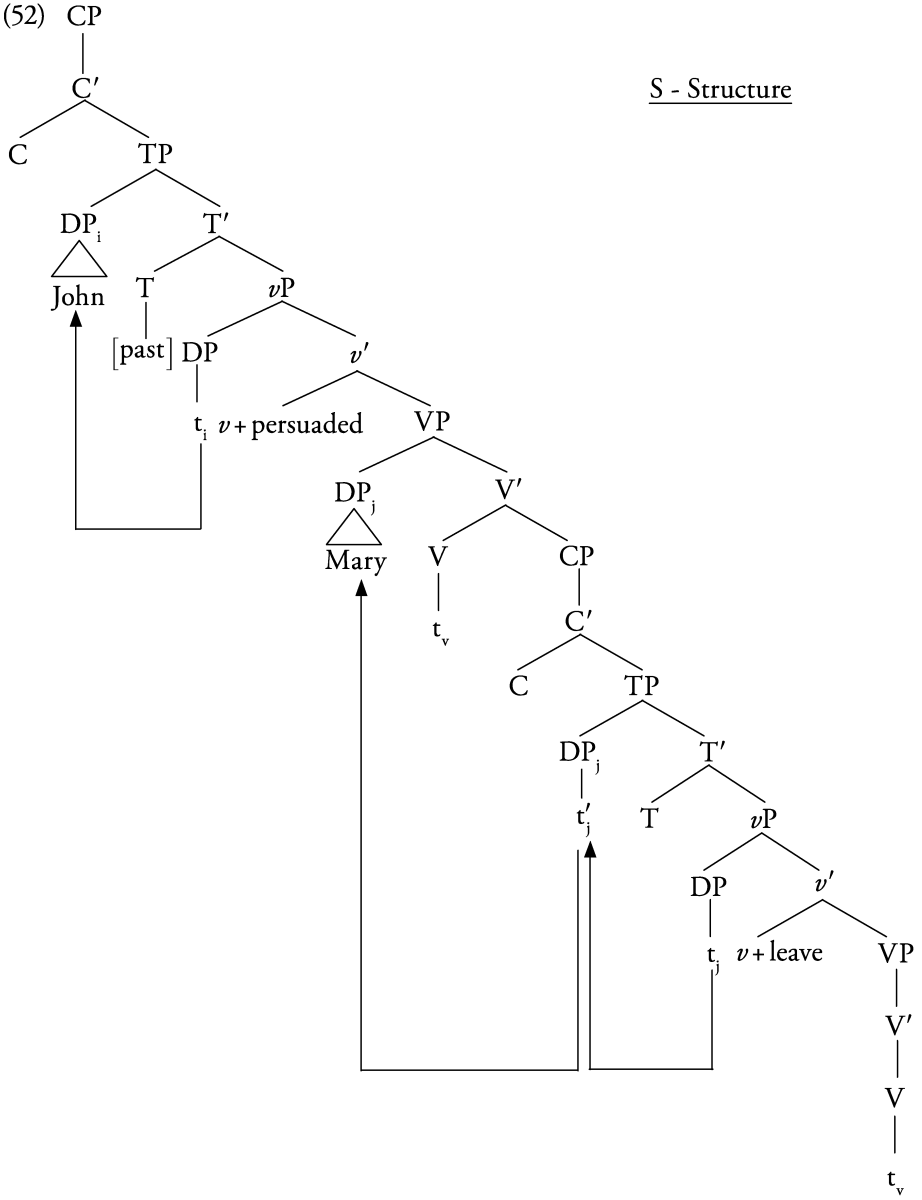
(51)



Remember from the previous chapter that we've now got VP-shells to consider, so the VP *leave* is immediately dominated by *v*P, and *John* starts life as the specifier of the embedded *v*P. That's how it gets the Agent θ-role of *leave*. It then raises to the specifier of the defective embedded TP because of the Extended Projection Principle. It then raises to the specifier of the *v*P *promised [John] to [John] leave* and picks up the Agent θ-role of *promise*, and moves from there to the specifier of the matrix TP, again because of the EPP.

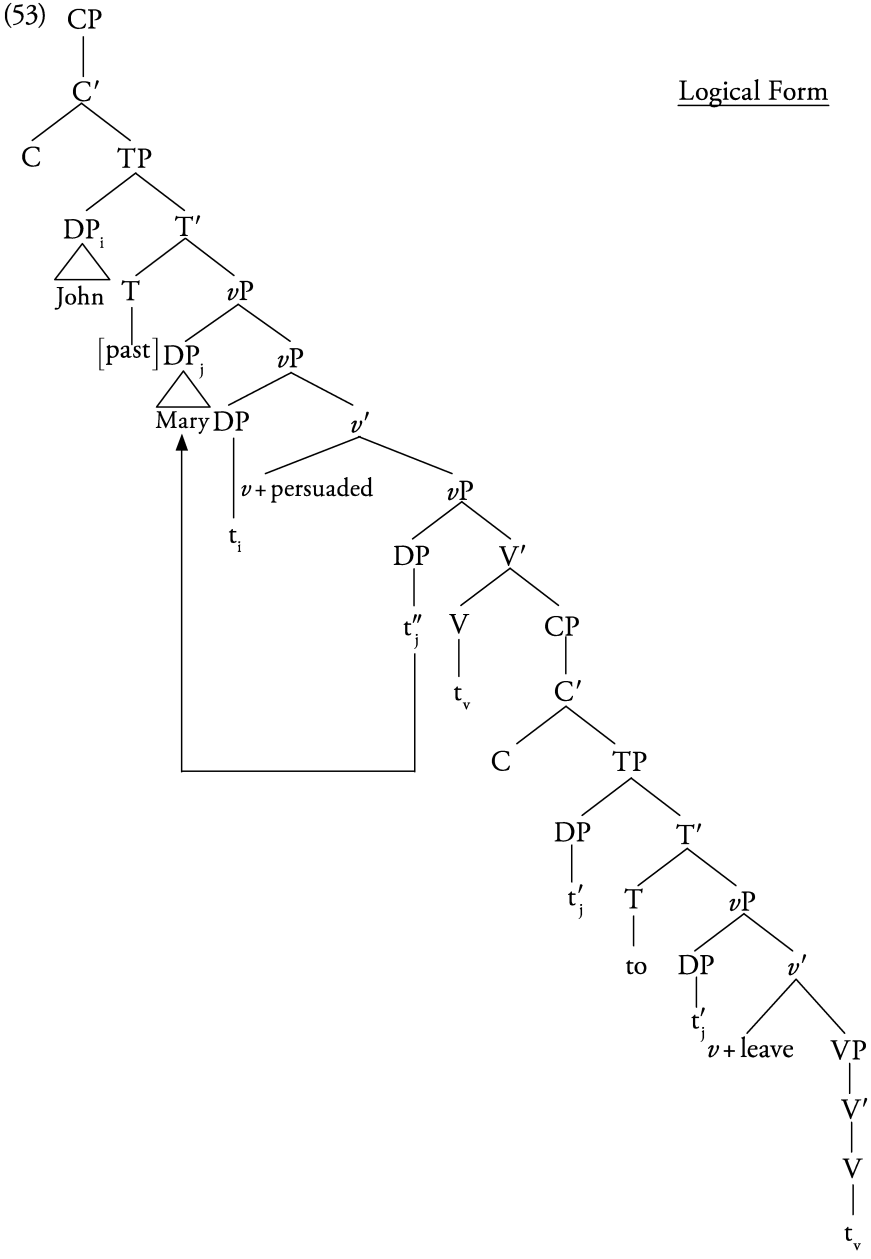
(Obligatory) object control also works pretty much as you'd expect:

(52)



Mary, just like *John* in the previous example, starts as the specifier of the embedded *v*P, where it gets the Agent θ-role of *leave*. Again it then raises to the specifier of the embedded TP to satisfy the EPP. However, now *Mary* is marked for Accusative Case and needs to get the Patient θ-role of *persuade*. To do that, it moves to the specifier of VP. Finally, as we discussed in the previous chapter, in order to be assigned Accusative Case, it moves at LF to the outer specifier of *v*P.

(53)



Logical Form

Modifying the θ-criterion

As mentioned above, collapsing obligatory control and movement means that we're going to have to modify the θ-criterion just a little bit. As we discussed at the beginning of the section, the θ-criterion enforces a one-to-one correspondence between arguments of a predicate and semantic roles that the predicate makes available. If we now want to say that there's *movement*

in obligatory control, we can't have this, since we have DPs (*John* in (51) and *Mary* in (52) respectively) which get assigned two θ-roles. We need the θ-criterion so that DPs get assigned *at least one* θ-role (we still don't want to generate **John seems that Bill left*), and we also need to make sure that all θ-roles get assigned to *something* (so that predicates can't just decide not to assign a θ-role).

(54) Revised θ-criterion: Every θ-role must be assigned and every argument chain must have at least one θ-role.

However, in addition to 'forcing' us into an account of non-finite clauses in terms of PRO, as we discussed back in Chapter 4, there's at least one other construction that the old version of the θ-criterion accounted for: the impossibility of sentences like (55) and (56):

(55) *John hit [John].

(56) *John believes [John] to be a great actor.

We want to make sure that we don't predict, for example, that (55) can mean *John hit himself*, with *John* being assigned both the Agent and Patient θ-roles by *hit*.

As usual, the way forward is to look at the differences between sentences like (55) which we want to rule out and sentences like (26) which we want to allow. Is there a difference between the two that we can hang something on? Well, it looks like there is: Case. Since non-finite Ts don't assign Nominative Case, the control-as-movement examples all involve chains with only one Case-marked position (Spec TP for subject control verbs and the outer specifier of *v*P for object control verbs). (55) and (56) are different, in that they both involve movement from one Case position (Accusative) to another (Nominative), so the chains in (55) and (56) contain *two* Case-marked positions.

If we assume that Case is a sort of lexical feature that DPs have, on analogy with person and number features, we could use this as a way of prohibiting sentences like (55) and (56). This is essentially the approach taken in Hornstein (1999). It seems reasonable to assume that elements can't come out of the lexicon with conflicting features (or two features of the same type, or however you want to think about it). Just as a verb can't come out of the lexicon marked for both first person plural and third person singular, we could say that a DP can't come out of the lexicon bearing features for both Nominative and Accusative case. (Remember that just because English doesn't morphologically mark these things doesn't mean that the DPs don't have them.) That will then rule out (55), because there's only one DP *John* and two Case features that need checking, and *John* can't be marked for both of them. In control contexts, the DP, despite getting two θ-roles, the DP only ever has one Case feature to check (Nominative in the case of subject control and Accusative in the case of object control).

