Locality | Edited by Enoch Oladé Aboh, Maria Teresa Guasti, and Ian Roberts

OXFORD STUDIES IN COMPARATIVE SYNTAX

Locality

OXFORD STUDIES IN COMPARATIVE SYNTAX
Richard Kayne, General Editor

The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2
Edited by Luigi Rizzi

The Syntax of Anaphora
Ken Safir

Principles and Parameters in a VSO Language: A Case Study in Welsh
Ian G. Roberts

Structures and Beyond: The Cartography of Syntactic Structures, Volume 3
Edited by Adriana Belletti

Movement and Silence
Richard S. Kayne

Restructuring and Functional Heads: The Cartography of Syntactic Structures, Volume 4
Guglielmo Cinque

Scrambling, Remnant Movement and Restructuring in West Germanic
Roland Hinterhölzl

The Syntax of Ellipsis: Evidence from Dutch Dialects
Jeroen van Craenenbroeck

Mapping the Left Periphery: The Cartography of Syntactic Structures, Volume 5
Edited by Paola Benincà and Nicola Munaro

Mapping Spatial PPs: The Cartography of Syntactic Structures, Volume 6
Edited by Guglielmo Cinque and Luigi Rizzi

The Grammar of Q: Q-Particles, Wh-Movement, and Pied-Piping
Seth Cable

Comparisons and Contrasts
Richard S. Kayne

Discourse-Related Features and Functional Projections
Silvio Cruschina

Functional Heads: The Cartography of Syntactic Structures, Volume 7
Edited by Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro, Cecilia Poletto

Adverbial Clauses, Main Clause Phenomena and Composition of the Left Periphery: The Cartography of Syntactic Structures, Volume 8
Liliane Haegeman

Variation in Datives
Edited by Beatriz Fernández and Ricardo Etxepare

Locality
Edited by Enoch Oladé Aboh, Maria Teresa Guasti, and Ian Roberts

Locality

Edited by Enoch Oladé Aboh, Maria Teresa Guasti, and Ian Roberts

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trademark of Oxford University Press
in the UK and certain other countries.

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

Library of Congress Cataloging-in-Publication Data

Locality / Edited by Enoch Oladé Aboh, Maria Teresa Guasti, and Ian Roberts
pages cm. — (Oxford Studies in Comparative Syntax)
Includes bibliographical references and index.
ISBN 978-0-19-994528-3 (pbk. : alk. paper)—ISBN 978-0-19-994526-9 (hardcover : alk. paper)
1. Grammar, Comparative and general—Syntax. 2. Cognitive grammar.
3. Generative grammar. I. Roberts, Ian G., editor of compilation. II. Aboh, Enoch Olad?,
editor of compilation. III. Guasti, Maria Teresa., editor of compilation
P291.L623 2013
415—dc23

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid-free paper

The papers in this collection deal with the concept of locality in syntactic theory, but more specifically they relate to the various contributions that Luigi Rizzi has made in this connection over the past three-and-a-half decades. All the authors are either former students of Luigi's or colleagues and friends who have collaborated with him closely over the years. We, his friends, students, and collaborators, are convinced that Luigi's influence on the development of syntactic theory has been extremely profound.

Very few others have influenced our field as much as Luigi has. And so, we who are fortunate enough to consider ourselves Luigi's friends and collaborators would like to offer him this volume, in recognition of our gratitude.

CONTENTS

CONTRIBUTORS

Enoch Oladé Aboh	University of Amsterdam
Anna Cardinaletti	Università Ca' Foscari Venezia
Guglielmo Cinque	University of Venice
Maria Teresa Guasti	University di Milano-Bicocca
Liliane Haegeman	Ghent University/FWO
Virginia Hill	University of New Brunswick
Richard S. Kayne	New York University
Anne Mette Nyvad	Institute for Aesthetics and Communication, Aarhus University
Jean-Yves Pollock	Paris Est Marne-la-vallée
Ken Ramshøj Christensen	Institute for Aesthetics and Communication, Aarhus University
Ian Roberts	University of Cambridge
Ur Shlonsky	Université de Genève
Tal Siloni	Tel Aviv University
Dominique Sportiche	University of California, Los Angeles
Sten Vikner	Institute for Aesthetics and Communication, Aarhus University

Locality

CHAPTER 1

Locality: An Introduction

ENOCH O. ABOH, MARIA TERESA GUASTI, AND IAN ROBERTS

The papers in this collection all deal with the concept of locality in syntactic theory and more specifically relate to the various contributions Luigi Rizzi has made in this connection over the past three and a half decades. The authors are all either former students of Rizzi's or colleagues and friends who have collaborated with him closely over the years. Luigi's influence on all our work, and on the development of syntactic theory as a whole, has been profound; this volume is a small attempt to recognise and show our gratitude for that influence.

In this brief introduction we will try to set the chapters in this book, and thus indirectly the nature of Rizzi's influence on the field, in context. Accordingly, §1 deals with the notion of locality in general and how this has developed over the past five decades. In §2, we focus more directly on Rizzi's contributions. Finally, §3 summarises the papers collected here.

1. LOCALITY IN GENERATIVE GRAMMAR

Arguably, some notion of locality is conceptually necessary in generative syntax. If syntax relates sound and meaning over an infinite domain, syntactic dependencies and operations must be restricted in such a way as to apply over limited, finite domains in order to be detectable at all (although of course they may be allowed to iterate indefinitely). The theory of what these finite domains are and how they relate to the fundamentally unbounded nature of syntax is the theory of locality.

The notion of locality also relates to explanatory adequacy as originally conceived in Chomsky (1964). The goal of generative grammar is the construction of Universal Grammar (UG), a theory that will explain the human faculty of language. A major aspect of such a theory will be its restrictiveness: the more limited the possibilities of the grammars of natural languages, the smaller the number of hypotheses about the grammar of his or her language the acquirer has to entertain and so the easier the task of language acquisition. One of the main concerns in generative grammar is thus a concern for the restrictiveness of the rules and principles which constitute the mechanisms of grammar. It is in this connection that the study of locality has been important.

The empirical domain in which the question of locality is most readily apparent involves a subclass of movement relations, known as A'-movement. A fundamentally important property of A'-movement dependencies is the fact that they appear to operate over an indefinitely large structural domain: that is, they appear to be unbounded. *Wh*-question formation, one instance of A'-movement, is such an operation. The apparently unbounded nature of this operation is illustrated in (1); examples are from the 1986 edition of Ross's (1967) Massachusetts Institute of Technology (MIT) PhD dissertation and are cited from the published version (copies of moved elements are in parentheses):

(1) a. What did Bill buy (what)?
b. What did you force Bill to buy (what)?
c. What did Harry say you had forced Bill to buy (what)?
d. What was it obvious that Harry said you had forced Bill to buy (what)?

(Ross 1986: 5)

In these examples, the *wh*-expression appears to be fronted over an indefinitely large amount of material.

However, *wh*-expressions cannot be fronted over just any sequence of material, as (2) illustrates:

(2) a. *What did that Bill wore (what) surprise everyone?
b. *What did John fall asleep and Bill wear (what)?

(Ross 1986: 6)

The enterprise of accounting for facts like those illustrated in (1) and (2)—that is, the construction of a theory of unbounded dependencies—is central to the concerns of generative grammar for three reasons.

First, we require the theory to be descriptively adequate; that is, it must account for the facts of English and other languages and any

typological generalisation which can be observed. Therefore, operations like *wh*-movement have to be constrained, and the formulation of the constraints is in part an empirical matter.

Second, as Ross (1986: 6) points out, whatever constraints are formulated are likely to hold more widely than just in English:

> The constraints on variables which I will propose are often of such a complex nature that to state them as constraints on rules in particular languages would greatly increase the power of transformational rules . . . So, from my investigations of the few languages I am familiar with, I will tentatively assume that the constraints I have arrived at are universal.

Another reason to think that the theory of unbounded dependencies is directly connected to UG comes from the nature of the data. Given the rather exotic nature of the data relevant to the formulation of constraints on variables, it is implausible that these constraints are acquired on the basis of primary linguistic data. Therefore, they must be innate, be part of the language faculty, or be derived from third-factor considerations of computational optimisation of the kind adumbrated in Chomsky (2005).

Third and most important, the existence of a class of constraints on variables clearly reduces the class of possible languages. The constraints therefore introduce an element of greater restrictiveness into the theory of UG. As we said already, this is a desirable step toward the overall goal of explaining knowledge and acquisition of language, since language acquirers thus have fewer hypotheses to consider in the process of grammar construction.

Although Chomsky (1964) contained some discussion of this issue and proposed the A-Over-A Principle and an early version of the *wh*-island constraint, Ross (1967) represents the first major attempt to formulate a system of locality constraints on rules, which became known as island constraints.

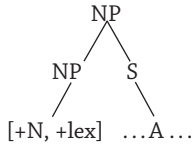
An island is a piece of structure out of which movement is impossible. The first island constraint discussed by Ross is the Complex NP Constraint (CNPC). This is stated as follows:

(3) No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

(Ross 1986: 76)

This means that A cannot be extracted in (4) (Ross 1986: 77):

(4)



The CNPC accounts for two main classes of facts:

(i) the impossibility of extraction from relatives, as in (5):

(5) *Which writer did you write [$_{NP}$ a play which [$_{S}$ was about (which writer)]]?

(ii) the impossibility of extraction from sentential complements to nouns like *claim, fact, story*, etc., as in:

(6) *Which writer did you believe [$_{NP}$ the claim that [$_{S}$ we had met (which writer)]]?

The second island constraint discussed by Ross is the Coordinate Structure Constraint (CSC), stated as follows:

(7) In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(Ross 1986: 99)

This constraint rules out the following kinds of example:

(8) a. *The lute which Henry plays (which) and sings madrigals is warped.
b. *The madrigals which Henry plays the lute and sings (which) sound lousy.

(Ross 1986: 98)

Three further island constraints were formulated by Ross: the Left Branch Condition (LBC), the Sentential Subject Constraint (SSC), and the Right Roof Constraint (although the last of these was not named as such by Ross).

The LBC is stated as follows:

(9) No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.

(Ross 1986: 127)

The LBC accounts for the following contrasts:

(10) a. The boy whose guardian's employer we elected (whose guardian's employer) president ratted on us.
b. *The boy whose guardian's we elected (whose guardian's) employer president ratted on us.
c. *The boy whose we elected (whose) guardian's employer president ratted on us.

The SSC, one of whose effects was noted by Chomsky (1964: 46), is formulated as follows:

(11) No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S.

(Ross 1986: 149)

This accounts for the following contrast:

(12) a. The teacher who the reporters expected that the principal would fire (who) is a crusty old fuzzlebotch.
b. *The teacher who that the principal would fire (who) was expected by the reporters is a crusty old fuzzlebotch.
c. The teacher who it was expected by the reporters that the principal would fire (who) is a crusty old fuzzlebotch.

(Ross 1986: 148)

The constraint later known as the Right Roof Constraint is introduced in terms of the notion of upward boundedness. A rule is upward-bounded if it cannot move elements over the first S-node dominating the base position of the element to be moved. Thus, extraposition, for example, cannot move an element out of the sentence dominating it. This is illustrated in (13):

(13) *A proof that the claim (that John had lied) has been made was given that John had lied.

Ross (1986: 179) generalises the upward-boundedness constraint on extraposition to all cases of rightward-movement (notably Heavy NP-Shift), as follows:

(14) Any rule whose structural index is of the form . . . A Y, and whose structural change specifies that A is to be adjoined to the right of Y, is upward-bounded.