Non-obligatory control

Having discussed the motivation behind treating obligatory control as movement, I want to close the discussion in this section by looking at what's left: the account of non-obligatory control. We observed above that while PRO in obligatory control contexts behaved like an anaphor from the point of view of interpretation, it behaved like a pronoun in non-obligatory control contexts, and this is precisely the direction that Hornstein (1999) takes. In non-obligatory contexts, PRO is in fact *pro*, the null pronominal of null-subject languages.

First, let's remind ourselves of the some of the non-obligatory control contexts that we were talking about earlier:

(32) It was believed (in the 1950s) that PRO shaving oneself was important.

(34) [John$_i$'s mother]$_j$ said that PRO behaving himself/herself was important.

(36) John$_i$ knew that it was time PRO$_i$ to pay himself$_i$ a lot of money.

(38) John$_i$ told Mary$_j$ that PRO$_{i+j}$ washing themselves$_{i+j}$/each other$_{i+j}$ would be fun.

In non-obligatory control contexts, as we mentioned, the antecedent of PRO *doesn't* have to c-command PRO (34), be 'close' to PRO (36) or even exist at all (32). Non-obligatory control also allows 'split' antecedents (38).

As Hornstein (1999: 91) notes, this is the precisely the range of possible interpretations that you'd expect a pronoun to have.

(57) [John$_i$'s mother]$_j$ said that his$_i$ behaving himself$_i$ was important.

(58) John$_i$ told Mary$_j$ that their$_{i+j}$ leaving together would cause a scandal.

Furthermore, the position of PRO in non-obligatory control contexts looks like a position from which movement is not possible, typically the subject position of a sentential subject. For example:

(59) *John was believed that [John] leaving tomorrow was a good idea.

(60) *John seemed that it was time [John] to leave tomorrow.

Other non-obligatory control contexts are also typically places from which movement is impossible, such as wh-islands:

(61) a. John told Sam how PRO to hold oneself erect at a royal ball. (Hornstein 1999: 92)

b. *Who did John tell Sam how [who] to hold himself erect at a royal ball?

All of these facts would fall together if the element in the specifier of TP in non-obligatory control contexts were a null pronoun.

10.3 Introduction to verb movement

Finally, I want to finish the discussion of movement by going back and picking up a topic that we only mentioned in passing back in Chapter 6. Since the introduction of the concept of movement in that chapter, we've seen many different kinds of movement taking place both between D-Structure and S-Structure (overt) and between S-Structure and LF (covert). Now all of these movements have something in common. They consist of XP movement; that is to say, movement of maximal projections. I now want to come back and say a little more about the movement of *heads*, also known as X°-movement.

A review: raising from T-to-C

The most frequent type of head movement, and the one which we'll be entirely concerned with here, is verb movement. We saw our first instance of verb movement briefly back in Chapter 6, when we introduced wh-movement in English. In English there are S-Structure word order alternations between a wh-question and its corresponding declarative, as illustrated in (62):

(62) a. What can John buy?
b. John can buy a book.

We know that the wh-word in (62a) must be in Spec CP, and we also know the subject *John* must be in Spec TP (the structural position for subjects). Based on this information, we can feel pretty confident in saying that the auxiliary has moved from T to C. The word order tells us that the auxiliary must be somewhere between Spec CP and Spec TP, and the only available position is the empty head of the CP, C°. Since C° is a head position and the auxiliary *can* is a head, there's no phrase-structure incompatibility in moving it from T to C.[4]

The logic is pretty obvious, but it's worth making it explicit because it illustrates a standard procedure for figuring out where an element has moved. Rather like navigating across the ocean, what you need to do is figure out where your fixed points of reference are and then triangulate based on them. So, in the above example, our fixed reference points are the wh-word in Spec CP and the subject in Spec TP. We know that those elements are in those positions. Therefore, based on the word order, the position of the auxiliary *can* in the phrase-structure tree must be somewhere where it can end up between *what*

and *John*. And, because we're talking about the movement of a head, the position that it's moving to must be a head position. In this case, there's only one position which fits the requirements, and that's C.

Raising from V to T

In addition to raising from T to C (in questions, for example), we also have instances where verbs raise from V to T. French is a language which is commonly used to illustrate this kind of verb movement. In French, main verbs obligatorily appear to the left of both VP-oriented adverbs and sentential negation. They may not appear to the right of either. (The analysis of negation in French is complicated slightly by the fact that the standard dialect contains two negative elements *ne* and *pas*. Following normal practice, we'll assume that it's *pas* which represents the sentential negation. We'll take *ne* to be a part of the verb.[5])

(63) a. Jean *va* souvent a l'école.
Jean goes often to the school
'Jean often goes to school.'

b. *Jean souvent *va* *a* l'école.
Jean often goes to the school.

(64) a. Jean ne va pas au magasin.
Jean *ne* go not to the store
'Jean isn't going to the store.'

b. *Jean ne pas va a la magasin.

Like Spec CP and Spec TP, adverbs and negation are often used as fixed points of reference to help us navigate in the wide-open seas of movement and word order. It's standardly assumed that so-called 'VP-oriented' adverbs (which we'll discuss in more detail below) establish the left-hand boundary of *v*P. (The usual assumption, which we'll adopt here, is that these adverbs are adjoined to *v*P. Negation, on the other hand, is normally assumed to head its own function projection between TP and *v*P.) Therefore, the fact that the verb in French must appear to the left of these adverbs and *pas* (as (63) and (64) illustrate) is taken to show that they've raised out of the *v*P to some higher position. In this subsection, we'll examine this issue in detail, and look at some similarities and differences between French and English with respect to verb movement.

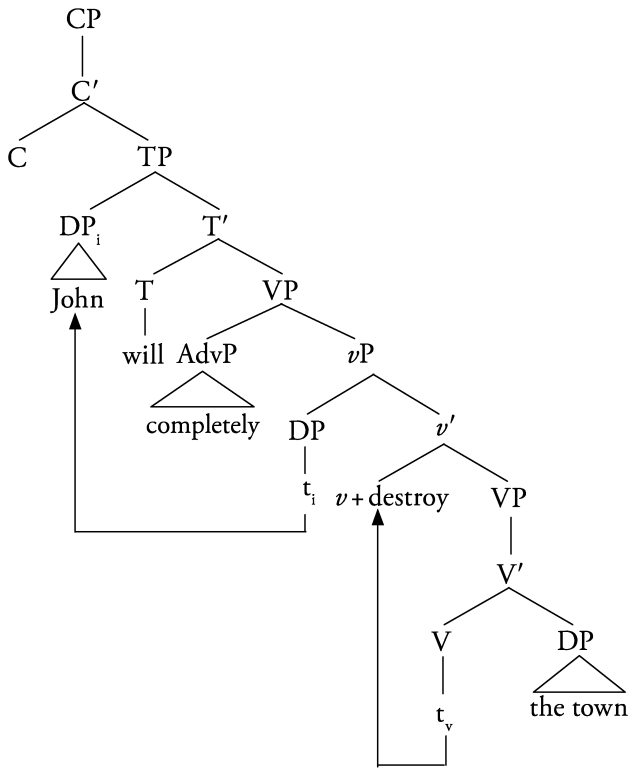
Adverbs and negation: two more reference points

Let's start by looking at *VP-oriented* adverbs. These are adverbs which specifically modify the verb phrase, as opposed to the entire sentence. Such a canonical VP-oriented adverb is *completely*. VP-oriented adverbs are one of the key reference points when looking at V raising to T because they're

usually assumed to be adjoined to *v*P. Therefore, the sentence in (65) has the tree in (66):

(65) John will completely destroy the town.

(66)

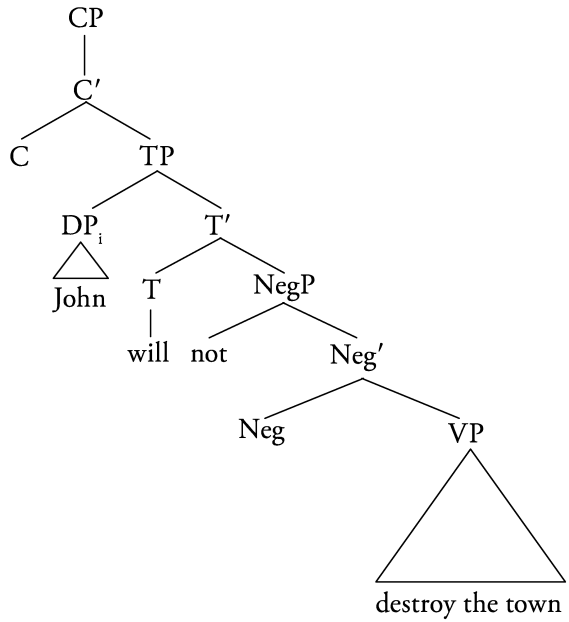


Completely is serving as an adverbial modifier to the *v*P *destroy the town*. We would therefore expect it to be a sister to the *v*P, which it is under the VP-adjoined analysis. With respect to the linear order, *completely* comes to the right of the auxiliary in T and to the left of the verb in VP, which is also as expected. (VP-oriented adverbs are different from *sentential* adverbials such as *probably*, which modify the entire sentence and are not always adjoined to *v*P.)

Sentential negation is also standardly used as a diagnostic for the left edge of *v*P. The only difference between negation and adverbs is that negation is standardly assumed to have its own functional projection NegP, rather than being simply adjoined to *v*P. So the tree for a sentence like (67) would be (68):

(67) John will not destroy the town.

(68)



So in the context of the question of whether a verb has raised from V to T, VP-oriented adverbs and negation play a key role. If the verb appears to the left of the adverb/negation, then we assume that it is in a higher position like T. On the other hand, if the verb appears to the right of the adverb/negation, then we assume that it is within the *v*P.