This concept of upward-boundedness, and its alleged nonapplication to leftward movement, is relevant both to Chomsky (1973) and to typological work on *wh*-movement.

The *wh*-island constraint was proposed, without being named as such, in Chomsky (1964: 37ff.). Ross (1986: 19) cites Chomsky's discussion of this constraint and goes on to point out certain examples which suggest it is too strong. The examples given by Ross are of the following type:

(15) He told me about a book which I can't figure out . . .
. . . whether to buy or not.
. . . how to read.
. . . where to obtain.
. . . what to do about.

Ross also notes that infinitival *wh*-complements like those in (15) are better than finite ones (compare, e.g., ?**He told me about a book which I can't figure out whether I should buy or not*, with (15)). Since Ross does not deny the ungrammaticality of the examples given by Chomsky to motivate the constraint (although it should be pointed out that an example like **What did you wonder where John put?* also involves a crossed dependency, increasing the ungrammaticality), the discussion is inconclusive.

Ross's island constraints prevented transformational rules from applying in certain contexts. The logical next step in research on these phenomena was to attempt to characterise what the various island constraints have in common. In this way, it was hoped that an intensional characterisation of the notion of island could be arrived at, in place of an extensional list.

The central idea in this line of research was subjacency, introduced in Chomsky (1973). In order to see how subjacency works, two assumptions must be spelled out. One concerns the structure of clauses and the other the operation of transformations.

First, it has been assumed since Bresnan (1970, 1972) that all sentences are preceded by a special complementiser position, COMP (or C); since Chomsky (1986) it has been assumed that C is the head of a phrasal category CP, taking TP as its complement. In subordinate clauses, this position is filled by the subordinating conjunction; in matrix clauses it is often empty, but in *wh*-questions it (or its Specifier, if C is taken to be a head) is filled by the fronted *wh*-element. COMP and S (or TP in more recent terminology), the clause introduced by the complementiser, form a larger constituent S' (CP). So the rules that we have been referring to as *wh*-question formation and relative-clause formation both involve fronting a *wh*-element to COMP (until Chomsky 1986 *wh*-movement was seen as substitution of the *wh*-phrase into COMP). Because of this essential feature shared by the two rules, the rules were collapsed as *wh*-movement.[1]

Second, the assumption that *wh*-movement operated in an unbounded manner was abandoned. This might seem strange, given the data in (1). However, it was proposed instead that *wh*-movement operates successive-cyclically, moving a *wh*-word through a series of C-positions in the generation of a sentence like (1c), repeated here (here we indicate the positions from which *what* has moved as containing copies of *what*):

(1c) [$_{S'1}$ What did [$_{S1}$ Harry say [$_{S'2}$ (what) [$_{S2}$ you had forced Bill [$_{S'3}$ (what) [$_{S3}$ to buy (what)]]]]]]?

Here *what* moves from its base position, the direct-object position of *buy*, first to the COMP of S'$_3$ and then to that of S'$_2$ and then to the matrix COMP, S'$_1$. Two conditions were imposed on this process of successive-cyclic movement:

(16) a. *COMP-to-COMP Condition*: a phrase in COMP can only move to a higher COMP.
b. *Strict Cycle Condition*: No rule can apply to a domain dominated by a cyclic node *A* in such a way as to affect solely a proper subdomain *A* dominated by a node *B* which is also a cyclic node.

(Chomsky 1973:243)

(The cyclic nodes are those that determine the domain of operation of the transformational cycle, NP and S). Condition (16a) forces successive COMP-to-COMP movement rather than COMP-to-subject movement, for example. Condition (16b) tightens the earlier notion of cyclic application of rules, whereby the ordered transformational rules were thought to apply *en bloc* to the lowest S and then, on the next cycle, to the next S up and so on, so that there would be no possibility of a rule on a higher cycle accidentally applying on a lower cycle.

Subjacency can be stated as follows (this is a simplified version of the formulation in Chomsky (1973: 247f 271ff.):

(17) No rule can relate X and Y in the following structure:
$\ldots X \ldots [_B \ldots [_B \ldots Y \ldots]$
where X is separated from Y by more than one bounding node B.

Chomsky proposed that bounding nodes were all and only cyclic nodes, that is, S and NP (these correspond to TP and DP under current assumptions).

Consider now the derivation of a sentence involving extraction from a relative clause:

(18) *Who did Mary read the book which we gave to?

Moving *which* on the lower cycle gives rise to an intermediate structure like (19):

(19) Mary read [$_{NP}$ [$_{NP}$ the book] [$_{S'}$ which [$_{S}$ we gave to who]]]

The only way to derive (18) from (19) is to move *who* up to the matrix COMP position in one step. However, such a movement violates subjacency, as two S-nodes separate the base position of *who* from the matrix COMP, as can be seen in (19). The Strict Cycle is violated if *who* moves first to the lower COMP and on to the matrix one, followed by movement of *which* to the lower COMP. Thus, the ungrammaticality of (18) is explained by subjacency along with the assumptions in (16).

Subjacency operates in a similar way to account for the complement cases of the CNPC and of the *wh*-island constraint. Relevant examples are given in (20):

(20) a. *Which race [$_{S}$ did you hear [$_{NP}$ the announcement [$_{S'}$ that [$_{S}$ John won (which race)]]]]?
b. *Who did [$_{S}$ you wonder [$_{S'}$ which books [$_{S}$ to give (which books) to (who)]]]?

In both of these examples, subjacency is violated. In (20a), *which race* moves from the lower COMP to the matrix one, crossing an NP-node and an S-node. In (20b), *who* moves to the matrix COMP directly from its base position over the filled lower COMP (the Strict Cycle prevents it from moving to the matrix COMP first, followed by the shorter movement to the lower COMP on the matrix cycle, and the assumption that there is only COMP position blocks successive cyclic movement of *who*). The one-step movement of *who* to the matrix COMP crosses two S-nodes and so violates subjacency. Subjacency thus appears able to unify the CNPC and the *wh*-island constraint.