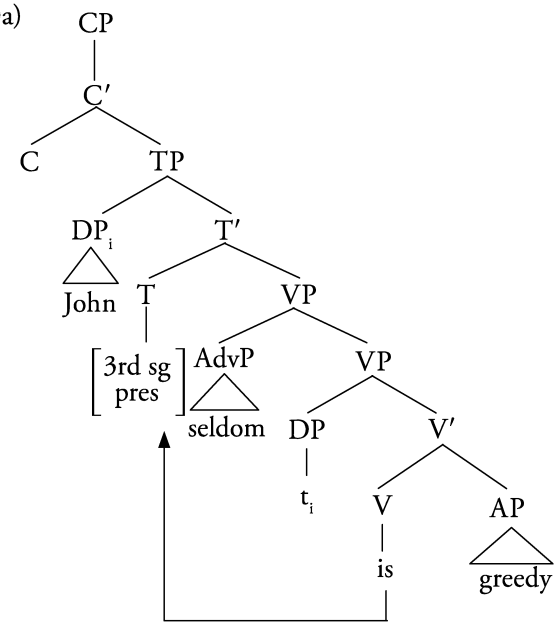
V-to-T raising in English and French

Let's now take a look at some facts about French and English verbs, and where they appear relative to the fixed points that we've been discussing. Compare the following English sentence with its French translation:

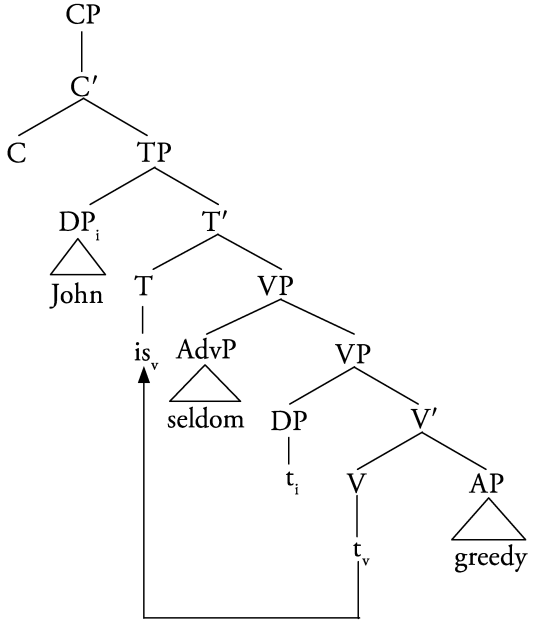
(69) a. John is *seldom* greedy.
b. Jean est *rarement* avide.
J. is seldom greedy.

If *seldom* is adjoined to *v*P both in French and in English, and *John/Jean* is in Spec TP, then the verb *is/est* must therefore be in T. But could it have been base-generated there at D-Structure? That would be unlikely. If it were, we would be left with a verb phrase with no verb head – which is not possible. So we must assume that *is* was base-generated at D-Structure as the V head of VP, and then moved via *v* to T at S-Structure. That gives us the representations in (70) for D-Structure and S-Structure (which are the same for French and for English, aside from the words of course).

(70a)



(70b)



So (70) shows the raising of *be* from V to T both in English and in French. In a yes/no question, V to T raising combines with T to C raising, resulting in V-to-T-to-C raising:

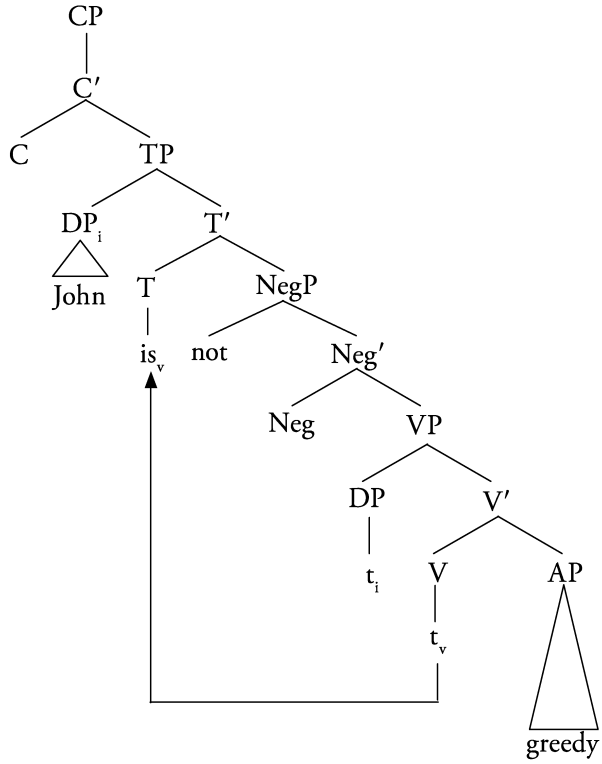
(71) $[_{CP}$ Is$_v$ $[_{IP}$ John t′$_v$ $[_{VP}$ seldom $[_{VP}$ t$_v$ greedy]]]]?

Given the situation in (70), we'd expect the same facts to obtain when we turn from adverbs to negation, and our prediction seems to be borne out. In English, auxiliaries in T appear to the left of negation and not to the right:

(72) a. John is *not* greedy.
b. *John *not* is greedy.

The only difference in the structure between adverbs and negation is that negation is assumed to head its own functional projection, NegP. So instead of a tree as in (70), we have (73):

(73)

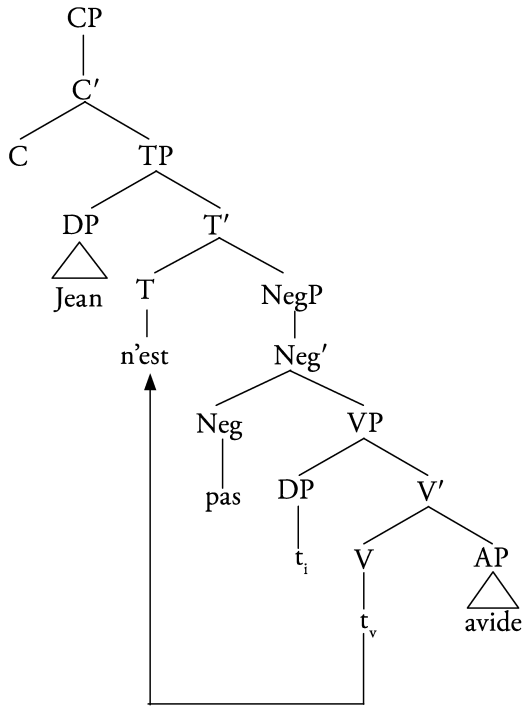


In French, the grammatical translation of (72a) is (74):

(74) Jean n'est *pas* avide.
Jean *ne* is not greedy

As expected, the verb *to be* appears to the left of negation, suggesting again that it has raised from V to T, as illustrated by (75):

(75)



Again, this raising is obligatory, as illustrated by (76):

(76) *Jean ne *pas* est avide.

It therefore seems that, with respect to the verb *to be*, both English and French behave identically. The adverb and negation placement tests suggest that in both cases the verb obligatorily raises from V to T.

French and English differ however with respect to the behaviour of other verbs. Main verbs in English, for example, do *not* raise to T, but in French they do. Consider first the adverb placement situation, this time using another VP-oriented adverb *completely*:

(77) a. John *completely* answered the question.
b. *John answered *completely* the question.

(78) a. Jean répondait *complétement* à la question.
Jean answered (imperf) completely the question
'Jean completely answered the question.'
b. *Jean *complétement* répondait à la question.

In (77), we see that the English main verb appears to the right of the adverb and may not appear to its left. In French, we have the reverse situation. This seems to show that the main verb in French must raise to T, but cannot in English. Just as we'd expect, French main verbs can also undergo further raising to C if there is a yes/no question, but English main verbs do not.

(79) a. Répondait-il à la question?
b. *Answered he the question?

(Interestingly, the fact that (79b) may feel 'Shakespearian' to you is a reflection of the fact that, at an earlier stage of its history, English was like French in allowing main verbs to raise from V-to-T and from T-to-C.)

Turning now to negation, the word-order facts with respect to verb raising are identical to what we've seen with adverbs. As expected, English main verbs may not raise to T past the NegP:

(80) a. *John goes *not* to the store.
b. John does *not* go to the store.

Precisely here we see the 'dummy' verb *do*, which we saw in Chapter 6, which is needed because the full verb is not raising from V to T. French verbs, on the other hand, not only *can* appear to the left of NegP, they *must* appear there:

(81) a. Jean ne va *pas* au magasin.
Jean ne goes not to the store
'Jean doesn't go to the store.'
b. *Jean ne *pas* va au magasin.

From this, we conclude that French main verbs must be raising past the NegP from V to T.

The copula verb *to be* in both French and English appears to the left of adverbs and negation, corresponding to raising from V to T. However, there appears to be a difference with respect to main verbs. In English they may not raise from V to T, but in French they must.

LF V-raising and a 'checking' theory of morphology

In comparing French and English verb raising, we've seen that French verbs raise from V to T, but full verbs in English do not. But we need to qualify this statement. English main verbs don't raise from V to T at S-Structure, but they do at LF.

Remember that back in Chapter 7 we saw instances of XP movement which took place between S-Structure and LF – that is, (covert) XP movement that you don't 'see'. There's no a priori reason why the same shouldn't be true of head movement. The possibility should exist that there are head movements that you don't 'see' because they take place between S-Structure and LF. English verb raising is such a case.

First, there are good conceptual reasons for thinking that English verbs *do* raise at LF from V to T. Throughout the book, we've been assuming that I contained two important items: the subject-verb agreement information and the tense information. One reason, that we mentioned back in Chapter 3, was that this information seems to be left behind in VP deletion/ellipsis constructions – (82), for example:

(82) John practises for three hours every day and Bill does too.

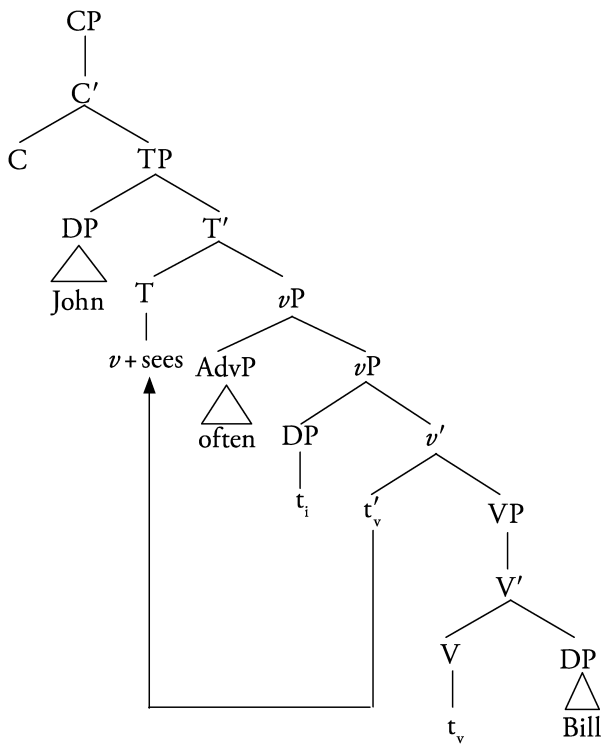
In (82), the VP *practise for three hours every day* has been elided, but the agreement and tense information seem to have been left behind. They're not deleted, but instead are realized on the 'dummy' verb *do*. It shows third person singular agreement (it has *does* and not *do*) and it is in the present tense (*does* and not *did*). We took this as an argument that this agreement and tense information was in I (as it was then) and not V, and therefore did not get elided.

But at the same time, this information *is* realized on the verb when there's no auxiliary. In the first conjunct of the above sentence, the verb *practises* is also third person singular and present tense. If that verb is in V and the agreement and tense information is in T, how do the two things ever get hooked up? In French, you have V-to-T raising in the syntax, so it's not really a problem. But how does English work?

There's an obvious answer. It's not that what comes out of the lexicon is *practise*, which then merges somehow with the agreement and tense information to give you *practises*, *practised* or whatever. Instead, the verb is selected from the lexicon 'fully formed', for example as *practises*, and then, after V-to-T raising, it checks its tense and agreement features with the features in T. If they match, everything's OK. If they don't match, then the sentence is ungrammatical. In French, this happens overtly between D-Structure and S-Structure, because we see the effects of V-to-T in the overt linear order of the words. But in English this checking process happens at LF. So, therefore, an English sentence like (83) really has (84) as its LF:

(83) John [$_{VP}$ often [$_{VP}$ sees Bill]].

(84)



This is a nice conclusion because it dovetails perfectly with some of the arguments that we gave for LF movement of XPs back in Chapter 7. LF encodes those aspects of the syntax which are relevant for interpretation. But it's hard to imagine that there's any difference with respect to LF between a simple English declarative sentence and its French translation. But if we don't postulate LF raising, then the LFs *will* be different. For example, negation would have wide scope relative to the verb in English because negation c-commands the verb. In French, on the other hand, the verb would have wider scope than negation (because the c-command relations are reversed). We don't expect to see LF differences like this, and they don't seem to exist. By positing LF V-to-T raising in English, we then make the claim that French and English *are* the same at LF, which looks like the right thing to say.

Summary

As a complement to the previous chapter, which considered several major theoretical modifications relating to phrase-structure, in this chapter we've looked at a couple of major theoretical modifications relating to movement. In Section 10.1, we introduced the idea that 'traces' are actually complete, but unpronounced, copies of the moved item. Looking particularly at QR, we saw that quantifiers can in fact undergo QR from any position in their chain. This is only possible if each of the links in the chain is a complete copy of the quantifier itself. We also looked at some evidence from Romanian and German that, under certain circumstances, intermediate traces and tails of chains can be spelled out as full copies of the moved element. This is explained naturally if they are simply copies of the original item. In Section 10.2, we looked at the movement approach to control, in which PRO is eliminated from the theory and the θ-criterion is relaxed so that DPs can be assigned more than one θ-role. We discovered that obligatory control contexts, such as subject and object control, can be reanalysed as instances of movement, while, in non-obligatory control contexts, PRO can be replaced by *pro*. Finally, in Section 10.3 we looked at head movement in more detail, focusing on verb raising from V to T.

Open issue: head movement as PF movement?

We ended this chapter by coming back to an issue that had kind of been lurking around in the background since we first discussed movement back in Chapter 6: the question of head movement. We suggested that the verb must raise overtly from V to T in French, but waits until LF to do so in English, on the basis that there seems to be no interpretive difference between a French sentence with verb-raising and its English counterpart without verb-raising.

However, beginning with work in the mid-1990s, Chomsky began to suggest another way of looking at the problem. Rather than explaining the lack of interpretive differences by postulating that the verb actually moved in *both* languages, he suggested that in fact the verb didn't really move in *either*

language. Specifically, he suggested that verb-movement, and head-movement in general, was a post-syntactic *PF* phenomenon. The change in word order happens after the syntax is handed over to the phonology. As far as LF is concerned, French and English are identical because *neither* verb has moved.

With just the introductory discussion about head movement that we had in this chapter, we haven't really encountered some of the more technical reasons why PF head movement might look appealing. There is, for example, a long discussion in the literature about whether head movement is substitution (like wh-movement or raising) or whether it's adjunction (like QR) – and it looks like there are problems either way you do it.

However, it's an important and radical idea within the Minimalist Program, and with the aid of some of the discussion in the Open Issue from Chapter 9 we can see one empirical argument that's been given in the literature in favour of the idea from Boeckx and Stejpanović (2001). As ever, we'll need to adapt the discussion a little bit, but not so much that the basic flavour of the argument is lost.

Boeckx and Stejpanović claim that there's a problem explaining so-called 'pseudo-gapping' sentences, under the assumption that objects move overtly to the specifier of *v*P (the conclusion reached in Chapter 9's Open Issue).

(85) Debbie ate the chocolate and Kazuko did the cookies
(Boeckx and Stejpanović 2001: (1))

The traditional analysis of (85) is as a simple instance of VP-ellipsis, which we've seen before in Chapter 2 and Chapter 8. The difference, though, is that, in a pseudogapping construction, the object is able to escape the VP in the second conjunct, and so it's only the verb *eat* that ends up getting elided.

If objects move in the syntax to the outer specifier of vP, as Lasnik and other collaborators have suggested, then we have a mechanism in place for explaining why the object might be able to escape. However, this solution raises its own problem when we consider a simple coordinate sentence like (86)

(86) Debbie ate chocolate and Kazuko drank milk.
(Boeckx and Stejpanović 2001: (5))

If the objects in both conjuncts must raise in the syntax to the outer specifier of vP, then the word order suggests that the verbs must raise also (a point we made in the Open Issue for Chapter 9), otherwise you would predict that you should be able to get the order in (87), which is ungrammatical.

(87) *Debbie chocolate ate and Kazuko milk drank.

But if the verbs obligatorily leave the VP, then they can't be left behind to be deleted by VP-ellipsis in (85).

What we want is for the verb to be forced to leave the VP, in order to explain (87), but not all the time, in order to make (86) possible. It's hard to imagine

how this is going to work if head movement is a syntactic operation. It can't just be that the movement is optional, full stop, because then you have no explanation for the ungrammaticality of (87).

What Boeckx and Stejpanović propose is that the verb has a feature which says 'I cause the syntactic derivation to crash when we get to PF if I'm still in the VP', and this feature is always present. (This probably seems here a little more stipulative than it is 'in real life', as there is a long literature in the Minimalist Program which we haven't discussed concerning different features that have a similar character.) This must be a PF feature of some kind, perhaps connected with morphology or prosody, as opposed to a syntactic feature, because one way that the feature can be eliminated is through the PF operation of ellipsis – i.e. just not pronouncing the element. That's the option taken in (85).

However, it seems as if the feature can also be satisfied by having a linear order in which the verb is no longer linearly a VP element. So in (86) both the objects and the verbs have left the VP. The objects move in the syntax, but the head movement of the verb is a PF operation. Otherwise you have the strange and unwelcome situation of the syntax 'seeing' a problematic PF feature (and doing something about it), as if the syntax were reacting to a vowel that was specified both [+high] and [+low]. If the change in linear order of heads is *in general* a PF displacement rather than an instance of syntactic movement, then we can explain why this should be the case.

FURTHER READING

Like the previous chapter, the bibliography for this chapter is about continuing on the transition from what we've talked about here to the Minimalist Program. Even more so than in Chapter 9 I think, even the work that we've cited here in this chapter may not be immediately accessible because of the adapting that I've done. As such, really the books that I would recommend you consult are the same ones as the last chapter: the textbooks on the Minimalist Program that I mentioned (Hornstein et al., and Radford and Adger) and the edited collection in Bošković and Lasnik. They'll help you to get familiar with some of the ideas from the Minimalist Program that we haven't discussed in this book, which you can then use to work your way into the primary literature.

EXERCISES

Exercise 1

Draw the S-structure trees for the following sentences using the revised clause structure of this chapter, not forgetting the Copy Theory of Movement:

1. John promised to return the book.
2. Mary is going to the store.

3. Alice persuaded Bill to help.
4. What did John send?
5. John's brother's book arrived.
6. Who did Bill ask Mary to visit?

Exercise 2

Draw the LF trees for the following sentences using the revised clause structure of this chapter:

1. John bought many computers.
2. William has always been claiming that Alice is innocent.
3. Kevin is likely to be in the garden.
4. Some industries may not always be honest.
5. Robert is often here.
6. Every person who bought a book has arrived.

Exercise 3

How might the following piece of data be relevant for a movement analysis of control?

1. I persuaded the ministers all to resign.

Exercise 4

In some Scandinavian languages, like Icelandic, sentential negation is not its own functional projection NegP. Instead, it's a VP-adjoined adverb, like *always*. Given this information, what does the following suggest to you about the issues discussed in Section 10.3?

1. Jólasveinninn borðaði ekki hattinni.
   the Christmas troll ate not the hat
   'The Christmas troll didn't eat the hat.

2. Jólasveinninn borðaði hattinni ekki.
   The Christmas troll ate the hat not
   'The Christmas troll didn't eat the hat.'

3. a. *Jolasveinnin hafa [$_{VP}$ bordað hattinni ekki]
      the.Christmas.troll has eaten the.hat not
   b. Jolasveinnin hafa [$_{VP}$ ekki bordað hattinni]

Notes

2 Phrase Structure and Constituency

1. The alternative would be to say that *do so* can substitute for either VP or V′, but, as I say, the two are essentially equivalent.

4 θ-Theory and Case Theory

1. In fact, there's a joke in *Alice in Wonderland* based on precisely this misunderstanding about expletives, where two characters argue about what the thing was that the Archbishop of Canterbury found when he 'found it advisable that...'.
2. You may be wondering about examples such as *John is likely to be home now* or *John seems to be doing well*. There's something a little different going on in these cases, and we'll discuss them in Chapter 6.
3. In fact, as we'll see in Chapter 8, IP appears to be *generally* defective, in that even finite IP is not a barrier to government.