The SSC can also be derived from subjacency. Schematically, the SSC prevents extraction of the lower NP in the following configuration:

(21) [$_{S}$ [$_{NP}$ [$_{S}$ . . . NP . . .] . . .] . . .]

Clearly, any such extraction will cross at least one S-node and an NP-node.

The SSC was generalised as the Subject Condition in Chomsky (1973: 250). This condition prevents extraction of any part of a subject NP, whether that subject is sentential or not. Relevant examples are as follows:

(22) a. *Who did stories about (who) amuse John?
b. *Who did your interest in (who) annoy Bill?

Again, it is clear that the *wh*-element moves across an NP-node and an S-node here, in violation of subjacency.

It is also possible to derive the general constraint that all rightward rules are upward-bounded, in Ross's sense, from the Subjacency Condition

combined with the fact that COMP appears only on the left of S. In that case, there is no possibility of successive-cyclic rightward movement, so elements can move rightward only as far as the first bounding node up, that is, the first S up. Any further movement violates subjacency. But in fact there is no real asymmetry between leftward and rightward movement: all rules are upward-bounded, but leftward-movement has the possibility of moving through the COMP escape hatch.

Subjacency can also take care of the Left Branch Condition. An NP immediately dominated by another NP must cross that barrier and the next S-node up in order to get to the nearest COMP. It will thus necessarily cross two bounding nodes and therefore violate subjacency.

On the other hand, subjacency can derive only certain cases of the CSC. Assuming that conjoined categories form a larger category of the same kind as those conjoined (i.e., that *John and Mary* is an NP (or DP) just like *John* and *Mary*), then extraction of a conjoined NP will violate the Subjacency Condition, as would extraction of an NP out of a conjoined S. However, subjacency cannot directly handle extraction of an NP from a coordinate VP.

In general, then, subjacency went some way toward unifying Ross's island constraints. The account relies on the idea that the bounding nodes are NP and S. It is clear that the logical next step is to try to see if we can give an intensional definition of the bounding nodes. This brings us to the *Barriers* theory of Chomsky (1986).

The main question addressed in *Barriers*, in the context of the Government-Binding (GB) theory, is that of the relation between government and the characterization of the bounding nodes for subjacency. One aspect of this question corresponds to the point raised previously: can we arrive at an intensional definition of bounding nodes? Here we concentrate on this aspect of the proposals in Chomsky (1986) at the expense of others (notably the question of the relationship between the Empty Category Principle (ECP) and subjacency; for more on the ECP, see §2).

One consideration that led to the definition of barrier given in Chomsky (1986) was the discovery of a new class of islands, the adjunct islands. Huang (1982) noted that extraction of a part of an adjunct was impossible:

(23) *Which warning did you leave [despite (which warning)]?

Huang combined this with the Subject Condition, briefly mentioned in the previous section, in his Condition on Extraction Domains (CED), which prevents extraction from non-complements, that is, subjects and adjuncts.

Chomsky (1986: 14) first defines barriers in terms of Blocking Categories (BC):

(24) X is a BC for Y iff X is not L-marked and X dominates Y.

(25) X is a barrier for Y iff (a) or (b):
a. X immediately dominates Z, Z a BC for Y;
b. X is a BC for Y, X ≠ IP.

(IP, or Inflection Phrase, corresponds to the earlier S and the later TP). In order to understand how these definitions work, we need a definition of L-marking (Chomsky 1986: 15):

(26) X L-marks Y iff X is a lexical category that θ-governs Y.

θ-government is the relation between a lexical head and its sister. The basic form of subjacency remains the same, in that movement across two barriers is not permitted (Chomsky 1986: 30).

We now briefly show how this system derives the same results as the earlier subjacency theory just outlined as well as a few more.

Consider first the Subject Condition. Subjects of finite clauses are not directly θ-marked, since they are not complements. Therefore, they are not L-marked. Therefore, by (24), a subject NP in a finite clause is a BC for anything contained in it and therefore is a barrier, by (25b). Moreover, the IP immediately dominating the subject is a barrier, by (25a). So any element moved out of a subject NP into the nearest Specifier of CP (SpecCP) position will violate subjacency, as two barriers will be crossed. This derives the Subject Condition.

The Adjunct Condition, relevant in (23), is derived in a precisely analogous way. Adjuncts are not complements, so they are not L-marked. Since they are not L-marked they are BCs and therefore barriers for extraction of material from inside them. Moreover, the IP immediately dominating an adjunct is also a barrier, given (25a). So extraction from adjuncts is impossible. This is a good result because it is clear that various categories serve as adjuncts and thus create islands, so the earlier approach of simply listing which categories are barriers could not work here. In particular, consider the following contrasts (see Huang 1982):

(27) a. *Who did you meet John [$_{AP}$ angry at (who)]?
b. Who did you make John [$_{AP}$ mad at (who)]?

(28) a. *Which city did you meet a man [$_{PP}$ from (which city)]?
b. Which city did you see the destruction [$_{PP}$ of (which city)]?

In each example, extraction is sensitive to whether the AP or PP is an argument or not: the category itself seems to be irrelevant. These contrasts can be captured in terms of the *Barriers* framework but could not be in the framework of Chomsky (1973).

Consider next the relative-clause case of the CNPC, repeated here (with S and S' changed to IP and CP, respectively):

(19) Mary read [$_{NP}$ [$_{NP}$ the book] [$_{CP}$ which [$_{IP}$ we gave to who]]]

Relative clauses are not arguments of the NPs they modify. Because of this, CP in (19) is not L-marked, so it is a BC and a barrier, and the NP immediately dominating it is therefore also a barrier for material extracted from inside CP. As a result, *what* cannot move to the matrix SpecCP in one step without violating subjacency.

The complement case of the CNPC is more problematic, precisely because the CP complement to the head noun of the complex NP, being a complement, is L-marked and therefore is neither a BC nor a barrier. Neither is NP a barrier. Chomsky (1986: 36) suggests that the CP complement of N may be an inherent barrier. It may thus be that only one barrier is crossed, leading to a 'weaker' violation.

Wh-island violations work largely as in the earlier approach. The presence of one *wh*-element in an embedded SpecCP forces the other *wh*-element to cross at least one barrier, namely, CP (which inherits barrierhood for material extracted from within IP from the non–L-marked node IP). Since IP is a BC by (24) but not a barrier, given (25b), it may be that only one barrier is crossed here, leading to a weaker violation. Given the examples in (15), this may be the correct conclusion. We return to the distinction between 'weak' and 'strong' islands in the next section.

Here we have sketched the development of the theory of locality of A'-movement from Chomsky (1964), through Ross (1967) to Chomsky (1973), and, finally, Chomsky (1986). Rizzi's contributions stem from the late 1970s, as we will now see.

2. RIZZI'S CONTRIBUTIONS TO THE THEORY OF LOCALITY

Rizzi is primarily responsible for three central innovations in the theory of locality: the observation that the bounding nodes/barriers for subjacency in Italian are different from those of English; the connection between complementiser-trace effects and the null-subject parameter (NSP); and relativised minimality. Each of these contributions has been extremely

influential in the general development of syntactic theory since the late 1970s. We will now look at them in turn.

2.1 Subjacency in Italian and the beginning of parametric theory

One class of island constraints which falls under the standard version of subjacency, as in (17), is the *wh*-island constraint. This is illustrated in (29):

(29) *The only job which you didn't know who they were going to give (which) to (who) has actually been given to you.

This kind of example is ungrammatical for many speakers of English (including one of the current authors), and it involves a violation of subjacency, as we can see from (30):

(30) *The only job [$_{S'1}$ which [$_{S1}$ you didn't know [$_{S'2}$ who [$_{S2}$ they were going to give (which) to (who)]]]]

Here, *who* moves from its base position marked by parentheses to the COMP of the lower clause, so *which* must move in a single step to the COMP of the higher clause. This movement crosses two bounding nodes, S_2 and S_1, so subjacency is violated.

Rizzi's key observation in his 1982 paper (originally written in 1977; see Rizzi 1982:xii) was that the analog of (29) is grammatical in Italian:

(31) Il solo incarico che non sapevi a chi avrebbero affidato è poi finito proprio a te.
The only task that not knew-2sg to whom would-have-3pl entrusted is then finished exactly with you.
'The only task that you didn't know to whom they would entrust has been entrusted exactly to you.'
(Rizzi 1982: 50)

Rizzi considers and rejects, on empirical grounds, the possibility that Italian relaxes the Strict Cycle Condition (see (16b)) or the ban on moving through a COMP already filled by a *wh*-constituent (cf. the doubly filled COMP filter of Chomsky and Lasnik 1977). He also provides evidence, including violations of the CNPC, that NP is a bounding node in Italian.

The crucial grammaticality contrasts in the paper involve *wh*-movement from an indirect question embedded inside a declarative as opposed to a declarative inside an interrogative. Schematically, the relevant parts of the two structures are given in (32) (see Rizzi 1982: 55):

(32) a. [$_{S'3}$ $COMP_3$ … [$_{S'2}$ $COMP_2$[-wh] … [$_{S'1}$ $COMP_1$[+wh] wh_{rel} wh_Q]]]
b. [$_{S'3}$ $COMP_3$ … [$_{S'2}$ $COMP_2$[+wh] … [$_{S'1}$ $COMP_1$[-wh] wh_{rel} wh_Q]]]

In both structures, wh_Q is attracted to the [+wh] COMP, that is, $COMP_1$ in (32a) and $COMP_2$ in (32b), while wh_{rel} is attracted to $COMP_3$ (the whole structure being a relative clause). In (32a), wh_Q moves to $COMP_1$ and wh_{rel} to $COMP_2$ and on to $COMP_3$; in (32b), wh_Q moves to $COMP_2$ and wh_{rel} to $COMP_1$ and then on to $COMP_3$, giving rise to ungrammaticality. Examples (33a, 33b) instantiate the schemata in (32a, 33b), respectively (Rizzi's 1982: 56, adapted):

(33) a. La macchina [$_{S'3}$ che credo [$_{S'2}$ che Gianni si domandi [$_{S'1}$ se potrà
The car that I-think that Gianni self ask if he-can-FUT
utilizzare nel weekend]]] è la mia.
use in-the weekend is the mine
'The car that I think Gianni wonders whether he will be allowed to use during the weekend is mine.'

b. *La macchina [$_{S'3}$ che mi domando [$_{S'2}$ se Mario creda [$_{S'1}$ che potrà]
the car that me I-ask if Mario believe that he-can-FUT
utilizzare nel weekend è la mia.
use in-the weekend is the mine

Rizzi argues that these contrasts can be explained only if we assume that the Strict Cycle, the doubly filled COMP filter and subjacency, hold in Italian just as they do in English but, crucially, that the bounding nodes for subjacency in Italian differ from those of English in that in Italian they are S' and NP. Thus, in the Italian example in (31) 'long' *wh*-movement to the COMP of S'_1 is allowed since it crosses just one bounding node, S'_2, while the corresponding movement in the English, seen in (29) and (30), crosses two bounding nodes S_2 and S_1.