6 Movement and Chains

1. We will ignore for the moment the transformation that also moves I to C. We will return to this though in Chapter 10.
2. This is a considerable oversimplification. Modern Irish has a lot of different complementizers for different kinds of contexts, but they're not relevant for the present discussion. See Chung and McCloskey (1987), as well as McCloskey (1979) for more details.
3. The capital L does not appear in the orthography, but is merely present to indicate that this is the complementizer *a* which triggers lenition mutation on the following word (as opposed to a different complementizer *a* which does not).
4. We won't concern ourselves here with the mechanism by which the element in the by-phrase gets assigned the θ-role that would normally be assigned to the subject.
5. The attentive reader will have noticed that this is not quite true, strictly speaking. We still have things like (62a) in the text to account for. There is much disagreement in the literature about how to handle these cases. Some recent accounts are Lasnik (1995), Moro (1997) and my own work (Poole 1996). We'll take a closer look at some of the issues involved in Chapter 7.

7 Logical Form

1. Czech (Toman 1981) and Hungarian (Kiss 1994) are other examples.
2. A true semanticist would probably be screaming obscenities at this point, as this is precisely the way that people like Geach (1972) say that you *shouldn't* think about

quantifiers. However, since we're only interested here in the *syntactic* behaviour of quantifiers, this issue, along with much else, will be swept under the rug. For anyone who'd like to take a peek under said rug, there are a number of good textbooks on formal semantics, including Larson and Segal (1995) and Heim and Kratzer (1998).

3. In fact, what's even more natural in my dialect is to use a *plural* pronoun, even when it's explicit that only one person is involved. So (53) is most natural for me as:

   (i) As for the new employee, their duties will include...

   The problem is that even a plural pronoun doesn't *force* a bound-variable interpretation in the way that 'he/she' does. In (ii) *they* can still be interpreted as referring to some unmentioned group of people:

   (ii) Who thinks that they're smart?

   For this reason, I'll continue with 'he/she', even though it's a little less natural.
4. I'm putting aside cases where the verb *be* contracts with *there*. In my dialect, for example, (i) is fine:

   (i) There's three things I need to do.

   However, without contraction, (i) is impossible.
5. It should be mentioned that although superiority is a topic which received a lot of attention in the literature (Lasnik and Saito 1984 is one of the best known), its effects are not well understood at all, and judgements about sentences like (79b) in particular can vary under different discourse situations. Chomsky (1995: ch. 4, n. 69) goes so far as to suggest that the phenomenon doesn't really exist. See particularly Kuno (1987) for some discussion.
6. To my ear, *what on Earth* can't be substituted in this way, so apologies to anyone offended by the profanity. It's all in the name of science.

8 Empty Categories and their Constraints

1. The sentences usually given as illustrations of strong crossover are a little simpler, something like (i), which cannot by synonymous with (ii):

   (i) *Who$_i$ does he$_i$ like t$_i$?
   (ii) Who$_i$ t$_i$ likes himself?

   These simpler sentences don't illustrate the point we need to make about wh-traces, though; hence the need for the more complicated (5) and (7).
2. Dropping the subject is sometimes possible in English. For example, in describing your day in an email message, you might say 'Went to the store. Saw a good movie', etc. This is clearly something different from true null subject languages for several reasons, not the least of which is that the subject can usually only be dropped if it's I or sometimes expletive *it*.
3. But see the Open Issue section of this chapter.
4. See Rizzi's (1986) original article for more extensive argumentation.
5. As I mentioned way back in Chapter 1, there are some rare dialects of American English which don't have the 'that-trace' effect for reasons which are not clear.
6. In Lasnik and Saito's terminology, when a trace is checked by the ECP, it receives the feature $+\gamma$ if it satisfies the ECP and $-\gamma$ if it does not. We won't adopt that terminology here, but I mention it because you will often see references to 'gamma marking' in the

literature on the ECP as another way of saying 'checking to see if a trace satisfies the ECP'.

7. (62) is deliberately vague about exactly how successive cyclicity for QR would be represented (does it move through Spec CP or just from IP-adjoined position to IP-adjoined position, etc.?) because we're about to argue that this analysis isn't right anyway.

9 Towards the Minimalist Program: Clause Structure Revisited

1. A similar example that you may already have thought of is (i):

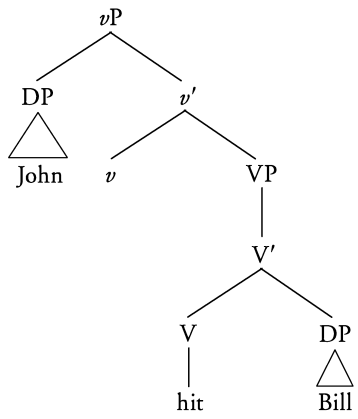
   (i) The children are *all/both/each* believed to have gone to the store.

   Unfortunately, it would take us too far afield to discuss (i) here, but the right analysis is surely related to the fact that English allows objects to shift leftward across passive participles, as in (ii):

   (ii) There were six books$_i$ written t$_i$ about the incident.

2. There's an obvious question that you might ask at this point: wouldn't it be more efficient to just have the Agent θ-role and Accusative Case assigned to the same DP? Then we wouldn't have to worry about this whole question of multiple specifiers. How come this doesn't happen? Unfortunately, the answer is a little complicated and, to a certain extent, relies on assumptions from the Minimalist Program which we haven't really discussed. However, the gist of the answer is this: assume you have something like the D-structure in (i)

(i)



   If *v* assigned accusative Case to *John*, in addition to the Agent θ-role, that would create a problem for I. I needs to have something to assign Nominative Case to. But if *John* moves to the specifier of IP, it gets assigned Case twice, which can't happen. But because *John* is 'closer' to I than *Bill*, *John* prevents *Bill* from moving to the specifier of IP. Only the 'closest' DP can move to the specifier of IP. (This is basically *Relativized Minimality*, from Chapter 8.) The only way that everything can work out is if *John* gets the Agent θ-role, but *not* Accusative Case, moves to the specifier of IP to get Nominative Case, and then *Bill* moves in later as a second specifier of *v*P to get Accusative Case.

3. I'm presenting here a slightly simplified picture for the purposes of introducing the concept. There is some debate in the literature about exactly how many functional categories CP should be broken into, what they encode, and whether any of them are recursive (i.e. whether you can have more than one of a given type).
4. I've slightly modified from Rizzi's original both the glosses and translations of the Italian examples in this section for ease of comprehension.
5. Rizzi in fact suggests that FocusP is both preceded *and* followed by TopicPs, but there is some debate about what the correct analysis is of those elements which follow FocusP, and so for purposes of simplicity I assume simply that TopicP precedes FocusP.

10 Towards the Minimalist Program: Movement Revisited

1. The phenomenon of reconstruction is more usually illustrated with Binding Theory examples within the Minimalist Program (particularly Principle C effects; see e.g. Chomsky 1995; Fox 2000). However, these arguments crucially require the assumption that Binding Theory is satisfied at LF, not S-Structure. This is the standard assumption within the Minimalist Program, but which would take us too far afield to motivate here.
2. Just to emphasize, a non-obligatory control context doesn't mean that you *have* to have PRO_{arb}, just that you *can*.
3. You could imagine a context in which you know that John had an important job interview, and you saw Mary after she spoke with John, and her appearance suggested that John was happy. But no context like that makes the sloppy identity reading of (49) possible.
4. This principle of phrase-structure compatibility, which says that only heads can move into head positions and that only XPs can move into XP positions (like specifiers), was first explicitly discussed by Emonds (1976) under the name of the Principle of Structure Preservation.
5. See Pollock (1989, n. 3) for some discussion. Also potentially relevant is the fact that some non-standard dialects of French which have only one negative marker use only *pas* and not *ne*.

References

Abney, S. (1987) 'The English Noun Phrase in its Sentential Aspect', unpublished PhD dissertation, Massachusetts Institute of Technology.

Adger, D. (2003) *Core Syntax* (Oxford: Oxford University Press).

Aoun, J., N. Hornstein and D. Sportiche (1981) 'Some Aspects of Wide Scope Quantification', *Journal of Linguistic Research* 1, pp. 69–95.

Baker, C.L. (1970) 'Notes on the Description of English Questions: The Role of an Abstract Question Morpheme', *Foundations of Language* 6, pp. 197–219.

Barss, A. and H. Lasnik (1986) 'A Note on Anaphora and Double Objects', *Linguistic Inquiry* 17, pp. 347–54.

Bobaljik, J. (2003) 'Floating Quantifiers: Handle with Care', in L. Cheng and R. Sybesma, eds, *The Second Glot International State-of-the-Article Book* (Berlin: Mouton de Gruyter), pp. 107–48.

Boeckx, C. and S. Stjepanović (2001) 'Heading Toward PF', *Linguistic Inquiry* 32, pp. 345–55.

Borsley, R. (1994) 'In Defense of Coordinate Structures', *Linguistic Analysis* 24, pp. 218–46.

Bošković, Ž. (2002) 'On Multiple Wh-Fronting', *Linguistic Inquiry* 33, pp. 351–83.

Bošković, Ž. and H. Lasnik (2007) *Minimalist Syntax: The Essential Readings* (Oxford: Blackwell).

Bošković, Ž. and J. Nunes (2007) 'The Copy Theory of Movement: A View From PF', in N. Corver and J. Nunes, eds, *The Copy Theory of Movement* (Amsterdam: John Benjamins).

Burzio, L. (1986) *Italian Syntax* (Dordrecht: Foris).

Butt, J. and C. Benjamin (2000) *A Reference Grammar of Modern Spanish* (London: Hodder).

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

Chomsky, N. (1965) *Aspects of the Theory of Syntax* (Cambridge, MA: MIT Press).

Chomsky, N. (1970) 'Remarks on Nominalization', in R.A. Jacobs and P.S. Rosenbaum, eds, *Readings in English Transformational Grammar* (Waltham, MA: Ginn), pp. 184–221.

Chomsky, N. (1973) 'Conditions on Transformations', in S. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle* (New York: Holt, Reinhart and Winston), pp. 232–86.

Chomsky, N. (1981) *Lectures on Government and Binding* (Dordrecht: Foris).

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding* (Cambridge, MA: MIT Press).

Chomsky, N. (1986) *Barriers* (Cambridge, MA: MIT Press).

Chomsky, N. (1988) *Language and Problems of Knowledge: The Managua Lectures* (Cambridge, MA: MIT Press).