In addition to clarifying the status of apparent *wh*-island violations in Italian in an elegant and parsimonious fashion, this analysis was the first application of the idea of principles and parameters. Chomsky (1976) had first sketched out the idea of a parameter of UG, but it is in Rizzi's paper that the idea is applied for the first time. 'English and Italian differ in the choice of the bounding nodes which count for subjacency in that such nodes are NP and S for English, and NP and S' for Italian' (Rizzi 1982: 73, n. 25). In other words, subjacency is a principle of UG, with the options for (some) bounding nodes left open. English and Italian select differing options, with the result that *wh*-movement operates rather differently in the two languages giving rise to surface differences in certain *wh*-island constructions as observed.

To see the full importance of this idea, we need to consider Chomsky's (1964: 28ff.) definitions of levels of adequacy for linguistic theory. These were observational, descriptive, and explanatory adequacy. An observationally adequate grammar presents the data correctly, while a descriptively

adequate grammar 'specifies the observed data . . . in terms of significant generalizations that express underlying regularities in the language' (Chomsky 1964: 28). Explanatory adequacy 'can be interpreted as asserting that data of the observed kind will enable a speaker whose intrinsic capacities are as represented in th[e] general theory to construct for himself a grammar that characterizes exactly this intuition' (Chomsky 1964: 28); in other words, attaining explanatory adequacy involves showing how a given empirical phenomenon can be deduced from UG.

The postulation of parametric variation in UG principles was a very large step in the direction of explanatory adequacy, since, one could assume, if we can say that *this* syntactic feature of *this* language is due to setting *that* parameter to *that* value we have provided an explanatorily adequate account of the syntactic feature in question in that we have related it directly to UG as well as a descriptively adequate account to the extent that the analysis of the relevant property of the language is correct. Moreover, the parametric account has immediate cross-linguistic implications, since it implies that another language lacking the property in question will set the parameter in question to a different value. Now, if each parameter value determines a cluster of disparate syntactic features, then explanatory adequacy is enhanced, especially if certain features are readily accessible to acquirers on the basis of impoverished evidence while others are hardly likely to be easily accessible. In this case, arriving at a parameter value determining both the accessible and relatively inaccessible feature gives us a simple account of how the inaccessible feature can be acquired, thus accounting for an aspect of the poverty of the stimulus to language acquisition and thereby, again, reaching explanatory adequacy. At the same time, other things being equal, a 'typological' prediction is made: the inaccessible feature will be acquired whenever the acquired one is, since both reflect the same abstract property of Universal Grammar, the setting of a given parameter to a given value.

Chomsky has often stated that the move to the principles-and-parameters conception of UG and cross-linguistic variation was a major breakthrough. Rizzi's pioneering work in the Italian–English contrasts seen already was a crucial ingredient in this major conceptual advance.

2.2 Complementiser-trace effects and the null-subject parameter

Another very important strand of Rizzi's work, which again partially implicates locality, concerns the NSP. It has played a prominent role in the theoretical study of comparative syntax in recent years, not just because of the characterization it gives us of languages like Italian and how they differ from languages like English but primarily because it has been seen as a

good example of the way rather abstract grammatical properties may have proliferating effects, unifying apparently unrelated surface phenomena.

Rizzi (1982, Chapter 4), building on earlier work by Perlmutter (1971), proposed a cluster of surface properties determined by the NSP, as follows:

(34) a. The possibility of a silent, referential, definite subject of finite clauses.
b. 'Free subject inversion.'
c. The apparent absence of complementiser-trace effects.
d. Rich agreement inflection on finite verbs.

(34a) refers to the well-known feature of Italian, Spanish, Greek, and many other languages that allows a definite pronominal subject to drop, as in:

(35) Parlano italiano.
Speak-3pl Italian
'They speak Italian.'

Here a common intuition is that the content of the pronoun is expressed or perhaps recovered by, the 'rich' agreement inflection on the finite verb.

(34b) refers to the general possibility of expressing an overt subject, usually with a focus interpretation, in postverbal position:

(36) a. Hanno telefonato molti studenti.
b. *Ont téléphoné beaucoup d'étudiants.
Have telephoned many students.
'Many students have telephoned.'

Free inversion is in fact subject to slightly differing constraints in different languages, being more freely available in Spanish and Greek than in Italian (see, e.g, Sheehan 2006, Chapter 6 and the references given there).

(34c) relates to Perlmutter's generalisation, since it originates in Perlmutter's (1971) pioneering work. Perlmutter's generalisation expresses the fact that in non–null-subject languages the subject of a finite clause cannot undergo *wh*-movement if the complementiser introducing the clause is present. This constraint holds of English and French, as the following examples show:

(37) a. *Who did you say that—wrote this book?
b. *Qui as-tu dit qu'—a écrit ce livre? (=(37a))

Here the questioned constituent (*who/qui*) corresponds to the subject of the subordinate clause, so there is a gap in that position. The ungrammaticality of (37a) is known as a complementiser-trace effect, since in many versions of the theory of movement it is held that the empty subject position at the

movement site in the complement clause contains a trace of the moved *wh*-element. The idea that the presence of the complementiser determines the ungrammaticality of such examples is supported by the fact that (37a) becomes grammatical if *that* is omitted. In French, (37b) can be rendered grammatical by altering the form of the complementiser from *que* to *qui*. These points are illustrated in (38):

(38) a. Who did you say—wrote this book?
b. Qui as-tu dit qui—a écrit ce livre? (=(38a))

In null-subject languages, as Perlmutter observed, it appears that complementiser-trace effects are not found. Rizzi (1982, Chapter 4) showed that in fact this is not true if certain quantificational structures covertly derived at the level of Logical Form (LF) are taken into consideration. The subject of a finite clause introduced by a complementiser can readily be questioned in these languages however:

(39) Chi hai detto che— ha scritto questo libro? (Italian)
Who have-2sg said that—has written this book

This feature of the null-subject cluster can be reasonably thought of as relatively inaccessible in the PLD, while rich agreement inflection is presumably very accessible (especially given the known sensitivity of acquirers to inflections; see Hyams 1986; Guasti 1993–1994; Wexler 1998), and the other two properties may be somewhat accessible.

The principles and parameters approach to UG can take us towards explanatory adequacy, in the sense of Chomsky (1964) as given already. But it is also clear that this approach defines language types. In this way, typology, in the sense of the establishment of cross-linguistic relations and of a structure to cross-linguistic variation, and acquisition become intrinsically related. This is a very positive development as it clearly opens the way to a two-pronged empirical approach to understanding the nature of UG.

In Rizzi's (1982) terms, all the properties associated with the null-subject parameter are connected by the presence of the silent pronoun *pro* in the subject position. This element is licensed by rich agreement inflection and can satisfy the general requirement for a subject position (the Extended Projection Principle of Chomsky 1982: 10), allowing an overt subject to appear in the freely inverted position and indeed to be *wh*-moved from this position. Thus, the formal property which underlies the null-subject parameter, on this analysis, is the availability of *pro* subjects. Once acquirers deduce this (on the basis of the universal principles determining the availability of null pronouns), they will immediately deduce the existence of the other

properties in the cluster, and the implicational links among the properties in (34) follow. Hence, we expect typological correlations to support parametric clusters and thereby to motivate analyses of the general type instantiated by Rizzi's account of the cluster associated with null subjects.

2.3 Relativised minimality

Rizzi's other major contribution to linguistic theory, and indeed to the theory of locality, was the postulation of the notion of relativised minimality, initially in his 1990 monograph (a significantly updated version is presented in Rizzi 2000; see also Starke 2001 for major refinements).

The initial observation goes back to Huang (1982), who observed that there are important differences between arguments of certain types and adjunct elements with regard to extraction. These differences emerge if we compare direct object extraction from a *wh*-island, seen in (40a), with extraction of an adverbial element from the same island, as in (40b):

(40) a. ?*Whose car were [$_{IP}$ you wondering [$_{CP}$ how [$_{IP}$ you should fix (whose car) (how)]]]?
b. *How were [$_{IP}$ you wondering [$_{CP}$ whose car [$_{IP}$ you should fix (whose car) (how)]]]?

The difference between (40a) and (40b) seems to be as follows: while argument-extraction examples as in (40a) are very awkward, they are intelligible; in examples like (40b), on the other hand, it is all but impossible to grasp the correct interpretation (with *how* interpreted as modifying the lower clause, looking for an answer like 'with a spanner' in each case). This distinction has been interpreted as indicating that the constraints on adjunct-extraction involve conditions relating to the level of LF.

There are islands which block only adjunct extraction; these are known as weak islands. Cinque (1991: 1–2) lists negative (or 'inner') islands (first observed by Ross 1986), factive islands, and extraposition islands. These are illustrated in (41)–(43):

(41) a. Which car don't you know how to fix (which car)?
b. *How don't you know to fix your car (how)?

(42) a. Which car do you regret that you fixed (which car)?
b. *How do you regret that I fixed your car (how)?

(43) a. Who is it believed these days that John likes (who)?
b. *How is it believed these days that John fixes his car (how)?

The distinction between arguments and adjuncts was handled in GB theory by the Empty Category Principle (ECP), which required that traces of movement be properly governed. Proper government is a restricted form of the general government relation and has two components: lexical government and antecedent government. Lexical government is government by a lexical head: complements are always lexically governed, while adjuncts never are and subjects usually are not. In terms of lexical government it is possible to capture many aspects of the argument–adjunct distinction, leaving aside certain complexities involving subjects (including notably complementiser-trace phenomena, which played an important role in work on the null-subject parameter as we saw in the previous section).

The ECP was seen as a condition on representations, holding of LF representations as well as S-structure ones (in fact, analyses differed on this last point). The idea that the ECP was, at least in part, an LF condition was supported by the extensive work on the interpretation of *wh*-elements (or their equivalents) in languages lacking overt *wh*-movement, such as Chinese, beginning with Huang (1982). The fact that Chinese lacks overt *wh*-movement but has some process that allows *wh*-elements to be interpreted in a fashion comparable to that of English and other languages showing overt *wh*-movement can be seen from the interpretations of examples like (44):

(44) a. Zhangsan yiwei Lisi mai-le shenme?
Zhangsan think Lisi bought what
'What does Zhangsan think Lisi bought?'

b. Zhangsan xiang-zhidao Lisi mai-le shenme
Zhangsan wonder Lisi bought what
'Zhangsan wonders what Lisi bought?'

In both examples in (44) the *wh*-expression *shenme* is unmoved; it occupies the normal direct object position in Chinese (henceforth I refer to unmoved *wh*-phrases as *wh-in-situ*). The selectional properties of the main verbs force different interpretations on these two occurrences of *shenme*. In (44a), *shenme* cannot be interpreted as having scope only over the embedded clause, so it must be interpreted as a matrix question, as indicated in the translation. In (44b), the main predicate requires an interrogative complement, so the scope of *shenme* is restricted to the subordinate clause. Huang suggested that the differences in scope and sensitivity to selectional properties were best accounted for by assuming covert *wh*-movement in the derivation of LF from S-structure.

Huang supported this idea by showing that, although many island effects are not found in *wh-in-situ* languages such as Chinese, other locality effects

associated with *wh*-movement are found in these languages. In particular, adjunct *wh*-elements cannot be interpreted with wide scope in certain islands. The following example illustrates this for the adjunct *weishenme* ('why') in a complex NP in Chinese:

(45)	ni	zui	xihuan	[weishenme	mai	shu	de	ren] ?
	You	most	like	why	buy	book	Prt	person

'Why do you like the man who bought the books?'

Here *weishenme* cannot be interpreted as modifying the predicate inside the relative clause, that is, as asking why the man bought the books. Facts of this kind show that, while overt movement is sensitive to islands, further locality constraints on *wh*-interpretation which are independent of overt movement. Huang influentially proposed that these effects are due to the application of the ECP to traces of covertly moved *wh*-phrases at LF, while the lack of standard island effects is due to the lack of overt *wh*-movement and the idea that subjacency applies only to overt *wh*-movement.