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

Chomsky, N. (2001) *On Nature and Language*, A. Belletti & L. Rizzi, eds (Cambridge: Cambridge University Press).
Chung, S. and J. McCloskey (1987) 'Government, Barriers, and Small Clauses in Modern Irish', *Linguistic Inquiry* 18, pp. 173–237.
Demonte, V. and O. Fernández-Soriano (2009) 'Force and Finiteness in the Spanish Complementizer System', *Probus* 21, pp. 23–49
Emonds, J. (1976) *A Transformational Approach to English Syntax* (New York: Academic Press).
Epstein, S.D. (1984) 'Quantifier PRO and the LF Representation of PRO_{arb}', *Linguistic Inquiry* 15, pp. 499–505.
Epstein, S.D. (1990) 'Differentiation and Reduction in Syntactic Theory: A Case Study', *Natural Language and Linguistic Theory* 8, pp. 301–13.
Epstein, S.D. and N. Hornstein (1999) *Working Minimalism* (Cambridge, MA: MIT Press).
Fiengo, R. and H. Lasnik (1972) 'On Non-Recoverable Deletion in Syntax', *Linguistic Inquiry* 3, p. 528.
Fox, D. (1999) 'Reconstruction, Binding Theory and the Interpretation of Chains', *Linguistic Inquiry* 30, pp. 157–196
Fox, D. (2000) *Economy and Semantic Interpretation*. (Cambridge, MA: MIT Press).
Freeborn, D. (1992) *From Old English to Standard English* (Basingstoke: Macmillan).
Fukui, N. (1986) *A Theory of Category Projection and its Application*, unpublished PhD dissertation, Massachusetts Institute of Technology.
Fukui, N. & M. Speas (1986) 'Specifiers and Projection', in N. Fukui, T. R. Rappaport & E. Sagey, eds *MIT Working Papers in Linguistics* 8 (Cambridge, MA: Department of Linguistics & Philosophy, Massachusetts Institute of Technology), pp. 128–72.
Geach, P.T. (1972) *Logic Matters* (Berkeley and Los Angeles, CA: University of California Press).
Gregg, K. (1989) 'Second Language Acquisition Theory: The Case for a Generative Perspective', in S.M. Gass and J. Schacter, eds, *Linguistic Perspectives on Second Language Acquisition* (Cambridge: Cambridge University Press), pp. 15–37.
Grimshaw, J. (1990) *Argument Structure* (Cambridge, MA: MIT Press).
Haegeman, L. (1994) *Introduction to Government and Binding Theory* (Oxford: Blackwell).
Hale, K. and S.J. Keyser (1993) 'On Argument Structure and the Lexical Expression of Syntactic Relations', in K. Hale and S.J. Keyser, eds, *The View from Building 20: A Festschrift for Sylvain Bromberger* (Cambridge, MA: MIT Press), pp. 53–108.
Hale, K. and S.J. Keyser (2002) *Prolegomena to a Theory of Argument Structure* (Cambridge, MA: MIT Press).
Harada, K. (1972) 'Constraints on WH-Q Binding', *Studies in Descriptive and Applied Linguistics* 5, pp. 180–206.
Harbert, W. (1995) 'Binding Theory, Control, and pro', in G. Webelhuth, ed., pp. 177–240.
Heim, I. and A. Kratzer (1998) *Semantics in Generative Grammar* (Oxford: Basil Blackwell).
Hornstein, N. (1995) *Logical Form* (Cambridge, MA: MIT Press).
Hornstein, N. (1999) 'Movement And Control', *Linguistic Inquiry* 30, pp. 99–127.
Hornstein, N., J. Nunes & K. Grohmann (2005) *Understanding Minimalism* (Cambridge: Cambridge University Press).
Huang, J. (1982) 'Logical Relations in Chinese and the Theory of Grammar', unpublished PhD dissertation, Massachusetts Institute of Technology.
Huang, J. (1983) 'A Note on the Binding Theory', *Linguistic Inquiry* 14, pp. 554–61.

Huang, J. (1995) 'Logical Form', in G. Webelhuth, ed., pp. 125–77.
Jackendoff, R.S. (1977) *X'-Syntax: A Study of Phrase Structure* (Cambridge, MA: MIT Press).
Jackendoff, R.S. (1984) *Semantics and Cognition* (Cambridge, MA: MIT Press).
Jackendoff, R.S. (1990) *Semantic Structures* (Cambridge, MA: MIT Press).
Jackendoff, R.S., J. Maling and A. Zaenen (1993) 'Home is Subject to Principle A', *Linguistic Inquiry* 24, pp. 173–77.
Jaeggli, O. (1982) *Topics in Romance Syntax* (Dordrecht: Foris).
Kiss, K. (1994) 'Sentence Structure and Word Order', in F. Kiefer and K. Kiss, eds, *Syntax and Semantics 27: The Syntactic Structure of Hungarian* (San Diego, CA: Academic Press).
Kitagawa, Y. (1986) 'Subjects in Japanese and English', unpublished PhD dissertation, University of Massachusetts, Amherst.
Kuno, S. (1987) *Functional Syntax,* (Chicago, IL: University of Chicago Press).
Larson, R.K. (1988) 'On the Double Object Construction', *Linguistic Inquiry* 19, pp. 335–91.
Larson, R.K. and G. Segal (1995) *Knowledge of Meaning* (Cambridge, MA: MIT Press).
Lasnik, H. (1989) *Essays on Anaphora* (Dordrecht: Kluwer Academic Publishers).
Lasnik, H. (1995) 'Case and Expletives Revisited: On Greed and Other Human Failings', *Linguistic Inquiry* 26, pp. 615–33.
Lasnik, H. (1999) *Minimalist Analysis* (Oxford: Blackwell).
Lasnik, H. and M. Saito (1984) 'On the Nature of Proper Government', *Linguistic Inquiry* 15, pp. 235–89.
Lasnik, H. and M. Saito (1991) 'On the Subject of Infinitives', in L.M. Dobrin, L. Nichols and R.M. Rodriguez, eds, *CLS Part I, The General Session* (Chicago, IL: Chicago Linguistics Society) [reprinted in Lasnik (1999)].
Lasnik, H. and M. Saito (1992) *Move* α (Cambridge, MA: MIT Press).
Lasnik, H. and J. Uriagereka (1988) *A Course in GB Syntax* (Cambridge, MA: MIT Press).
Martin, R. (1999) 'Case, the Extended Projection Principle, and Minimalism', in Epstein and Hornstein (1999), pp. 1–26.
May, R. (1977) 'The Grammar of Quantification', unpublished PhD dissertation, Massachusetts Institute of Technology [reprinted (1990) New York: Garland].
May, R. (1985) *Logical Form* (Cambridge, MA: MIT Press).
McCloskey, J. (1979) *Transformational Syntax and Model Theoretic Semantics* (Dordrecht: Reidel).
Moro, A. (1997) *The Raising of Predicates* (Cambridge: Cambridge University Press).
Oehrle, R. (1976) 'The Grammatical Status of the English Dative Alternation', unpublished PhD dissertation, Massachusetts Institute of Technology.
O'Neil, J. (1995) 'Out of Control', in J. Beckman, ed., *Proceedings of NELS 25* (Amherst, MA: GLSA, University of Massachusetts).
O'Neil, J. (1997) 'Means of Control: Deriving the Properties of PRO in the Minimalist Program', unpublished PhD dissertation, Harvard University.
Pesetsky, D. (1987) 'Wh-in-situ: Movement and Unselective Binding', in E. Reuland and A. ter Meulen, eds, *The Representation of (In)definiteness* (Cambridge, MA: MIT Press), pp. 98–129.
Pollock, J.-Y. (1989) 'Verb Movement, UG, and the Structure of IP', *Linguistic Inquiry* 20, pp. 365–424.
Poole, G. (1996) 'Transformations Across Components', unpublished PhD dissertation, Harvard University.

Reinhart, T. (1976) 'The Syntactic Domain of Anaphora', unpublished PhD dissertation, Massachusetts Institute of Technology.
Reinhart, T. (1983) *Anaphora and Semantic Interpretation* (London: Croom Helm).
Rizzi, L. (1982) *Issues in Italian Syntax* (Dordrecht: Foris).
Rizzi, L. (1986) 'Null Objects in Italian and the Theory of pro', *Linguistic Inquiry* 17, pp. 508–51.
Rizzi, L. (1990) *Relativized Minimality* (Cambridge, MA: MIT Press).
Rizzi, L. (1997) 'The Fine Structure of the Left Periphery', in L. Haegeman, ed., *Elements of Grammar* (Dordrecht: Kluwer), pp. 281–387.
Ross, J.R. (1967) 'Constraints on Variables in Syntax', unpublished PhD dissertation, Massachusetts Institute of Technology.
Russell, B. (1905) 'On Denoting', *Mind* 14, pp. 479–93.
Saito, M. (1989) 'Scrambling as Semantically Vacuous A′-Movement', in M. Baltin and A. Kroch, eds, *Alternative Conceptions of Phrase Structure* (Chicago, IL: University of Chicago Press).
Smith, N. and I. Tsimpli (1995) *The Mind of a Savant* (Oxford: Blackwell).
Toman, J. (1981) 'Aspects of Multiple Wh-Movement', in R. May and J. Koster, eds, *Levels of Syntactic Representation* (Dordrecht: Foris), pp. 293–302.
Torrego, E. (1984) 'On Inversion in Spanish and Some of its Effects', *Linguistic Inquiry* 15, pp. 103–30.
Webelhuth, G., ed., (1995a) *Government and Binding Theory and the Minimalist Program* (Oxford: Blackwell).
Webelhuth, G. (1995b) 'X′-Theory and Case Theory', in G. Webelhuth, ed., pp. 15–96.
Zagona, K. (1982) 'Government and Proper Government of Verbal Projections', unpublished PhD dissertation, University of Washington, Seattle.

Index

A-
  A-bound, 231, 233
  A-free, 231, 233
  A-movement, 172, 190, 216, 217, 227, 228, 230, 232, 233, 235, 240, 242, 243, 271, 296, 299
  A-position, 153, 172, 216, 228, 229, 230, 231, 233, 257, 259
A′-
  A′-movement, 172, 216, 227, 228, 230, 231, 232, 233, 235, 239, 242, 243, 245, 299
  A′-position, 113, 172, 216, 228, 232, 244, 257, 258, 259
Abney, S., 73, 86
acceptability, 6, 7, 11
adjunct, 63–5, 151, 200, 250–6
Adger, D., 297, 330
adverbs, 45
  sentential, 321
  VP-oriented, 320–2, 324–6
analogy, 9–11, 239, 241
anaphor, 120–37, 144, 150, 155, 183, 227, 230, 273, 295, 300
anaphora, 147
antecedent, 53, 120ff.
  'split' 311, 312, 318
antecedent contained deletion, 281–3, 296
Aoun, J., Hornstein, N. and Sportiche, D, 219
Arabic, Palestinian, 80
argument structure, 88, 89–90, 91, 92, 115, 118, 179, 188, 218, 263
auxiliary verb, 15, 85, 270

Baby Ray, 207
Baker, C. L., 198
barrier, 102–4, 106, 108, 112, 116–17, 129, 230, 238, 242, 247–8, 251, 253, 256, 258, 280, 332
  IP not a barrier, 106–7, 240, 246, 247
Barss, A. and Lasnik, H., 54, 273
base-generated, 151, 154, 155, 159, 169, 173, 174, 178, 179, 183, 266, 267, 270, 271, 278, 292, 296, 309, 322
Beatles, The, 98, 186, 201, 202
behaviour, linguistic, 3–5
Berber, 17
binding, 126, 142, 143, 150, 233, 273, 296
Binding Theory, 77, 119–49, 150, 154, 155, 176, 183, 186, 221, 226, 227–39, 242, 244, 245, 266, 273–4, 294–6, 299–300, 304, 310, 335
  see also Principle A, B, C
'black box', 3, 5, 6
Bobaljik, J., 297
Boeckx, C. and Stepanović, S., 329–30
Borsley, R., 84, 86
Bošković, Ž., 306
Bošković, Ž. and Lasnik, H., 265, 297, 306, 330
Bošković, Ž. and Nunes, J., 305–6
bounding node, *see* nodes, bounding
brain, 2–3, 5, 11, 18
Burzio, L., 174, 281
Burzio's Generalization, 174, 281
Butt, J. and Benjamin, C., 72

Carrier, J., 7
'cartographic' approach to CP, 287, 292
Case Filter, 108–9, 110–12, 113–14, 117, 178, 180, 183, 196, 239
Case Theory, 89, 98–118
Case, abstract
  accusative Case assignment, 99–104, 106–8, 128, 216, 334
  genitive Case assignment, 100, 101, 108
  nominative Case assignment, 99, 101, 104, 106, 107, 109, 136, 280
  oblique Case assignment, 99, 100, 101, 102, 106
Case, morphological
  accusative case, 89, 99
  genitive case, 98
  nominative case, 89, 98
  oblique case, 98
causative/inchoative alternation, 277–8
c-command, 33–4, 53–4, 58, 59, 101–2, 124–7, 129, 136, 139, 142–4, 150, 151, 154, 176, 183, 196–7, 203–4, 207–8, 210, 212, 214, 215, 225, 229, 231–2, 245, 257, 259, 267, 273–4, 277, 292–6, 300, 301, 303, 304, 310–13, 318, 328
chain, 177–80, 196, 216, 254, 257, 300, 305–7, 317, 328
Chinese, 189, 190–3, 196, 199, 200, 202, 222–3, 225, 305
Chomsky, N., 1, 2, 3, 5, 8, 11, 12, 20, 54, 55, 62, 86, 93, 106, 108, 110, 113, 114, 115, 119, 134, 144, 147, 178, 183, 217, 219, 228, 239, 258, 263, 265, 276, 277, 280, 300, 328, 333, 335
Chomsky, N. and Lasnik. H., 280
Christopher (autistic patient), 17–18, 20
Chung, S. and McCloskey, J., 168, 332
coindexation (hypothetical), 134–6, 137, 144, 148
'comma' notation, 82
competence/performance distinction, 3–4, 6, 10

complement (of a head), 63, 65
complementary distribution between pronouns and anaphors, 145–7
complementizer (C), 10, 66, 70–1, 85, 96, 106, 108, 128, 159, 166, 168, 170, 172, 240, 243–9, 250, 254, 255, 288, 290, 292, 332
'conceptual' argument, 18, 77, 202, 206, 300
constituent, 27–31, 33, 34–5, 37, 38, 40, 42, 43, 44, 45, 46, 48, 50, 52, 57, 58, 61, 67, 82, 85, 273, 275
coordination test, 27, 29–31, 38, 43, 45, 48, 52, 67, 79
control
  obligatory, 310–13, 316, 317
  non-obligatory, 318–19
  movement approach to, 307–18
  see also PRO, subject-control verb, object-control verb
'copy' theory of movement, 299, 300, 305, 306
crossing, 161–2, 165, 180, 182, 193, 200
Czech, 80, 332

D(eep)-Structure, 152–6, 160, 161, 164, 173, 174, 176–80, 183, 185–9, 191, 192–6, 199, 202, 212, 215–18, 226, 233, 240, 266, 267, 269, 270, 271, 276, 277, 281, 284, 293, 299, 300, 302, 303, 306, 319, 322, 327, 334
d(iscourse)-linking (of wh-words), 219–22
Demonte, V. and Fernández-Soriano, O., 71
determiner, 22–4, 38, 47, 55, 65, 72, 73, 75–9, 98, 101, 108, 139, 270
do so, 40–2, 45, 49–52, 64, 332
domination, 32, 33, 34, 62, 66, 101, 103, 125, 161, 182, 196, 207, 210, 273
  immediately, 32, 62, 72, 74
double-object verb, 52–5, 271–5
DP-hypothesis, 72, 75

ease of use, role in grammaticality of, 9
echo question, 153, 198
ellipsis, 68, 69, 80, 282, 311, 313, 326, 329, 330
  see also antecedent contained deletion
Emonds, J., 335
'empirical' argument, 18, 77, 80, 81, 83, 139, 366, 269, 280, 281, 300, 329
Empty Category Principle (ECP), 240, 246–59, 263, 333, 334
English, Old, 98
Epstein, S. D., 113–14, 260–3
Exceptional Case Marking (ECM), 104–10, 114, 117, 127–9, 215, 238, 248, 252, 253, 259, 271, 283, 284, 292–6
expletives, 92–4, 104, 111, 112, 173, 176, 254, 332, 333
  Case transmission by, 111
  expletive replacement, 216–18
Extended Projection Principle (EPP), 93–5, 106, 314–15

falsification, 3
Fox, D., 300, 302, 303, 335
free (not bound), 138, 143–5, 147, 228, 230, 231, 233–6, 238, 245
Freeborn, D., 98
French, 232, 233, 239, 320, 322–9, 335
Fukui, N., 86
Fukui, N. and Speas, M., 86
functional categories, 61, 66, 74, 75, 78–80, 84, 86, 100, 245, 246, 265, 266, 280, 281, 287, 289, 291, 292, 335

γ (gamma)-marking, 333
'garden path' sentences, 9
Geach, P., 332
German, 13, 78, 89, 306–7, 328
governing category, 129–34, 136–42, 144–8, 150, 155, 186, 227–36, 238–40, 242, 263–4, 273
government, 106, 129, 151, 152, 237, 238, 240, 242, 243, 246, 248, 256–8, 267, 280, 332
  antecedent, 247–59
  lexical proper, 246–8, 250–3, 255, 256, 258, 259
  minimality requirement on, 247
  proper, 243, 245, 249, 255–7, 259
  requirement on traces, 243–6
grammar (of a language), 3, 7, 11, 19, 20, 89, 98, 110, 112, 113, 188, 189, 218, 239, 249, 287
Gregg, K., 4
Grimshaw, J., 115
Guided by Voices, 175, 207

Haegeman, L., 20, 55, 106, 115, 147, 183, 222
Hale, K. and Keyser J., 277
Harada, K., 196
Harbert, W., 263
Harris, Z., 2
head (of a phrase), 22, 23, 25, 33, 35, 36, 40, 45, 47, 52, 61–7, 69, 70, 72, 74, 76, 77, 80–6, 101–3, 112, 129, 139, 153, 155 , 159, 177, 179, 203, 246–8, 256–8, 268, 276, 280, 283, 286–90, 292, 305–7, 320, 322, 324, 326, 328–30, 335
head-complement relationship, 280
head-initial vs head-final languages, 81–3, 190
head-movement, 319, 326, 328–30
Heim, I. and J. Kratzer, 333
Hornstein, N., 222, 281, 296, 308, 310, 312, 317, 318
Hornstein, N. et al., 265, 297, 330
Huang, J., 145, 146, 190, 196, 200, 222
Hungarian, 204, 332

Icelandic, Modern, 118, 331
idioms, 273
immediate domination, *see* domination, immediate
'inclusive' language, 211–13
Infl (I), 66–9, 75, 81, 84, 100–2, 104, 106, 107, 109, 110, 112, 116, 153, 155, 177, 233, 237, 242, 244–7, 250, 262, 280, 281, 288, 326–7, 332, 334
intuitions, native speaker, 6, 11, 22, 52, 113, 115, 239, 249–50, 261, 295
inverse linking, 207–11
Irish, Modern, 166, 168, 180, 332
Italian, 181–2, 228, 232, 234–5, 258, 259, 288–91, 335

iterate, 38, 39, 45
I-within-I Filter, 134–7, 147, 230

Jackendoff, R. S. et al., 149
Jackendoff, R. S., 55, 115
Jaeggli, O., 219
Japanese, 18, 22, 66, 72, 81–3, 189–92, 196, 199, 200, 202, 210, 218–22

katakana, 22
Kayne, R., 219
Kiss, K., 332
Kitagawa, Y., 297
knowledge, unconscious, 12
Kuno, 333

language acquisition device, 12, 20, 239
language acquisition, logical problem of, 11, 12, 239
  see also Plato's Problem
language faculty, 18, 52, 201, 218
Larson, R., 88, 273, 276
Larson, R. and Segal, K., 333
Lasnik, H., 147, 265, 280, 295, 329, 332
Lasnik, H. and Saito, M., 199, 220, 254, 259, 263, 282, 284–6, 333
Lasnik, H. and Uriagereka, J., 115, 147, 183, 222, 263
Latin, 89, 98, 99
lexical categories, 61, 66, 86, 245–8, 253, 259, 265
lexical item, 48, 89, 93, 300
lexical proper government, *see* government
lexicon, 89, 115, 179, 188, 189, 218, 317, 327
'likely' vs 'probable', 9–11, 242
Logical Form (LF), 188–226, 254–6, 259–61, 264, 282–5, 295–6, 302–4, 315, 319, 326–9, 335