Examples like (45) were analysed and connected to examples like (40b) in languages with overt movement in terms of the ECP. Adjunct traces are never lexically governed, and in examples of the type in question it was argued that they failed to be antecedent governed. Various analyses of antecedent government were proposed. (Different proposals, with differing and overlapping empirical coverage, were put forward in, e.g., Kayne 1981, 1983; Pesetsky 1982; Lasnik and Saito 1984, 1992; Chomsky 1986.) Rizzi's (1990) system of relativised minimality arguably provided the most elegant and satisfactory analysis and has proven extremely influential ever since.

The central idea of relativised minimality is that antecedent government of β by α is blocked by the presence of an intervening element γ where γ is the same type of element as α. Clearly, the two notions of *intervention* and *being of the same type* must be defined. Intervention is defined in terms of asymmetric c-command: γ intervenes between α and β if γ asymmetrically c-commands β and α asymmetrically c-commands both β and γ.[2] We return to the definition of types of element later; for the moment we need to assume only that distinct *wh*-phrases are of the same type as one another.

To see how relativised minimality works, consider again an example like (40b) (here we use trace notation to illustrate the workings of the ECP as a condition on traces):

(31b) *How$_j$ were [$_{IP}$ you wondering [$_{CP}$ whose car$_i$ [$_{IP}$ you should fix t$_i$ t$_j$]]]?

The trace of *whose car*, t_i, satisfies the ECP by being lexically governed by the verb *fix*. The trace of *how* fails the ECP by failing to be lexically governed

(since it is an adjunct) and failing to be antecedent governed owing to the presence in the lower SpecCP of an intervening element, *whose car*, which is of the same type as *how*. Antecedent government of the trace of *how* is thus blocked, the ECP is thereby violated, and the sentence is ungrammatical.

Rizzi (1990) defines the notion of *structural type* that is relevant for the creation of an intervention effect in terms of a three-way distinction between A'-positions, A-positions, and head-positions. A'-positions are specifier positions which do not bear a grammatical relation, while A-positions are specifier positions which do bear a grammatical relation (subject positions in nearly all the relevant cases). Hence, in (40b), *whose car*, being in an A'-position which asymmetrically c-commands t_j and is asymmetrically c-commanded by how_j, acts as an intervener for the relation between *how*, also in an A'-position, and t_j. Other cases where an element in an A'-position acts as an intervener for a *wh*-trace are negative islands as in (41), where the intervener is *not*, and 'pseudo-opacity' effects in French, where the intervener is a particular kind of quantificational adverb:

(41b) *How $_i$ did you not fix your car t_i?

(46) a. Combien de livres$_i$ a-t-il beaucoup consultés t_i?
How-many of books has-he very-much consulted
'How many books has he consulted a lot?'

b. Combien$_i$ a-t-il beaucoup consultés t_i de livres?
How-many has-he very-much consulted of books

In (41b), *not* occupies an A'-position. In (46), *beaucoup* similarly occupies an A'-position. (Rizzi 2000 extends the range of structural types to make distinctions amongst a range of adverb types.) In both cases the intervener asymmetrically c-commands the trace and is asymmetrically c-commanded by the antecedent. The contrast between movement of *combien de livres* and *combien* alone is a further case of an argument–adjunct asymmetry. Examples (40b), (41b), and (46) thus fall under a single generalisation.

Cases where an element in an A-specifier blocks A-movement include superraising, as in (47a), and that where a head blocks head movement underlies the head movement constraint:

(47) a. *John$_i$ seems that it is likely t_i to be late.
b. *Have$_i$ he could t_i done it?

In (47a) *it* is the intervener: it is in an A-position, like the antecedent *John*; it asymmetrically c-commands t_i; and it is asymmetrically c-commanded

by *John*. In (47b), *could* is the intervener: it is in a head position, like the antecedent *have*; it asymmetrically c-commands t_i; and it is asymmetrically c-commanded by *have*. Again, (47) falls under the same generalisation as the examples of adjunct *wh*-movement discussed already.

The notion of selective intervention, according to which an element is an intervener depending on the nature of the antecedent or the nature of the trace (or copy) of movement is maintained in Chomsky's (1995) Minimal Link Condition and generally in conceptions of locality in the minimalist program based on the notion of shortest movement or closest attractor (see, e.g., the non-intervention clause in the definition of Agree in Chomsky 2000, 2001). In fact, we can give a general definition of *intervener*, as follows:

(48) α is an intervener for a relation R(β, γ) iff [(R(β, α) & R(α, γ) & $\neg$R(α, β)]

(48) says that α 'breaks up' a relation of some kind between γ and β if α has that same relation with γ and β has it with α but α does not have that relation with β (clearly the relation question must be asymmetrical). This kind of notion is naturally compatible with the central minimalist idea that derivational or representational economy is a central part of the theory of syntax, where they are often connected to a general optimisation strategy of minimal search.

2.4 Conclusion

To conclude, we see from the previous sections very brief summaries of an impressive range of highly technically, empirically, and theoretically sophisticated work that Rizzi has made a series of seminal contributions to syntactic theory. He was the first to implement the notion of parameter of UG, in his work on *wh*-islands in English and Italian, a pioneer in further developing the approach in his work on the null-subject parameter and in particular the observation regarding extraction from postverbal position in complement clauses, and, finally, the originator of relativised minimality, which represents a profound generalisation about the nature of the human language faculty.

The chapters that follow attempt, in their various ways, to do justice to this impressive legacy.

3. THE CHAPTERS IN THIS VOLUME

In their chapter, Kayne and Pollock analyse inversion structures in French interrogatives such as *Cela la gêne-t-il*? ('that her bothers it' = 'does that