'main' verb, 9, 16, 69, 325
Martin, R., 111
May, R., 207, 208, 222, 282
McCloskey, J., 332
m-command, 102–4, 112, 116, 129, 246, 248, 251, 267–9, 280
meaning, 9, 13, 64, 88, 191, 192, 201, 230, 260, 279, 282, 284, 306, 312
'minimal IP' requirement, 120–4, 126, 127, 129, 130, 136, 262
minimality, 247–54, 256, 259
  relativized, 256–8
  rigid, 256–8
  see also government, minimality requirement on
Moro, A., 332
movement
  constraints on, 168, 191, 200, 206, 216, 221, 240, 243 (*see also* Subjacency, weak crossover, strong crossover, ECP)
movement, covert, 193 (*see also* Logical Form)
movement, overt, 189, 193, 196, 200, 202, 208, 210, 217, 219, 222
Movement test, 27, 30–1, 52, 57

Neg(P), 267–8, 320–6
negation, 268–70, 274, 295, 320–6, 328
negative polarity item (NPI), 53, 274, 295–6
nodes
  bounding, 161–2, 164–5, 180, 181–2, 193–4, 196, 220, 240, 242, 244, 250, 306
  daughter, 32, 85, 86
  mother, 32
  sister, 32–6, 40, 47–51, 54, 63–5, 79, 97, 105, 132, 133, 139, 150, 151, 153, 154, 169, 172, 178, 182, 190–4, 228, 268, 270, 277, 280, 286, 287, 293, 321
non-null-subject languages, 13, 232, 233, 259
noun phrase (NP), 22–9, 31–43, 45–8, 52–5, 61–9, 72–9, 82, 102, 103, 134, 136, 147, 182, 232
null-subject languages, 13, 166, 232, 233, 239, 258, 262–3, 333

object-control verb, 97, 304, 310, 317
Oehrle, R., 54
O'Neil, J., 307
'one' substitution, 35–9, 47, 48, 51, 52
operator, 172, 190, 202, 211, 274, 291, 303
operator-variable structures, 202, 213

parameters (of Universal Grammar), 12–13
'Parrot' Theory, 8
passive, 172–7, 179, 180, 202, 217, 228, 230, 231, 271, 299, 304, 334
percolation (of syntactic features), 80–1, 268
Pesetsky, D., 219–21
phrase-structure rule, 21, 23, 24, 26, 31, 39, 40, 44, 46, 61, 52, 66, 67, 82, 152, 154, 173, 176
phrase-structure tree, 21, 23, 24, 26, 31, 32, 24, 46, 68, 79, 89, 161, 319
Plato's Problem, 180, 239
  see also language acquisition, logical problem of
Polish, 199, 200, 204, 305
Pollock, J.-Y, 335
Poole, G., 332
'poss-ing' construction (gerundive), 73–6, 108, 287
postpositions, 82
predicate, 15, 16, 88–91, 93, 94, 110, 112, 113, 153, 175, 188, 235, 240, 277, 282, 307, 316, 317
Principle A, 119, 120–37, 138, 144, 145, 176, 227–31, 233, 236, 238, 240–2, 249, 273, 294, 295
Principle B, 137–41, 144, 145, 147, 227–32, 236, 242
Principle C, 141–4, 213, 227, 229, 231, 232, 244, 335
principles (of Universal Grammar), 12, 13, 111, 113, 115, 119, 176, 200, 217, 221, 238, 248, 249, 250
PRO (big pro), 94, 96, 97, 108–12, 227, 228, 232, 234–40, 242, 243, 258, 260, 261, 263, 304, 307–13, 317, 318, 328
  'arbitrary' PRO, 97, 236, 237, 260, 310
  'controlled' PRO, 97, 235, 237, 261
  PRO Theorem, 96, 236–9, 240, 243, 263
pro (little pro), 227, 234, 235, 239, 258, 259–63, 318, 328
proform, 29, 35, 40, 45, 67

pronouns, 7, 13, 27, 28, 76, 77, 98, 99, 104, 137–9, 141–7, 211–13, 231–6, 274, 275, 288, 310, 312, 318, 319, 333
'bound-variable' interpretation for, 211–13, 232, 333
see also pro
Proper Binding Condition, 196–7, 207, 210, 214, 2128, 270, 304
proper government, *see* government

Quantifier Raising (QR), 189, 201–11, 213–15, 225, 260–2, 282–4, 287, 300–3, 305, 328–9, 334
quantifier lowering, 303–5
quantifiers, 201–15, 260, 262, 266, 275, 284, 286, 301–3, 305, 333
existential, 300–2, 305
'stranded', 269–71
universal, 260–2, 300–2, 305

R(eferring)-Expression, 119, 141–4, 207, 227–9, 231–3, 235, 239, 244, 245, 258
definite description, 143–4
epithet, 143–4
proper name, 75, 77–9, 119, 141, 143, 144
Raising, 172, 175–80, 185, 187, 190, 202, 214–15, 217, 222, 228, 231, 240, 242, 260, 284, 299, 301, 308, 313, 319–20, 329
see also raising predicate, raising verb, 'superraising' violations
raising predicate, 175–6, 185, 240
raising verb, 175, 187, 214
reciprocals, 53, 119–21, 144, 227
reconstruction, 300, 335
reflexive(s), 53, 119–21, 144, 176, 274, 310
Reinhart, T., 33, 213
Rizzi, L., 181–2, 234, 256–8, 266, 287–92
Romanian, 305–6, 328
Ross, J. R., 183
Russell, B., 201
Russian, 72, 99, 102

S(urface)-Structure, 152–6, 161, 176–7, 179–80, 185, 188–9, 191–4, 196–204, 212, 216, 218, 220, 222, 254–6, 258–9, 262, 264, 274, 282, 284, 286, 296–7, 300, 302–3, 319, 322, 326–7, 330, 335
Saito, M., 196, 199, 220, 254, 259, 263, 292, 294–6, 333
science, 1–4, 21, 333
scope, 202–5, 208, 214, 221, 225, 260, 262, 300, 302–4
ambiguity, 204, 206, 214, 224–5
marker, 197
narrow, 203, 208
wide, 203–5, 207–11, 214–15, 224, 300–305, 328
Smith, N. and Tsimpli, I., 20
Spanish, 13, 18, 20, 71–2, 78, 87, 166, 180, 228, 232–3, 239, 259, 271
specifier (of a phrase), 59, 62–3, 65–6, 70–2, 76–7, 80, 84–5, 95, 100–1, 108, 130, 132, 139–40, 151, 153, 163–4, 169, 172, 191, 216, 237, 266–71, 275, 277–81, 283–4, 286, 291–4, 296, 304–7, 309, 314–15, 317, 319, 329, 334–5
multiple, 284–6, 334
specifier-head, 80–1, 84–5, 177, 216
agreement, 80–1, 84–5
'split' CP, 287–8, 291–2
strong crossover, 231, 333
Structure Preservation Principle, 335
structure-dependency, 17
stylistic factors, 4, 6, 7
Subjacency Principle, 161–3, 180, 249
subject (Binding Theory), 130–6, 139–41, 144–7, 230, 236
'accessible' subject, 130, 134–7, 144–7, 230, 236, 242
subject-control verb, 97, 308
subject inversion, 'free' (Spanish), 166
subject/object asymmetry, 245
subject-verb agreement, 68, 80, 84, 155–7, 177–9, 183, 216, 226, 286, 326
Substitution test, 27, 29, 37, 43, 47–8, 51–2
Successive Cyclic Hypothesis, 160–3, 165
successive-cyclicity, 160–2, 166, 167, 180, 194
superiority effects, 219, 222
'superraising' violations, 230, 313
Swahili, 16, 18
syntactic category, 22–3, 27, 29, 35, 38, 62, 73, 84, 89

TP, 266, 286, 287, 292–5, 301–5, 309, 314–17, 319, 320, 322
Tagalog, 7
'that-trace' effects, 243–5, 248, 250, 25, 253, 254, 255, 258, 259, 333
θ (theta)-Criterion, 92, 94, 95, 109–14, 173, 174, 176, 178, 180, 196, 234, 235, 281, 307, 310, 316, 317, 328
θ-position, 173–5, 178, 180, 233
θ-role, 88, 89, 90–4, 95, 109–14, 152, 173–8, 196, 216, 234, 235, 252, 253, 259, 266–9, 280, 281, 300, 305, 307, 310, 317, 328, 332, 334
Agent θ-role, 91, 94, 95, 113, 130, 233, 267, 277, 278, 280, 281, 286, 292, 307, 314, 315, 317, 334
Experiencer θ-role, 91, 95, 114
Goal θ-role, 92, 94
Instrument θ-role, 252
Location θ-role, 92, 252
Patient θ-role, 91, 95, 151, 152, 252, 281, 315, 317
Source θ-role, 92
Theme θ-role, 91, 94, 278
θ-Theory, 89, 94, 95, 97, 112, 115
Toman, J., 332
Topicalization, 30, 31, 186, 203
Torrego, E., 166
trace(s), 152, 160, 177–80, 185, 188, 196–7, 202, 208, 210–12, 227–33, 235, 239–40, 242–51, 253–9, 263–4, 283, 299–300, 303–7, 313, 328, 333–4
traces, intermediate, 160, 244, 246, 247, 248, 255, 306, 307, 328

transformations, 18, 151–2, 154–5, 172, 176, 179–80, 183, 188–9, 193, 207, 210, 217–18
Truth, The, xii
Turkish, 66, 81, 82, 83, 239, 259

ungrammaticality, degrees of, 113–15, 249
Universal Grammar (UG), 11–14, 16–18, 20, 72, 78–9, 82–3, 239

variable, 62, 202, 261, 262, 303
  bound-variable interpretation, *see* pronoun, 'bound-variable' interpretation for
verb raising, 320, 322–3, 325–8
verbal particles, 73
Vergnaud, J.-R., 108
Visibility Condition, 110–12, 113–15
*v*P, 266, 278–81, 283, 284, 286, 292–6, 314, 315, 317, 320–2, 329, 334
VP-internal subject, 266, 269, 271, 275, 281, 284, 286, 297
VP-shell, 66, 162, 271, 272, 275–6, 284, 292, 314
V-to-T raising, *see* verb raising

weak crossover, 211–13
Webelhuth, G., 55, 62, 115, 147, 222, 163
wh-cleft construction, 91
Wh-movement, 152–4, 156–7, 159–62, 165–73, 178–80, 183, 186–7, 189–94, 196, 198–200, 202, 206–7, 210–11, 213, 218–19, 222, 228, 231, 244, 257, 299, 319, 329

X′-Theory, 55, 61, 62, 66, 70, 74, 79, 81, 83, 85, 86, 89, 139, 179, 188, 189, 203, 218, 265
X^{0}-movement *see* head-movement

Yoda, 31

Zagona, K., 297